

DISS. ETH NO 23789

Quantitative Genetics of Larval Starvation Resistance in *Drosophila melanogaster*

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

Izzet Mehmet Akcay

M.Sc., Sabanci University
B.Sc., Middle East Technical University

born on 13.06.1985
citizen of Turkey

accepted on the recommendation of

Prof. Dr. Ernst Hafen
Prof. Dr. Matthias Peter
Prof. Dr. Trudy FC Mackay
Dr. Katja Koehler

2017

TABLE OF CONTENTS

1. SUMMARY	4
2. ZUSAMMENFASSUNG	6
3. INTRODUCTION	8
3.1. Starvation resistance in the animal kingdom	9
3.1.1. Physiological and behavioral adaptations to starvation	9
3.1.2. Metabolic adaptations to starvation	11
3.2. Starvation resistance in <i>Drosophila</i>	12
3.2.1. Genetic variation for starvation resistance	13
3.2.2. Environmental plasticity of starvation resistance	14
3.2.3. Physiological mechanisms underlying starvation resistance	14
3.2.4. Correlations of starvation resistance with other stress resistance traits and life history traits	18
3.3. <i>Drosophila</i> larval stage is characterized by extensive feeding and growth	19
3.4. Molecular mechanisms underlying metabolic homeostasis in <i>Drosophila</i>	22
3.5. Functional genomics studies identify many genes involved in starvation responses	24
3.6. Genome-wide association studies	27
3.6.1. Background on quantitative trait genetics	27
3.6.2. Basic concepts and methods used in GWAS	29
3.6.3. Next steps after a GWAS	33
3.7. GWAS in model organisms	34
3.8. The <i>Drosophila</i> Genetic Reference Panel	35
3.9. Summary and project motivation	37
3.10. Aims of the thesis	38
4. RESULTS	40
4.1. Optimizing the conditions for phenotyping larval starvation resistances	40
4.1.1. Starvation resistances during early larval stages	41
4.1.2. Screening the DGRP lines	42
4.2. Quantitative genetics analysis	43
4.3. Genome-wide association mapping	46
4.4. SNP-based validation	50
4.5. Sexual dimorphism in larval starvation resistance	52
4.6. Correlations with other traits	54
4.7. Micro-environmental plasticity of larval starvation resistance	56
4.8. Supplementary results	59
4.8.1. Sex ratio distortion as a minor confounder for phenotyping pooled sexes	59
4.8.2. Bimodal death curves are not due to sexual dimorphism	60
4.8.3. Sexual dimorphism in adult starvation resistance	61

4.8.4. Q-Q plot for GWAS on inversion-corrected mean larval starvation resistances.....	62
5. DISCUSSION	65
5.1. Phenotypic and genotypic variation in larval starvation resistance.....	65
5.2. Starvation resistance during early larval stages.....	67
5.3. Sexual dimorphism in larval starvation resistance.....	68
5.4. Genetic correlations with other traits in the DGRP	69
5.5. GWAS identifies many loci and genes involved in larval starvation resistance.....	71
5.6. Epistatic interactions between loci.....	73
5.7. Functional validation of associations.....	74
5.8. Micro-environmental plasticity for larval starvation resistance	76
5.9. Insights into molecular and cellular mechanisms underlying larval starvation resistance	78
6. CONCLUSION AND OUTLOOK.....	85
6.1. Conclusion.....	85
6.2. Outlook.....	86
6.2.1. Functional characterization of associated genes using classical genetics	86
6.2.2. Systems genetics.....	87
6.2.3. Linkage mapping analysis	88
6.2.4. Experimental evolution.....	89
6.2.5. Association mapping in other populations	90
7. MATERIALS AND METHODS.....	91
8. SUPPLEMENTARY DATA	95
9. REFERENCES.....	132
10. ACKNOWLEDGEMENTS.....	152
11. CURRICULUM VITAE.....	153

1. SUMMARY

Natural populations are very often confronted with reductions in nutrient levels. The physiology underlying the adaptation to starvation is of interest for evolutionary studies and has impacts on human health. Animals normally supply their energy demands by metabolizing their food, but when the access to food is limited, they need to use their internal energy reserves. The ability to tolerate prolonged periods of scarce food conditions, which is termed starvation resistance, is determined by the balance between energy supplies and expenditure.

Model organisms, such as *Drosophila melanogaster*, provide opportunities to study the physiological responses to starvation and determine the underlying molecular mechanisms. Although during evolution, the arthropod lineage separated from the vertebrate lineage more than 600 million years ago, genetic, molecular and behavioral analyses suggest that *Drosophila* and humans are not that unrelated. Studies using *Drosophila* have shown that starvation resistance is correlated positively with nutrient reserves, lifespan and other stress resistance traits, and negatively with fertility. These correlations are evolutionarily derived, and have a genetic basis.

Classical genetic screens and candidate-gene-based studies have identified various genes involved in starvation resistance; however, many components of the molecular mechanisms functioning in starvation resistance, and the interactions of these components with each other and with the environment remain unclear. A complex trait such as starvation resistance depends on the actions of multiple genes and their interaction with environmental factors. Such traits are also called quantitative traits because they exhibit a continuous distribution of phenotypes. The genetic basis of quantitative traits can be studied using genome-wide association studies (GWAS), where genetic variants across the whole genome are tested for association to phenotypic variation in a population. Associated genetic loci are then tested for causality in follow-up studies.

Recently, GWASs became feasible also in *Drosophila melanogaster*, owing to the generation of genetic reference populations and the corresponding sequence information. One of these populations is the *Drosophila* Genetic Reference Panel (DGRP), which we used to dissect the genetic basis of starvation resistance at an early larval stage. We standardized the experimental setup to minimize potential confounding effects of environmental determinants and phenotyped 129 DGRP lines, detecting substantial genetic variation. We also found genetic correlations between larval starvation resistances and triglyceride levels, alcohol sensitivity,

and locomotory behavior. The GWAS identified 94 genetic variants, most of which are located in intronic and intergenic regions, potentially playing regulatory roles. We also identified 62 associated genes, the functions of which are diverse, illustrating the complex etiology of the trait. Some of these candidate genes are involved in starvation responses, regulation of metabolism and feeding behavior, and some of them have human homologs that are associated with metabolic traits and diseases, such as obesity, BMI, blood lipids, etc. Given the evolutionary conservation of major physiological and metabolic pathways, the results of this project will likely have implications beyond *Drosophila*.



2. ZUSAMMENFASSUNG

Natürliche Populationen sehen sich oft mit Nährstoffmangel konfrontiert. Die Physiologie der Anpassung an Hunger ist für evolutionäre Studien von Interesse und hat Auswirkungen auf die menschliche Gesundheit. Tiere decken ihren Energiebedarf normalerweise dadurch, dass sie ihre Nahrung verstoffwechseln. Wenn der Zugang zu Lebensmitteln begrenzt ist, müssen sie ihre körpereigenen Energiereserven zurückgreifen, um zu überleben. Die Fähigkeit von Organismen, längere Zeit mit knappen Nahrungsmittelressourcen auszukommen, nennt man Hungerresistenz. Sie wird letztendlich durch das Gleichgewicht zwischen Energiezufuhr und -verbrauch bestimmt.

Modellorganismen wie *Drosophila melanogaster* bieten Möglichkeiten, die physiologischen Reaktionen auf Hunger zu untersuchen und die zugrunde liegenden molekularen Mechanismen zu bestimmen. Obwohl sich während der Evolution die Arthropodenlinie von der Wirbeltierlinie vor mehr als 600 Millionen Jahren getrennt hat, deuten genetische, molekulare und verhaltensbiologische Analysen darauf hin, dass *Drosophila* und der Mensch sich nicht so unähnlich sind. Studien mit *Drosophila* haben gezeigt, dass die Hungerresistenz eng mit anderen Merkmalen wie den Nährstoffreserven, der Lebensdauer und den Resistenzen gegen andere Stressfaktoren korreliert. Hungerresistenz steht jedoch im Konflikt mit der Fruchtbarkeit. Diese Korrelationen sind evolutionär bestimmt und haben eine genetische Basis.

Mithilfe genetischer Untersuchungen konnten viele Gene identifiziert und charakterisiert werden, die eine Rolle in der Hungerresistenz in adulten Fliegen spielen. Trotz dieser Fortschritte sind noch immer viele Komponenten der Mechanismen, die an der Hungerresistenz beteiligt sind, unbekannt. Auch ist grösstenteils noch immer unklar, wie die Komponenten miteinander oder mit den Umweltbedingungen zusammenwirken. Es wird angenommen, dass ein komplexes Merkmal, wie die Hungerresistenz, aus den Wirkungen von mehreren Genen und ihrer Wechselwirkung mit Umweltfaktoren entsteht. Solche Merkmale werden auch als quantitative Merkmale bezeichnet, da sie eine kontinuierliche Verteilung von Phänotypen aufweisen. Die genetische Basis der quantitativen Merkmale kann mit genomweiten Assoziationsstudien (GWAS) untersucht werden, bei denen genetische Varianten über das gesamte Genom auf Assoziation mit der phänotypischen Variation einer Population untersucht werden. Die assoziierten, genetischen Loci werden dann in Folgeuntersuchungen auf ihre Kausalität getestet.

Aufgrund sinkender Kosten der Genomsequenzierung und der Erzeugung von genetischen Referenzpopulationen wurden GWASs auch in *Drosophila*

melanogaster möglich. Eine dieser Populationen ist das *Drosophila* genetic Reference Panel (DGRP), das wir verwendet haben, um die genetische Basis der Hungerresistenz in einem frühen Larvenstadium zu untersuchen. Um den Einfluss möglicher störender Umweltfaktoren zu minimieren, vereinheitlichten wir die Kulturbedingungen und den Versuchsaufbau. Wir untersuchten die Hungerresistenz in 129 DGRP Linien und fanden erhebliche genetische Variation. Weiterhin fanden wir genetische Korrelationen zwischen larvaler Hungerresistenz und dem Triglyceridspiegel, der Alkoholempfindlichkeit sowie zwischen Hungerresistenz und mehreren Merkmalen körperlicher Bewegung. Mithilfe der GWAS identifizierten wir 94 genetische Varianten, von denen sich die meisten in Introns und intergenischen Regionen befinden, möglicherweise also regulatorische Funktionen haben. Wir fanden 62 assoziierte Gene mit unterschiedlichen Funktionen, was die vielschichtigen Ursachen der Hungerresistenz verdeutlicht. Einige dieser Kandidatengene sind an der Hungerresistenz, an der Regulation des Metabolismus und am Fressverhalten beteiligt. Andere der identifizierten Gene haben menschliche Homologe, die mit Stoffwechseleigenschaften und Erkrankungen wie Übergewicht, BMI oder Blutfettwerten assoziiert sind. Aufgrund der evolutionären Konservierung der wichtigsten physiologischen und metabolischen Vorgängen können die Ergebnisse dieses Projekts Auswirkungen weit über *Drosophila* hinaus haben.

3. INTRODUCTION

The ability to tolerate prolonged food shortage is called starvation resistance. *Drosophila* has been a lead model organism to study starvation resistance owing to its short generation time, ease of culture, well-annotated genome, amenability to genetic manipulation, and lack of ethical issues to perform acute starvation experiments (Rion & Kawecki, 2007). Research on *Drosophila* starvation resistance has been implemented using different approaches such as artificial selection and population studies, candidate gene-based functional experiments, genetic screens and expression profiling. These studies provided important insights into the genetic basis of starvation resistance, even though still far from providing a complete picture.

Starvation resistance is a quantitative trait that depends on the cumulative actions of multiple genes and the interactions of these genes with each other and the environment. Dissecting the complex genetic basis of quantitative traits is challenging. In classical genetic studies, genetic manipulations are utilized to alter or suppress the functions of individual genes, and effects of these manipulations on the phenotype are assessed. An alternative strategy to identify the genes important for the trait is to perform a genome-wide association study (GWAS). In GWAS, phenotypic variation in a large population and the sequence knowledge of the individuals in this population are utilized to assess the effects of natural polymorphisms on the phenotype by making association tests. In this project, we take a GWAS approach to identify the genes underlying starvation resistance at the larval stage of *Drosophila*.

This introduction will begin with a general outline for the adaptive physiological mechanisms animals exploit to cope with the starvation stress. This discussion will particularly focus on the vertebrates (Chapter 3.1). Subsequent chapters will focus on *Drosophila* and depict to what extent these mechanisms show similarities across species. First, insights gained from studies on adult starvation resistance will be discussed; these include the multifactorial nature of its genetic basis, its environmental plasticity and its genetic correlations with other traits. Acute and adaptive physiological responses of flies to starvation will also be discussed (Chapter 3.2). As the focus of this thesis is starvation resistance at the larval stage, the physiological role of this stage in *Drosophila* life cycle and the importance of nutritional balance at this stage will be detailed (Chapter 3.3). Next, the molecular and physiological mechanisms that regulate energy homeostasis in *Drosophila* and their similarity to mammals will be given briefly (Chapter 3.4). After covering the functional genomics studies addressed to identify genes involved in starvation responses (Chapter 3.5), genome-wide association studies will be introduced as a complementary approach to study

quantitative traits. The basics of association mapping will be explained, and the advantages and shortcomings of GWASs will be discussed (Chapter 3.6). The aspects of *Drosophila* that render it an optimal model organism for GWASs will be discussed (Chapter 3.7 and 3.8). Lastly, the motivation for this thesis project will be summarized, and the specific aims will be listed (Chapter 3.9 and 3.10).

3.1. Starvation resistance in the animal kingdom

3.1.1. Physiological and behavioral adaptations to starvation

Starvation resistance and the underlying mechanisms have been studied in many animal species. Given the diversity of vertebrates, their habitats and their lifespan, it is not surprising that vertebrates differ extensively in their starvation survival abilities (Mrosovsky & Sherry, 1980; McCue, 2010). For instance, some vole species can tolerate starvation for only one day (Mosin, 1982), whereas snakes and frogs can survive up to 2 years of starvation (McCue, 2007), and the European eel was reported to survive up to 1594 days of starvation at 14°C (Boétius & Boétius, 1985). Despite these large differences, the mechanisms they exploit to resist starvation are based on the same principles: increasing energy supplies, reducing energy expenditure, and optimizing resource utilization within the body. These mechanisms and the relationships between starvation resistance and other behavioral and physiological traits will be discussed here in more detail.

Feeding habits

Starvation resistance is highly interconnected with feeding habits. Small birds and mammals, for instance, feed at frequent and regular intervals to maintain their high metabolic rates. These endotherms are relatively susceptible to starvation. On the opposite end, large ectotherms that are also sit-and-wait predators, such as snakes and crocodiles, feed relatively infrequently, and typically consume very large meals. They withstand starvation for much longer periods (Secor & Diamond, 2000; Wang et al., 2006).

Body size

In mammals, body size shows strong correlation with starvation endurance. Large mammals have higher ratios of energy reserves to energy expenditure than small mammals (Lindstedt & Boyce, 1985). Therefore, larger body sizes are favored during periods of food shortage. It has been postulated that starvation

endurance has been one of the selective pressures for the evolution of body size (Millar & Hickling, 1990).

Physical activities

Animals must minimize their energy expenditure during starvation. Indeed, many animals reduce unnecessary physical activities (Westerterp, 1977; Dunlap, 1995). But some animals also display higher locomotor activities (Kettersen & King, 1977). From an energetics point of view, increased locomotion imposes a disadvantage for survival by depleting the endogenous resources faster. However, increased locomotion might be beneficial to escape from local pressures of resource shortage and to increase the chance of finding new resources (Lendvai et al., 2004; McCue, 2010). For example, captive rats reduced their activity in the initial phases of starvation, but increased their activity markedly in the late phase during which they ran out of the lipid reserves and switched to protein catabolism. Therefore, their hyperactivity might reflect an escape behavior (Sclafani & Rendel, 1978; Koubi et al., 1991).

Body temperature and dormancy

Thermoregulation is an effective energy saving strategy. Reduction in body temperature during inactive parts of the diurnal cycle, also known as daily torpor, has been reported in birds and mammals including humans (Tobey, 1951; Hainsworth & Wolf, 1970; Westerterp, 1977). Unlike in endotherms, thermoregulation in ectothermic vertebrates is behavioral; they choose suitable environments. When fishes, amphibians and reptiles were exposed to a thermal gradient during starvation, they preferentially selected lower temperatures compared to fed controls (Bicego et al., 2007).

When food shortage is due to seasonal changes and is accompanied with changes in ambient temperature or lack of water, some animals enter into prolonged states of dormancy; hibernation or aestivation (Humphries et al., 2003; Geiser, 2004). Others such as birds and some large mammals migrate to more desirable areas (Wang et al., 2006). Hibernating or migrating animals accumulate large internal lipid deposits before these predicted periods of food deprivation. Some hibernating mammals such as hamsters also hoard food (Baggott, 1975; Nedergaard & Cannon, 1990).

Metabolic rate

Many studies reported reductions in resting and basal metabolic rates without accompanying reductions in body temperature in starving animals (Mosin, 1984; McCue, 2007, 2010). This hypometabolic state might result from decreased cellular metabolism associated with decreased protein synthesis, turnover, and ATP supply rates (Hand & Hardewig, 1996). Alternatively, it might be due to large mass loss observed in tissues with high turnover rates and metabolic costs,

such as the gastrointestinal tract that undergoes marked atrophy upon starvation (Hobold et al., 2004; Wang et al., 2006; McCue, 2007). However, it is noteworthy that when concomitant reductions in body mass are also accounted for, these studies indeed reported a shallower hypometabolism, no change, or even hypermetabolism (McCue, 2010). These inconsistent outcomes in mass-specific metabolic rates might arise due to body composition changes that occur at different rates or patterns in different species (McCue, 2010) or between obese and lean individuals of the same species (Elia et al., 1999).

Reproduction

Lastly, starvation strongly inhibits reproduction, a highly energy-demanding behavior. In mammals, ovulation and sexual receptivity of females are highly susceptible to food deprivation (Morin, 1986; Jones & Wade, 2002). Most animals are indeed seasonal breeders; they only breed under food abundance (Wang et al., 2006).

3.1.2. Metabolic adaptations to starvation

As discussed above, animals adapt their metabolic rates under starvation. They undergo progressive changes in the primary metabolic fuel source (i.e. carbohydrate, lipid, amino acids) and in the rate of body mass loss. These metabolic responses to food deprivation were extensively studied in mammals (Cahill, 1976).

In the first phase of food deprivation, hepatic glycogen breakdown becomes the main source of blood glucose. Lipolysis in the adipose tissue is also induced. Peripheral tissues such as muscles and adipose tissue progressively diminish glucose utilization and switch to use fatty acids (Cahill, 1976), whereas the central nervous system remains dependent on blood glucose. Since hepatic glycogen stores are limited, liver starts to produce glucose also by gluconeogenesis. Initially, the fuel for gluconeogenesis is mainly amino acids from proteolysis of peripheral proteins, and partly glycerol from lipolysis (Hanson & Owen, 2013). However, this cannot be sustained for long, because high rates of protein catabolism would be detrimental for the body (Cahill, 2006).

With prolonged food deprivation, the body shifts to lipid-dominated catabolism. The liver starts to produce ketone bodies (i.e., acetoacetate and β -hydroxybutyrate) from fatty acids released by the adipose tissue. Ketone bodies can be used for energy production by most of the body tissues including the brain (Owen et al., 1967; Grey et al., 1975). The brain progressively shifts from

glucose utilization to ketone body utilization. Muscles progressively shift to fatty acid utilization to spare ketones for use in the brain (Owen & Reichard, 1971). These transitions, in turn, decrease the need for high rates of gluconeogenesis and, hence, muscle proteolysis. As protein catabolism is reduced and most of the metabolism is fueled by energy-dense fat, weight loss is minimized at this stage (Paoli et al., 2013). Thus, the body is 'adapted' to starvation (Wang et al., 2006). A lean adult can maintain this state for several weeks, whereas an obese adult can withstand much longer (Kerndt et al., 1982).

In the last phase of starvation, when lipid reserves are depleted, the body starts to utilize proteins for energy production. Body mass loss increases, and death occurs from multiple organ failure if not infection diseases (Koubi et al., 1991; Bertile et al., 2003).

The studies mentioned above highlighted the importance of lipids for starvation endurance. Lipids have certain advantages over other types of storage molecules: they have a higher caloric density (9.4 cal/g for triglycerides) and their compact storage in water-free droplets minimizes the metabolic requirements to store them (McCue, 2010). In contrast, glycogen is a poor means of storing energy, given its low caloric density (4 cal/g) and water-laden storage (Kreitzman et al., 1992). However, glycogen can be quickly metabolized to meet a sudden need of glucose. Proteins, on the other hand, are not suited for energy storage, because increasing the protein composition, i.e. by building muscles, would, in the first place, increase the body's energy expenditure (Cahill, 2006).

3.2. Starvation resistance in *Drosophila*

Drosophila has been an important model organism to study starvation responses. One assay that can be exploited readily in *Drosophila* is to measure the survival time under starvation, which would be a direct readout for the organism's ability to resist to starvation. Even though molecular and physiologic responses to starvation have been studied in both adults and larvae, resistance to starvation has almost entirely been assessed at the adult stage (Rion & Kawecki, 2007). In these assays, flies are subjected to acute starvation and their survival time is recorded after regular visual inspections. Here, male and female flies are tested separately, but groups of flies of the same sex can be kept in the same vials since *Drosophila* adults do not feed on dead carcasses (Huey et al., 2004). The flies are starved in vials that contain either non-nutritive agar medium or water-soaked filter papers to prevent desiccation, to which flies are more sensitive than they are to starvation (Da Lage et al., 1989). In laboratory selection experiments, flies are starved till 50–90% die, and the survivors are bred to establish the

subsequent generation. After a certain level of direct response to selection is achieved, correlated (indirect) responses are measured to understand the physiological mechanisms underlying or interacting with starvation resistance (Harshman & Hoffmann, 2000). These experiments provided important insights into the genetic basis of starvation resistance and its environmental plasticity, both of which will be discussed throughout this chapter.

3.2.1. Genetic variation for starvation resistance

There is substantial genetic variation for starvation resistance in *Drosophila* as documented by multiple lines of evidence. First, *Drosophila* populations exhibit rapid and strong responses to selection for starvation resistance in laboratory selection experiments (Chippindale et al., 1996; Harshman & Schmid, 1998; Baldal et al., 2006; Schwasinger-Schmidt et al., 2012). For instance, Chippindale et al. obtained a remarkable 4 to 6 fold increase in starvation survival time upon 60 generations of strong selection (Chippindale et al., 1996). Second, high heritability estimates were obtained in studies on isofemale lines derived from the same population (Hoffmann et al., 2001; Mackay et al., 2012). Mackay et al. assessed the starvation resistances of 203 isofemale inbred *Drosophila melanogaster* lines, collectively known as *Drosophila* Genetic Reference Panel (DGRP). They observed 3 fold differences among starvation survival times of these lines (Mackay et al., 2012). Third, *Drosophila melanogaster* populations collected from different geographical locations also showed large variations in their starvation resistances (Da Lage et al., 1990; Parkash & Munjal, 2000; Hoffmann et al., 2001). These variations might reflect the genetic variations arising from local adaptations to temperature, humidity and food resources. Lastly, significant variations were also observed among different *Drosophila* species (Van Herrewege & David, 1997; Matzkin et al., 2009). However, since species were represented by only one or a few populations in these studies, interspecific comparisons might be confounded by intraspecific variations (Hoffmann & Harshman, 1999).

Notably, starvation resistance in *Drosophila* shows sexual dimorphism. With a few exceptions (Da Lage et al., 1989; Matzkin et al., 2009), females were mostly reported to tolerate acute starvation longer than males (Harshman et al., 1999a; Robinson et al., 2000; Ballard et al., 2008; Schwasinger-Schmidt et al., 2012; Lee & Jang, 2014). Similarly, females survive longer than males on average in the DGRP population (Mackay et al., 2012). However, the magnitude of sexual dimorphism is not uniform across the DGRP lines, but varies greatly. Indeed, in 9 of the 203 lines, males are more resistant. Hence, there is also a large genetic variation in the sexual dimorphism of the starvation resistance. These

observations imply that there are sex-specific genetic factors underlying natural variation in the trait.

3.2.2. Environmental plasticity of starvation resistance

Starvation resistance is highly sensitive to environmental variables. In addition to temperature and humidity (Da Lage et al., 1989), the nutritive value of the food is a key factor. For instance, adult *Drosophila melanogaster* die within 3-4 days under acute starvation, but they can survive up to 3 weeks when starved on 20% sucrose PBS agar medium (unpublished observation, Tettweiler et al. 2005). Hence, flies rely greatly on carbohydrates for maintaining their somatic homeostasis and fueling their physical activities.

Besides the nutritional value of the starvation medium, the dietary history of flies before starvation also affects their resistance to acute starvation. For instance, feeding flies with higher amounts of protein prior to starvation decreases their starvation resistance (Chippindale et al., 1993). This decrease was found to parallel the decreases in their carbohydrate and lipid contents (Bradley & Simmons, 1997). Lee and Jang demonstrated that differences in the protein:carbohydrate ratios of the diet could give rise to differences in starvation resistance as extreme as 2 fold (Lee & Jang, 2014). They showed that higher P:C diets resulted in higher lipid-free body proportions (e.g. energy demanding muscles) while lower P:C diets resulted in higher lipid proportions (i.e. energy supplies). These differences in body composition ultimately determined the starvation resistance. Lastly, larvae that were raised on carbohydrate rich diet were also found to have higher starvation resistance as adults (Sisodia & Singh, 2012), indicating that effects surpassing from the larval to the adult stage are also important.

In conclusion, diet has a great influence on the metabolic traits in *Drosophila*. Both larval nutrition and adult nutrition affect the starvation resistance of adult flies. Therefore, comparative studies between experiments performed at different times and in different laboratories might be confounded by these effects, and must be interpreted with caution.

3.2.3. Physiological mechanisms underlying starvation resistance

Food limitation stress is an important source for natural selection, and animals have evolved certain physiological mechanisms to maximize their survival under

starvation (Rion & Kawecki, 2007; Gibbs & Reynolds, 2012). These mechanisms are diverse but mainly impinge on 3 strategies; increasing the energy supplies, minimizing the energy expenditure, and optimizing the resource allocation within the body. These mechanisms are discussed below.

As noted in Chapter 3.1.2, food-deprived animals prioritize their nutrient reserves for utilization. They initially use their glycogen stores, and then progressively switch to lipids. Lipids fuel most of the metabolism during prolonged periods of starvation. Therefore, starvation endurance is highly dependent on the amount of the body's lipid reserves. Catabolism of proteins is minimized until lipid reserves are depleted, since this is detrimental to the body (Cahill, 1976). A similar transition in the utilization of endogenous resources was observed in *Drosophila* (Lee & Jang, 2014). Food-deprived flies initially lost only their non-lipid body weight. During this period, they most probably digested and emptied the ingested food in their digestive tract, and also utilized their stored and circulating carbohydrate sources. Then, they started to lose only their lipid content, and survived until they completely depleted their lipids. Lipid-free body weight did not decrease during this lipid-consuming phase, implying that proteins were not used as energy reserves (Lee & Jang, 2014).

Given the central role of internal energy reserves in survival under starvation, it is not surprising that experimental evolution studies in *Drosophila* extensively documented a positive correlation between lipid reserves and starvation resistance; flies selected for starvation resistance adapted to accumulate significantly more lipid reserves (Chippindale et al., 1996; Djawdan et al., 1997; Djawdan et al., 1998; Harshman et al., 1999a; Hoffmann et al., 2005; Baldal et al., 2006; Schwasinger-Schmidt et al., 2012). Although selection was employed at the adult stage, selected flies acquired their excess lipid reserves during both adult and larval stages so that they had greater lipid reserves already at eclosion (Chippindale et al., 1996). However, there is no correlation between lipid levels and starvation resistance in isofemale lines generated from wild populations, including the DGRP lines (Hoffmann et al., 2001; Jumbo-Lucioni et al., 2010). Therefore, even though strong selection pressure predisposes lipid accumulation as the primary mechanism to increase resistance to starvation when food resources are abundant under laboratory conditions, accumulation of lipid reserves is only one aspect of this multifaceted trait, and may not be so prevalent in determining starvation resistance in more natural settings (Hoffmann et al., 2001). In accordance with this, several studies found correlations between levels of other macromolecules and starvation resistance. For instance, total carbohydrate and trehalose contents increased after selection for starvation resistance (Djawdan et al., 1997; Schwasinger-Schmidt et al., 2012). Furthermore, glycogen and glucose levels showed modest positive associations with starvation resistance in the DGRP population (Jumbo-Lucioni et al., 2010).

Resource accumulation as a strategy to survive food limitation is reflected also by the increased food seeking behavior of *Drosophila* when starved. *Drosophila* increased their physical activity and showed increased response to chemical attractants upon starvation (Farhadian et al., 2012). Moreover, when food is reintroduced, they consumed more food and at a higher rate. Starved flies stored more food in their crop compared to flies fed *ad libitum*. These effects were transient and returned to basal levels over time (Farhadian et al., 2012). Likewise, flies that were fed with low-calorie diets also compensated for insufficient caloric intake by eating more food (Carvalho et al., 2005). Therefore, starvation stimulates food intake. In the reverse situation, it is expected that higher food intake levels result in higher starvation resistances due to accumulation of larger endogenous reserves. However, food intake was found to have a strong negative correlation with starvation resistance in the DGRP population (Garlapow et al., 2015). Moreover, selection for starvation resistance resulted in dramatically reduced food intake (Masek et al., 2014). This paradox might be explained by examining how food intake was measured in both studies. In capillary feeding (CAFE) and proboscis extension reflex (PER) response assays, flies are deprived of food for some time prior to the measurement for proper quantification; therefore, fasting-induced feeding levels are measured. In other words, starvation resistant flies fed less than starvation sensitive flies upon fasting, implying that they felt hunger less than sensitive flies. In support of this view, starvation selected flies required longer fasting periods to show PER responses comparable to control flies (Masek et al., 2014). A positive relationship between starvation resistance and food intake might be observed when food intake is measured during fed conditions, but this experiment has not been done yet.

The nutritional status of animals has a large impact on their physical activity and sleep behavior. In *Drosophila*, food deprivation was found to result in elevated locomotor activity (Knoppien et al., 2000; Yang et al., 2015), reduced number of sleep bouts and reduced total sleep time (McDonald & Keene, 2010). Hyperactivity of flies was especially accentuated before starvation-induced death (Lee & Park, 2004). Higher mobility of flies increases their chance of finding food in nature. However, in laboratory conditions, it only results in higher energy expenditure and is disadvantageous since flies are entrapped in vials and cannot find extra food. Probably due to this, in artificial selection experiments, starvation-selected flies adapted to reduced locomotor activity and increased sleep compared to control flies (Williams et al., 2004; Masek et al., 2014). Moreover, starvation did not suppress sleep in selected flies, unlike in control flies (Masek et al., 2014). Therefore, selection for starvation resistance favored reduced activity and enhanced sleep as a means of energy conservation. In the DGRP population, several sleep traits showed correlations with starvation

resistance in the same direction as found in selection studies (Mackay et al., 2012; Harbison et al., 2013).

As a strategy to minimize energy expenditure, some animals can decrease their metabolic rates under starvation conditions. This relationship between metabolic rate and starvation resistance was investigated also in *Drosophila*; however, the results were inconsistent. Some studies found a reduction in mass specific metabolic rates upon starvation in both control and starvation-selected flies (Djawdan et al., 1997), whereas others observed sex-specific and different responses among different lines of selected and control groups (Baldal et al., 2006). The response of metabolic rate to selection for starvation resistance also varied between different studies. Baldal et al. did not find any difference in mass specific metabolic rates between selected and control flies in fed and starved conditions (Baldal et al., 2006). Others showed that selection resulted in lower mass specific metabolic rate in fed and starved conditions (Djawdan et al., 1997; Harshman et al., 1999a). But this difference between selected and control lines disappears when the non-metabolizing carbohydrate and lipid contents, which are higher in selected lines than in controls, are deduced from the total mass (Djawdan et al., 1997). These inconsistent results suggest that the metabolic rate responds to acute and chronic starvation conditions in a plastic manner.

Like other animals, *Drosophila* also adjust their reproduction according to nutritional status. Food availability promotes mating in both sexes. It has been shown that food-derived odors act as aphrodisiacs and increase male courtship (Grosjean et al., 2011). Likewise, females become more receptive to males when they are fed (Lebreton et al., 2015). Starvation immediately ceases these reproductive activities, as they are costly from an energy homeostasis perspective. Females stop laying eggs when food is withdrawn. They initiate programmed cell death in the ovaries at stage 8/9 to reabsorb their nutrients (Terashima & Bownes, 2005). Similarly, males respond to protein starvation by eliminating transit-amplifying spermatogonia while preserving a reduced pool of germline stem cells (Yang & Yamashita, 2015). Therefore, the use of endogenous resources is diverted more towards survival than reproduction during times of nutrient restriction.

In conclusion, starvation resistance reflects an animal's ability to store internal energy reserves and to utilize these resources efficiently during times of food limitation (Wang et al., 2006). *Drosophila* increase their food intake and storage, regulate their physical activity and metabolic rate, and decrease their reproduction in response to nutritional stress. These adaptations confer them the ability to survive starvation longer.

3.2.4. Correlations of starvation resistance with other stress resistance traits and life history traits

Several artificial selection and population comparison studies investigated the relationships between resistance to starvation and other environmental stressors. Mostly positive correlations were found between starvation and desiccation resistances (Harshman et al., 1999a; Harshman et al., 1999b; Bubliy & Loeschcke, 2005; Hoffmann et al., 2005; Hoffmann & Harshman, 1999) although at least one study reported a lack of correlation (Djawdan et al., 1998). One study found decreased resistance to cold in females but not in males upon selection for starvation resistance (Hoffmann et al., 2005); while another found a positive correlation between these two traits (Bubliy & Loeschcke, 2005). There seems to be no correlated response between resistance to starvation and heat shock (Bubliy & Loeschcke, 2005). Several studies reported a positive association between starvation resistance and oxidative stress (Harshman et al., 1999b); however this correlation is not universal since, at least in one study, only some starvation selected lines increased their resistance to paraquat (Baldal et al., 2006). Correlated or indirect responses in multiple stress resistance traits after selection for resistance to a specific stress suggest common physiologic and genetic mechanisms that underlie multiple traits. These common mechanisms might be related to reduced energy expenditure, increased energy conservation and increased energy reserves (Bubliy & Loeschcke, 2005; Baldal et al., 2006). However, in many cases, these correlations were inconsistent within or among different studies suggesting that multiple mechanisms underlie genetic variation in stress resistance, and only some of them are common among different traits (Bubliy & Loeschcke, 2005; Rion & Kawecki, 2007).

There are also many striking correlations between starvation resistance and life history traits. Laboratory selection for starvation resistance is predominantly based on increased resource accumulation; however, increased resource acquisition comes with its own costs. One such drawback is the prolongation of larval development to increase feeding time. This correlation was reported in many studies (Chippindale et al., 1996; Harshman et al., 1999a; Bubliy & Loeschcke, 2005; Gibbs & Reynolds, 2012) except one (Hoffmann et al., 2005). Another important life history trait affected by increased starvation resistance is fecundity; flies selected for either trait displayed reduction in the other trait (Wayne et al., 2006). The trade-off between fecundity and starvation resistance might partially be explained by differential allocation of resources between reproduction and survival (Partridge et al., 2005). Alternatively, several studies also demonstrated that reproduction and survival required a different nutritional balance in the diet for maximum performance; high P:C diets maximized reproduction whereas low P:C diets maximized survival traits such as starvation resistance and longevity (Lee & Jang, 2014). Longevity is an important

life history trait that was repeatedly shown to have correlate positively with starvation resistance (Service, 1987; Rose et al., 1992; Chippindale et al., 1993; Bubliy & Loeschcke, 2005), however one study showed that this correlation is sex-dependent (Baldal et al., 2006) while another study showed lack of correlation (Harshman et al., 1999b). A strong positive correlation between starvation resistance and virgin lifespan was also detected in the DGRP population (Arya et al., 2010). These correlations imply partially overlapping mechanisms for these survival traits. Indeed, it is well-known that dietary restriction and many genetic or chemical manipulations that mimic dietary restriction result in lifespan extension in diverse organisms indicating that mechanisms that play a role in starvation responses might also be important in extending lifespan (Broughton et al., 2005).

In conclusion, adult starvation resistance shows mostly positive correlations with lifespan and negative correlations with developmental rate and fecundity. Furthermore, correlations with certain stress resistance traits suggested both common and trait-specific stress resistance mechanisms. The molecular mechanisms underlying these relationships between organismal traits remain largely unclear and need further investigation.

3.3. *Drosophila* larval stage is characterized by extensive feeding and growth

Drosophila is a holometabolous insect; and its life cycle consists of morphologically distinct stages (Tennessen & Thummel, 2011). Embryogenesis is completed in the egg in one day, and the hatching larvae develop through 3 instars (L1, L2 and L3), each separated by molting of the larval cuticle. Under nutrient rich conditions, the first and second instars each last one day, and the third instar lasts two days. Larvae continuously feed to grow in size; their body mass increases approximately 200 fold (Church & Robertson, 1966). The first two molts allow the larvae to accommodate this enormous size increase. Towards the end of the third instar, larvae cease feeding and start wandering to search for a suitable site for pupation. With the third molt, the larval cuticle becomes the pupal case. The pupal stage lasts for 4 days during which the larvae metamorphose into the imago, or the adult fly. Adults start mating 1 day after eclosion from the pupae. They can live for more than 10 weeks (Tennessen & Thummel, 2011).

During metamorphosis, most larval tissues undergo programmed cell death, and new adult organs are formed by the proliferation and differentiation of precursor cells. These precursor cells are determined during embryogenesis. They are set apart in groups, and stay undifferentiated throughout the larval

stage (Edgar, 2006). Therefore, larvae can speculatively be considered as vessels for the adult-specific precursor cells feeding extensively to provide the nutrients and energy for their growth. Growth in larvae occurs by different mechanisms for the larval specific tissues and adult precursor cells. Larval specific gut, fat body, salivary gland, Malpighian tubule, trachea, muscle and epidermis tissues are made up by endoreplicating cells; they undergo successive rounds of DNA synthesis without mitosis, which results in polytene cells. Growth in these tissues is assumed by an increase in cell size (Smith & Orr-Weaver, 1991; Britton & Edgar, 1998). On the other hand, adult-forming precursor cells and the germ cells are mitotic diploid cells. Growth in these cells occurs through increase in cell number rather than cell size (Britton & Edgar, 1998). Most of the adult precursor cells are called imaginal cells since they are for the imago. During metamorphosis, imaginal disc cells give rise to adult head structures, appendages and genitalia; imaginal ring cells and histoblasts give rise to adult internal organs such as gut and salivary glands, and abdominal epidermal segments (Curtiss & Heilig, 1995). The larval nervous system is not lysed, but undergoes remodeling. Gonads are preserved; larval fat body cells also persist as dissociated cells only to be lysed within 2 days after eclosion (Aguila et al., 2007).

During the non-feeding pupal stage, *Drosophila* relies on endogenous sources. Nutrient reserves acquired during larval feeding and histolysis of larval-specific tissues provide the energy and materials necessary for building new adult-specific tissues. Larval fat body, indeed, accumulates more energy reserves than needed only for metamorphosis (Merkey et al., 2011). Newly eclosed adult *Drosophila* do not feed before approximately 8 hours after eclosion, and hence also rely on these reserves (Merkey et al., 2011). Persisting larval fat cells provide their lipids to the adult body till they undergo cell death within 2 days after eclosion. In addition, young females use these lipid reserves also for the maturation of the ovaries (Aguila et al., 2013). Importantly, starvation resistance of newly eclosed adults highly depend on these larval fat cells (Aguila et al., 2007).

Larval growth determines the final adult body size; therefore, it is tightly monitored and regulated (Edgar, 2006; Tennessen & Thummel, 2011). When larvae achieve a sufficient level of growth, they initiate metamorphosis. The weight at which larvae can successfully complete metamorphosis and develop into adults without further feeding is called the minimal viable weight. Larvae surpass this threshold weight in early L3 instar (Beadle et al., 1938). Under nutrient rich conditions, several hours after attaining the minimal viable weight, larvae attain the so-called critical weight. Critical weight is defined as the weight after which food withdrawal does not delay the onset of metamorphosis (Edgar, 2006; Mirth & Shingleton, 2012). Therefore, larvae become committed to metamorphosis once they achieve the critical weight. However, it takes some

time after attaining the critical weight to cease feeding and initiate pupation. During this lag period, larvae continue to grow exponentially if cultured on rich food. Hence, most of the increase in body size occurs during this period (Beadle et al., 1938; Bakker, 1959).

The timing of larval molts and the onset of metamorphosis are determined by pulses of the steroid hormone 20-hydroxyecdysone (20E) that is released from the prothoracic gland (Thummel, 2001). Many nutritional and developmental signals feed into this steroid hormone signaling, regulating the production of 20E, and, therefore, the duration of larval growth (Edgar, 2006). Food withdrawal before attaining the critical weight delays subsequent 20E pulses and the onset of pupation (Beadle et al., 1938). If larvae are starved during the early instars, their growth is arrested, and 20E pulses that initiate metamorphosis do not occur. Refeeding these larvae reactivates their growth, and they develop into normally sized adults. On the other hand, if larvae are starved after the second molt but before the critical weight is attained, they can still initiate metamorphosis, but the onset of metamorphosis is significantly delayed (Beadle et al., 1938; Bakker, 1959). As larval feeding time increases, pupal mortality decreases (Alpatov, 1930; Bakker, 1959). Only the larvae that achieve the minimal viable weight can succeed in developing into adults. Lastly, if larvae are starved after the critical weight attainment, the onset of metamorphosis is no longer delayed. Indeed, it is even accelerated (Beadle et al., 1938; Stieper et al., 2008). Since larvae starved after the minimal viable weight cannot further grow in size but survive into adults, these adults will be as small as one fourth to one fifth of the adults that derive from normally fed larvae (Bakker, 1959). Therefore, nutrient availability affects both the duration of the larval growth and the final body size.

Interestingly, larvae share some of the behavioral responses to food deprivation observed in adults. For instance, starvation induces intense muscle contractions and increased mobility, and larvae escape from their vials when starved (Bakker, 1959). They increase their feeding rate upon starvation and change their food preferences (Wu et al., 2005). Besides these common starvation responses, *Drosophila* larvae exhibit one distinct feature that adults do not: they can scavenge on dead fly carcasses (Gregg et al., 1990) and attack larger larvae upon food limitation under crowded laboratory conditions (Vijendravarma et al., 2013).

Research on starvation resistance at the larval stage has been very limited. However, one study examined the life history consequences of adaptation to larval dietary restriction (Kolss et al., 2009). They selected *Drosophila* lines that survived and developed within a limited time on poor larval food. This selection regime generated lines with increased egg-to-adult viability and development

rate on poor food. Intriguingly, selected lines had the same egg-to-adult viability as control lines on standard food, but preserved faster development. Moreover, they displayed some trade-offs with adult stage fitness traits on standard food such as smaller body size and lower fecundity, despite no change in lifespan. They were also less tolerant to adult starvation despite having the same proportional pupal lipid contents with controls (Kolss et al., 2009). Additionally, many other studies focused on the physiological and evolutionary correlates of larval crowding, for which dietary restriction might be one of the many stressors (Rion & Kawecki, 2007). Selection for larval crowding resulted in both increased and decreased adult starvation resistances in different populations (Zwaan et al., 1991; Borash et al., 2000; Borash & Ho, 2001). These observations point to alternative mechanisms for adaptation to high larval density. On the other hand, how adaptation to larval crowding affect larval starvation resistance has not been addressed so far. In conclusion, nutritional and metabolic status of *Drosophila* larvae has important life history consequences.

To sum up, larval stage is characterized by extensive feeding and growth to produce a correctly sized adult body. Hence, adult fitness is greatly dependent on larval nutrition and growth. Despite this, most of the metabolic, behavioral and stress resistance traits have not been well studied at the larval stage so far. Given the anatomical and physiological differences between adults and larvae, such traits might have different genetic underpinnings in larval stages than in adult stage. Therefore, larval traits need to be investigated in more detail.

3.4. Molecular mechanisms underlying metabolic homeostasis in *Drosophila*

Drosophila has become a powerful model organism to study metabolic homeostasis and pathogenesis (Bharucha, 2009). In *Drosophila*, food is digested and absorbed in the gut, which is the equivalent of the vertebrate intestine (Lemaitre & Miguel-Aliaga, 2013). The crop stores food before it enters the midgut. The fat body is the main storage organ for both lipids and glycogen. Hence, its storage function overlaps with those of vertebrate white adipose tissue and liver. Oenocyte cells also function as liver by accumulating lipids mobilized from fat body during starvation (Gutierrez et al., 2007). In this sense, the fat body-oenocytes axis resembles the mammalian adipose-liver axis. Malpighian tubules are the fly equivalents of kidney; they maintain osmoregulation. Metabolic homeostasis is largely regulated through insulin and glucagon homologs of *Drosophila* (Brogiolo et al., 2001; Lee & Park, 2004; Bharucha, 2009). Insulin is secreted from the insulin-producing cells (IPCs) residing in the pars intercerebralis of the fly brain. The fly homolog of glucagon, adipokinetic hormone (AKH), is secreted by the corpus cardiaca in the ring gland,

a major endocrine organ that also contains 20E-producing prothoracic gland. Neurons originating from pars cerebralis innervate corpus cardiaca, resembling the hypothalamus-pituitary axis in vertebrates (Bharucha, 2009; Owusu-Ansah & Perrimon, 2014). This neuroendocrine system processes signals related to the internal metabolic status and elicits appropriate systemic responses (Lee & Park, 2004; Bharucha, 2009; Yang et al., 2015). The insulin signaling pathway and other major central metabolic pathways are largely conserved between *Drosophila* and humans. However, insulin secretion in *Drosophila* is primarily dependent on the presence of amino acids rather than carbohydrates or fats (Gibbs & Reynolds, 2012). The fat body plays a central role in sensing the nutritional signals and conveying them to the brain (Géminard et al., 2009). It also coordinates growth in other organs by secreting humoral growth factors and metabolites (Britton & Edgar, 1998; Edgar, 2006; Rajan & Perrimon, 2012).

Metabolic homeostasis is, to a large extent, regulated by peptidergic signals (Bharucha, 2009; Owusu-Ansah & Perrimon, 2014). To begin with, Neuropeptide F, the *Drosophila* homolog of mammalian Neuropeptide Y, stimulates food consumption (Lee et al., 2004), whereas hugin, the homolog of mammalian neuromedin U, suppresses it (Melcher et al., 2006). Unpaired2 (upd2) is a functional equivalent of mammalian leptin despite a lack of a clear sequence-based homology. Upd2 is a cytokine secreted from the fat body in response to fat and sugar and triggers insulin secretion in the brain. Strikingly, human leptin can rescue the upd2 mutant phenotype (Rajan & Perrimon, 2012). Allatostatin A suppresses food intake (Hergarden et al., 2012) and also regulates food choices (Hentze et al., 2015). Sex peptide, which is found in the male seminal fluid and transferred into the female during mating, strongly stimulates female food intake (Carvalho et al., 2006). As a result of this, mated females have higher lipid reserves and increased starvation resistance compared to age-matched virgin females (Rush et al., 2007). Many other peptides, such as octopamine (norepinephrine homolog) and drosulfakinin (cholecystokinin homolog), play important roles in the regulation of feeding behavior and systemic energy homeostasis in *Drosophila* (Owusu-Ansah & Perrimon, 2014). Most of these peptidergic signals are common to both *Drosophila* and mammals. Therefore, since metabolic regulation and other fundamental processes have been conserved among metazoans, research on metabolic traits in *Drosophila* has potential implications in humans as well.

3.5. Functional genomics studies identify many genes involved in starvation responses

Drosophila melanogaster is well-suited to study the genes underlying starvation resistance. It is the first metazoan that has been fully sequenced (Adams et al., 2000). Diverse libraries of RNAi lines and transposable element-based mutant lines are available for reverse genetic screens to determine the genes involved in starvation resistance. *Drosophila* also offers many powerful tools for genetic manipulation, which might be exploited to study the functions of identified genes in depth in candidate-gene-based studies. Its well-annotated genome also facilitates transcriptome and proteome profiling studies to identify genes and pathways with altered expression during starvation. The current chapter covers these large-scale functional genomics studies.

In a large-scale P-element mutagenesis screen, Harbison et al. searched for genes that affect starvation resistance. Most of the genes they identified affected starvation resistance in a sex-specific manner (Harbison et al., 2004). These genes included not only those involved in metabolism and feeding behavior, but also those involved in cell fate specification and cell proliferation. They argued starvation resistance is not solely dependent on the genes involved in the physiological responses to starvation, but also on the genes that affect the development of organs and tissues that are involved in the responses to starvation and resource allocation (Harbison et al., 2004). For instance, one of the genes that influence the starvation resistance is *salm* (*spalt major*), which plays a critical role in the formation and proliferation of oenocytes, the hepatocyte-like cells of *Drosophila* (Elstob et al., 2001; Rusten et al., 2001). Therefore, developmental events might also influence adult starvation resistance.

Several groups conducted transcriptome profiling studies for adult starvation to identify genes regulated by the nutritional status (Groenke et al., 2005; Harbison et al., 2005; Fujikawa et al., 2009). Up to one quarter of the whole genome was found to be involved in the transcriptional response to starvation. Nearly half of the transcriptome showed sex-specific expression patterns (Harbison et al., 2005). Starvation-induced genes included metabolic enzymes, cytochromes, metabolite transporters, kinases, proteins involved in lipid metabolism, etc. On the other hand, starvation-suppressed genes included, along many proteases and some metabolic enzymes, those involved in reproductive activities and immune response (Groenke et al., 2005; Harbison et al., 2005; Fujikawa et al., 2009). This finding reflects a reduction in resource allocation towards sexual activity and immunity during food deprivation, which underlies the trade-offs between survival and reproduction (Wayne et al., 2006), and between stress resistance and protection against infections (Hoang, 2001).

Harbison et al. also compared the genes identified in the mutagenesis screen to the genes that showed altered expression during starvation (Harbison et al., 2005). They noted that many genes, although affecting starvation resistance, did not change their expression upon starvation. Similarly, many genes that were shown to respond to starvation did not affect starvation resistance. These discrepancies might be interpreted by several explanations. First of all, many starvation responsive genes, such as those involved in reproduction, are not directly contributing to the physiological responses to starvation and, hence, are unlikely to affect starvation resistance. Second, genes affecting starvation resistance might be regulated at the post-transcriptional or translational level rather than at the transcriptional level. Third, the mutagenesis approach might look over the genes with repressed expression during starvation if the effect of these genes on starvation resistance is mediated already through loss of their activity during starvation. Finally, mutagenesis of some genes might have positive effects before starvation and negative effects during starvation, or vice versa, and these counteracting effects may cancel out each other resulting in no net effect on the starvation resistance (Harbison et al., 2005).

In addition to the acute responses to starvation in gene expression, evolutionary responses to chronic starvation were also studied. Sorensen et al. profiled the transcriptomes of flies selected for resistance to starvation and other environmental stresses, and compared to the transcriptome of unselected control flies (Sorensen et al., 2007). They detected differentially expressed genes both specific to a particular selection regime and common for all selection regimes, suggesting that there are specific and general stress responses. They also found that selection for longevity, desiccation and starvation resistance resulted in similar changes in gene expression, reflecting the correlations of these traits at the phenotypic level. On the other hand, there was little overlap between the differentially expressed genes identified in this study and the starvation responsive genes identified in previous expression profiling studies that employed acute starvation (Zinke et al., 2002; Harbison et al., 2005; Sorensen et al., 2007). This discrepancy might have been partially a result of the differences in experimental procedures and genetic backgrounds. But, it also suggested that acute response to starvation stress and the long term adaptation to starvation may rely on distinct mechanisms (Sorensen et al., 2007).

The studies discussed so far were all performed with adult flies. Even though genes affecting starvation resistance at the larval stage were not investigated in genetic screens, starvation responsive transcriptomic changes were examined also at the larval stage. Handke et al investigated the effects of food deprivation on the protein composition of the larval hemolymph at the L2-L3 boundary (Handke et al., 2013). Hemolymph is the circulating extracellular fluid, the

equivalent to blood in vertebrates, that is in direct contact with all internal organs and mediates the inter-organ communication by transporting metabolites, nutrients, hormones, etc. (Leopold & Perrimon, 2007). Handke et al. identified changes in the hemolymph protein content upon starvation which will support future data mining (Handke et al., 2013).

In another study, Zinke et al. investigated the changes in the whole-genome transcript abundances at the L2 stage upon complete starvation in PBS medium and amino acid starvation in 20% sucrose medium (Zinke et al., 2002). Strikingly, they found an upregulation of lipases only in completely starved larvae, whereas upregulation of genes encoding the key enzymes of fat biosynthesis and downregulation of lipases only occurred in sugar-fed larvae. In both starved and sugar-fed larvae, genes involved in protein synthesis were downregulated. Finally, it is worth to note that sugar-fed larvae could survive much longer than completely starved larvae (over a week vs. 2 days), even though both were starving for amino acids and other essential molecules (Britton & Edgar, 1998; Zinke et al., 2002).

Lastly, Li et al. studied the nutritional control of gene expression by starving mid-L2 larvae for certain time periods in 20% sucrose, and re-feeding them for some time (Li et al., 2010). They found starvation-induced reversible changes in almost a quarter of the transcriptome over time. They also analyzed the transcriptome changes induced by PI3K overexpression or by deletion of TOR function. They found that amino acid starvation-induced changes overlapped greatly with the TOR-mutation-induced changes but only slightly with changes induced by PI3K overexpression. Additionally, they identified a consensus Myc-binding site and another sequence motif in the promoters of starvation-regulated genes (Li et al., 2010). They had previously shown that Myc-regulated genes were altered in starved larvae (Grewal et al., 2005). These results corroborated the roles of TOR and Myc signaling pathways in the regulation of metabolism.

Although the functional genomics experiments described here provide important insights into the genes regulating starvation responses, they have certain limitations. For instance, reverse genetic screens suffer from the incomplete set of genes for which a transgenic line is available. Detrimental effects of mutagenesis and lack of tissue specificity also impose a major challenge. The tissue specificity problem might be overcome by the use of RNAi lines, for which a suitable tissue-specific driver line is available for most tissues, but this multiplies the time and effort spent for screening. Likewise, temporal regulation of knockdown (before or during starvation) might also be maintained by heat-shock induced drivers. Most expression profiling studies were conducted using whole body RNA; therefore tissue-specific changes were overlooked. Profiling

studies at early larval stages also suffered from inability to separate the males and the females. In profiling studies in which the non-starved control groups are the flies before starvation, starvation-induced changes in expression levels might be confounded with changes due to diurnal cycle. In the case of larvae, developmental changes might also interfere with the changes due to starvation. Therefore, age-matched controls must also be included in these studies. Nevertheless, the drawbacks of these large-scale studies might be resolved by studying candidate genes and pathways in more tailored experiments.

In addition to the classical genetics methods, *Drosophila* also provides the necessary tools and features to employ a population genetics approach. In classical genetics, genes are subjected to radical genetic manipulations (e.g., knock-out, knock-down, misexpression, overexpression, point mutation) that result in strong changes in gene functions ultimately leading to a phenotypic change. In contrast, in the population genetics approach, variants in the DNA sequence as they are found in natural populations are the tools to study the effects of genes on the phenotype. These natural variants cause only subtle changes in the phenotype, resulting in inter-individual phenotypic variation. One way to identify which genetic variants influence phenotypic variation is genome-wide association mapping. The next chapters will deal with the basics of this method, its advantages and shortcomings, and its applications in *Drosophila* to study quantitative traits such as starvation resistance.

3.6. Genome-wide association studies

3.6.1. Background on quantitative trait genetics

Many biological traits have complex etiologies. They are controlled by the actions of multiple genes, the environment and interactions between genes and the environment. In natural settings, variation in each of these determinants makes relatively small contributions to the phenotype, producing a continuous distribution of phenotypes among individuals. Hence, it is a difficult task to dissect the genetic basis of these quantitative traits. This task was accomplished mostly by linkage mapping studies before the introduction of genome-wide association studies (GWAS). Linkage mapping is based on tracing the transmission of phenotypes with known genetic markers along a pedigree (Dawn Teare & Barrett, 2005). Linkage studies successfully detected variants with large effects and high penetrance, but were not adequately powered to detect variants with small effects (Hirschhorn & Daly, 2005). Since the advent of

next-generation sequencing and genotyping technologies made genetic association studies feasible, they became a powerful alternative to linkage studies. GWASs search for co-occurrence of a genetic marker with a phenotype in a large sample of unrelated individuals (Figure 1). They are especially well suited to detect common variants with small contributions to the phenotype (Hirschhorn & Daly, 2005). Moreover, GWASs usually refine associated loci to 10-100 kb-long segments that typically comprise a few genes whereas linkage studies usually identify 5-10 Mb-long segments that typically comprise tens to hundreds of genes. This is because GWASs use large samples of unrelated individuals so that the associated loci are identical by state due to many historical recombination events, whereas in linkage studies associated loci are identical by descent due to relatively few very recent recombination events (Stranger et al., 2011).

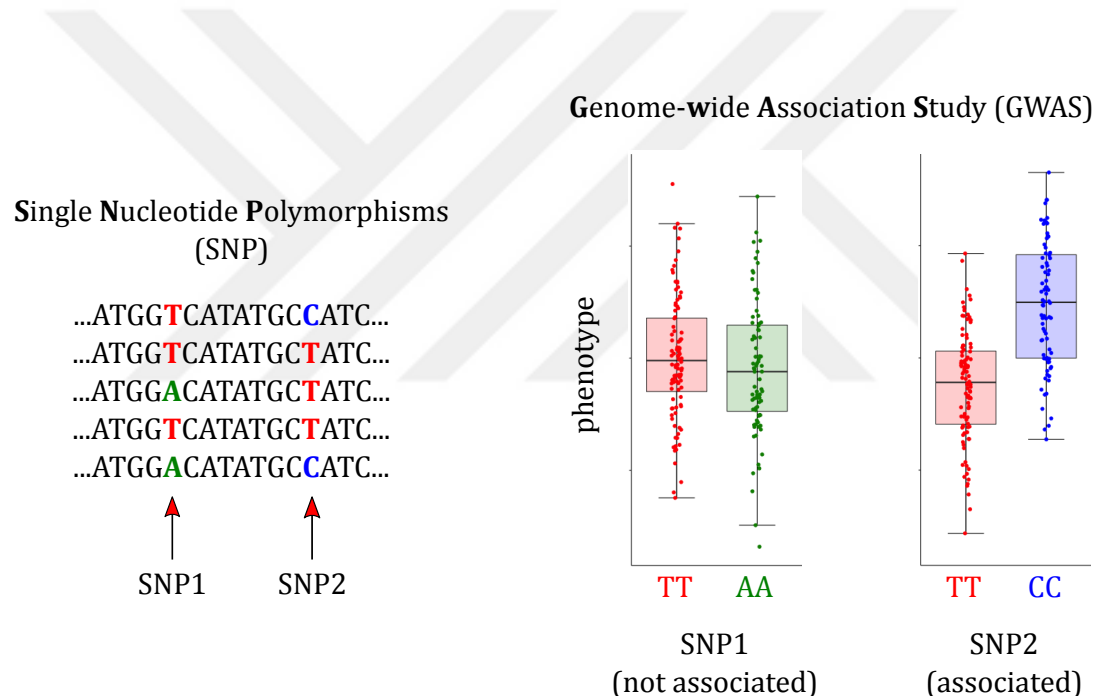


FIGURE 1. Genome-wide association study. This hypothetical example shows two SNPs, each with two alleles, tested for association to a quantitative phenotype. Individuals that have the T allele of SNP1 do not significantly differ in their phenotype from individuals that have the A allele of SNP1; therefore, SNP1 is not associated with the phenotype. In contrast, individuals that have the T allele of SNP2 and individuals that have the C allele of SNP2 significantly differ in their phenotype; therefore, SNP2 is associated with the phenotype. In a GWAS, the entire genome is screened for association. Heterozygotes are not shown for simplicity.

GWASs provide valuable insights into the genetic architecture underlying complex quantitative traits. They identify the number of loci involved in the trait, their relative contribution to phenotypic variation, the nature of alleles in these loci and their frequency in the population, the interactions between these alleles, and the heritability of the trait (Stranger et al., 2011). GWASs are hypothesis-generating tasks; they prioritize genomic loci and genes associated with the trait for further investigation. Functional characterization of these genes helps us to reveal the mechanisms underlying the studied trait (McCarthy & Hirschhorn, 2008). One major advantage of GWASs is that the whole genome, rather than specific candidate regions, is investigated for association (Kitsios & Zintzaras, 2009). As a result of this agnostic nature of the method, GWASs are able to identify previously unsuspected regions, such as loci that have genes with no anticipated function in the trait and loci in gene-depleted regions (McPherson et al., 2007). Such loci in gene deserts would never be considered in candidate gene-based studies. In this regard, GWASs are instrumental in revealing unprecedented mechanisms that underlie quantitative traits and diseases (Manolio et al., 2008). For example, several association studies increased the number of susceptibility loci from 3 to 71 (as of 2011) for Crohn's disease, a chronic inflammatory intestinal disease (Lee & Parkes, 2011). These studies identified genetic variants in autophagy-related genes (Hampe et al., 2007; Parkes et al., 2007), suggesting a link between autophagy and Crohn's disease. Autophagy is an intracellular degradation process that delivers parts of the cytoplasm to the lysosome. It had not been suspected in the pathophysiology of Crohn's disease before these GWASs. Many experiments showed later that defects in innate immune mechanisms, such as autophagic clearance of microbes, predispose to Crohn's disease, corroborating the involvement of autophagy in Crohn's disease (Lee & Parkes, 2011).

3.6.2. Basic concepts and methods used in GWAS

Genomes of individuals in a population differ from each other in many ways. There are substitutions in a single nucleotide at a specific position in the genome. These variations are called single nucleotide polymorphisms (SNPs), and the possible nucleotide variants existing in the studied population (e.g. TAGAGCT vs. TAGGGCT) are the alleles of these SNPs. Additionally, there are also so-called structural variations such as insertions/deletions (e.g. TAGAGCT vs. TAAGCT) and copy number variations (e.g. TAGAGCT vs. TAGAGAGAGCT). Most of these genetic variations are neutral; they do not affect phenotypes. However, some variants have effects on certain traits and are the reasons why individuals of a population assume different phenotypic values. GWASs can identify such variants if they are found more often than expected just by chance in individuals

with a particular phenotype (e.g. type 2 diabetes, higher body mass index (BMI), etc.). For categorical traits (e.g., healthy vs. affected as in case-control studies), associations are tested using chi-square or contingency-based tests and logistic regression tests. For quantitative traits (e.g., height, BMI, etc.), associations are tested using linear regressions.

Alleles of genetic variants that are in close proximity tend to co-occur across individuals; that is, they are in linkage disequilibrium (LD). LD is broken down by recombination events. In humans and other mammals, LD does not decay gradually with distance, but follows a block-like pattern (Daly et al., 2001; Guryev et al., 2006). Genetic polymorphisms segregate together within regions of diminished recombination (haplotype blocks) that are separated by recombination hot spots. Hence, a few marker SNPs can represent all of the common polymorphisms in each haplotype block. International genotyping consortiums such as HapMap (The International Hapmap Consortium, 2003) and 1000 Genomes Project (1000 Genomes Project Consortium, 2010; Auton et al., 2015) catalogued in large detail the genetic variations and the haplotype maps in human populations of different ethnicity. These studies enabled building genotyping arrays based on the most informative subset of SNPs for genetic variation across the genome. Furthermore, by comparing the alleles genotyped at marker SNPs to the densely genotyped reference haplotypes, alleles at un-genotyped SNPs can be predicted (Marchini & Howie, 2010). This is called genetic imputation and is used widely to increase the power and fine-map associations. It also enables meta-analyses of independent datasets that were genotyped on different array platforms (de Bakker et al., 2008). The large haplotype structure (100-200 kb) in humans facilitates genotyping efforts by greatly reducing the number of markers to be genotyped, but at the same time imposes a challenge to identify the causal variants and genes (de Bakker et al., 2005). In contrast to humans, LD in *Drosophila* decays very rapidly over short physical distances (Carbone et al., 2006). Therefore, the genetic diversity in *Drosophila* populations cannot be represented by a subset of marker SNPs, and genotypic information in all SNPs across the genome is necessary for GWASs. This increases the costs of association studies, but the reward is precise localization of causal variants (Mackay et al., 2009).

Typically thousands to millions of marker SNPs are tested in a GWAS. This large number imposes a challenge to derive a significance threshold that is corrected for multiple testing (Stranger et al., 2011). For a single test performed at the nominal 5% significance level, there is only 5% chance of incorrectly rejecting the null hypothesis (H_0 : no association) when the null hypothesis is true. However, when 1000 tests are performed, 50 of them (i.e., 5% of 1000) are expected to have p-values less than the significance threshold level of 0.05 just by chance, that is, even if the null hypothesis is true for all the tests (Sham &

Purcell, 2014). Therefore, significance thresholds at values typically used in small studies (i.e., $\alpha = 0.05$) would inevitably lead to many false positives when performing multiple tests. This necessitates correcting the significance threshold for multiple testing. One method of doing this is the Bonferroni correction, in which the new threshold is given by dividing the desired significance level by the total number of tests. For example, if 1000 SNPs are tested in a GWAS, the Bonferroni-corrected significance threshold would be $\alpha_{\text{corrected}} = 0.05/1000 = 5.e-05$. However, this method of correction is very stringent since not all the genetic variants are independent of each other (i.e., some of them are in LD; therefore tag the same locus). Therefore, less stringent methods (e.g. permutation-based statistical tests, etc.) might also be used.

The statistical power of GWAS, that is, identifying a true association, depends on several factors such as the effect size of the SNP, its allele frequency and the population size (Gibson, 2012; Korte & Farlow, 2013). The effect size of a SNP is the magnitude of phenotypic difference between its two alleles. The larger the difference is, the more the SNP contributes to overall phenotypic variation. The allele frequency of genetic variants is an important determinant for detection in the GWAS. The statistical power for the discovery of low frequency or rare variants is low if they do not have relatively large effects (Stranger et al., 2011). Since large haplotype blocks are represented by certain marker SNPs, the causal variant is almost always not the associated marker SNP itself, but is another genetic variant that is in LD with the marker. Hence, apart from the marker allele frequency, the causal allele frequency and its correlation with the marker allele are also important determinants for detection in a GWAS. Moreover, rare variants might also generate synthetic (false positive) associations when they are in strong association with other non-causal rare variants independent of their genomic locations (Dickson et al., 2010). An efficient way of identifying true rare variants is to utilize a candidate-based approach that involves extensive re-sequencing of candidate regions in selected cases and detailed functional assessment of these rare variants (Bodmer & Bonilla, 2008). In non-human organisms, rare variants with large effects can further be examined by linkage mapping after they are elevated to intermediate frequencies by crossing individuals that harbor them with reference individuals that do not (Korte & Farlow, 2013). Lastly, increasing the sample population size improves GWAS power to reliably detect SNPs that have smaller effect sizes and lower allele frequencies. The first GWASs used as little as a few hundreds of individuals (Ozaki et al., 2002; Klein et al., 2005), whereas most modern GWASs use as many as 200.000 individuals (Ehret et al., 2011). In parallel to this trend, the number of detected associations increased from a few to several tens. Meta-analyses of independent GWASs for the same trait also increase the power and can identify new associations (Stahl et al., 2010).

In a GWAS, relatively few loci are expected to have p-values below the stringent threshold. Therefore, the distributions of test statistics are inspected in Manhattan and Q-Q plots for any signs of systematic bias. Manhattan plots show the negative logarithm of p-value test statistics as a function of genomic location. Only the most significant associations are located above the threshold line that notifies the genome-wide significance level; therefore it is easy to visualize the number of significant associations in the GWAS. Q-Q (quantile-quantile) plots show the observed vs. expected distribution of the negative logarithm of p-values. The majority of the test statistics (corresponding to non-associated loci) are expected to align on the diagonal 1:1 line, and a few deviations at the end of the distribution correspond to the significant associations. Inflation of p-values in the GWAS due to a systemic bias, such as population structure (see below), becomes easily noticeable as an early deviation from the diagonal line (Bush & Moore, 2012).

In order to reduce the number of false positive associations in a GWAS (i.e., type I errors), potential complexities that lead to a systemic bias must be avoided (Korte & Farlow, 2013). Phenotypes must be quantified as precisely as possible using the same measurement methods for all individuals. Batch effects (i.e., experimental variations that stem from testing subjects in different groups using different batches of materials, chemicals, etc.) must be minimized or corrected for (Clayton et al., 2005). Individuals with extreme phenotypic values might add a bias to associations; they can be omitted from analysis. Genetic heterogeneity in the population is a major confounder when not corrected (Tian et al., 2008). The presence of subpopulations of different regional or ethnical background in a population is known as population stratification/structure, and can affect association results since allele frequencies differ between such subpopulations. For instance, when the numbers of cases and controls for each subpopulation are different, SNPs at which allele frequencies differ between subpopulations might result in significant associations even though they are not linked to the disease or trait examined in the GWAS. Population structure also leads to false negative results (i.e., type II errors). This happens when a true positive association for one subpopulation is diluted by other subpopulations in which the causal allele is absent (Stranger et al., 2011). Problems arising from population structure and also from cryptic relatedness among individuals are disentangled using statistical methods such as mixed models, which incorporate the pairwise genotypic similarities of individuals in the regression analysis (Sillanpää, 2011). Lastly, a good genotype data quality is a critical prerequisite for making GWASs. Base calling (i.e., assigning nucleotides/alleles to SNPs) must be accurate, and genotypes must be complete at the SNP and the participant (i.e., subject) level. In other words, the proportion of participants with a genotype call must be high for a SNP to be included in the GWAS, and the proportion of SNPs with a genotype call in the genome must be high for a participant to be included in the GWAS.

Incomplete genotypes are to be filtered out before association mapping (Tian et al., 2008), whereas SNPs with large deviations from Hardy-Weinberg equilibrium must not be eliminated and be flagged for further analysis since they might be indicative of potential genotyping errors, population structure but also of actual association to the trait (Turner et al., 2011). In sum, GWASs utilize many statistical tests to eliminate false positives and negatives, and identify the true positives with maximum efficiency.

3.6.3. Next steps after a GWAS

As pointed out earlier, GWASs are hypothesis-generating tasks and are followed by further investigations. The next step after a GWAS is usually to validate the associations by replicating in independent samples (McCarthy & Hirschhorn, 2008). Replications in populations of different ancestry would help to refine the location of the causal variants more sharply since different populations might have distinct patterns of LD (Cooper et al., 2008). In general, African populations have shorter LD blocks; hence they might be particularly helpful for fine mapping (Stranger et al., 2011). However, it should be noted that replication of associations might fail if disease etiologies or allele frequencies differ across populations of different ethnicities (Kochi et al., 2009).

As associations do not imply causation, GWAS results are devoid of causality by themselves. Therefore, the associated variants and loci must be investigated for a functional link to the trait. A naïve expectation is that causal variants affect the trait through their effects on nearby genes; hence, it is a common practice to assess the function of these candidate genes. However, in most cases, associated loci either encompass multiple genes or reside in gene-depleted regions; therefore, very often, it is not clear through which genes the associated loci exert their effects. Different strategies can be used to address this issue dependent on the particular case (Ioannidis et al., 2009). Prioritization of SNPs or genes based on their functional relevance might be helpful before progressing to functional experimentations. For instance, non-synonymous SNPs that result in radical amino acid substitutions are more likely to be causal variants than SNPs in non-coding regions. Likewise, some associated genes might have functions that are more likely to be involved in the studied trait. In order to infer the mechanistic basis of the studied trait, different types of datasets might be incorporated into association results. A common practice is to search for an enrichment of a particular biological function among candidate genes using gene ontology (GO) terms or protein-protein interaction data sets. Associated loci in gene deserts might be searched for sequence motifs such as enhancer regions, conserved sites, methylation/acetylation, etc. (Stranger et al., 2011). These analyses will ideally

narrow down the list of potential candidate SNPs and genes, or will give the first hunches to tackle in follow-up experiments.

An effective strategy to identify how associated variants function in the trait is the systems genetics approach (Civelek & Lusis, 2014). Genetic variants influence organismal phenotypes through intermediate molecular phenotypes, such as transcript or protein levels. Therefore, whole-genome transcript or protein levels can be quantified in individuals with known genotypes and phenotypes to be used in associations with the genotypic and phenotypic variations. Integration of these different types of associations fills the gap between genotypes and phenotypes establishing a link of causation based on the biological information flow. More specifically, correlations between genetic variants and transcript (or protein) abundances identify the genetic loci that affect expression levels. Correlations between expression levels and organismal phenotypes identify the transcripts (or proteins) whose levels affect the trait. Moreover, correlations among levels of transcripts (or proteins) themselves generate networks of co-regulated genes. Incorporation of these three correlations together with GWAS results explain the mechanism of action how the genetic variants affect the organismal phenotype (Mackay et al., 2009).

3.7. GWAS in model organisms

The success of initial GWASs in identifying genes associated with complex diseases provoked the idea of applying this methodology in model organisms. Genetic reference populations were formed for association mapping studies in many model organisms including *Arabidopsis thaliana* (Atwell et al., 2010), *Drosophila melanogaster* (King et al., 2012; Mackay et al., 2012; Huang et al., 2014), *Caenorhabditis elegans* (Rockman et al., 2010; Gaertner et al., 2012), mice (Churchill et al., 2004; Bennett et al., 2010; Threadgill et al., 2011; Collaborative Cross Consortium, 2012) and rats (Printz et al., 2003; Simonis et al., 2012). These populations are sets of inbred lines, or strains, each having a fixed and known genome. They are generated by crossing two or more inbred laboratory lines followed by repeated self- or sibling-mating, or by inbreeding directly from a natural population (Broman, 2005; Flint & Eskin, 2012). The resulting inbred lines are genotyped only once, but since they can be replicated continuously, they can be phenotyped many times. Genetic reference populations are community resources; they are accessible by all researchers to study diverse traits (Collaborative Cross Consortium, 2012).

Model organisms provide several advantages for association studies. The ability to strictly control environmental variables helps to maintain and treat all lines of

reference populations in the same manner. This increases the statistical power of genotype-phenotype associations by minimizing the contribution of environmental perturbations to phenotypic variation (Flint & Mackay, 2009). Measurements of multiple individuals of each inbred line provide an accurate estimate of the mean phenotypic value for each genotype and render association studies that employ inbred lines more powerful than those that employ outbred individuals. Consequently, GWASs in genetic reference populations require much smaller numbers of lines compared to human studies (Mackay et al., 2009; Korte & Farlow, 2013).

Genetic reference populations might be suitable tools to address many important questions that cannot be directly addressed with human subjects (Aitman et al., 2011; Weber et al., 2012). Additionally, they offer the opportunity for follow-up experiments to validate the initial findings of GWASs in very similar settings. They are very powerful tools for systems genetics studies. In theory, all tissues that are relevant to the studied trait or disease are accessible in model organisms for phenotyping, expression and molecular profiling. This is not the case in human subjects (Flint & Eskin, 2012). The inbred lines that show an extreme phenotype can also be used to construct mapping populations for linkage mapping analysis. Extreme and outlier phenotypes result from genetic variants with larger effects and lower frequencies; linkage mapping is the method of choice to identify such variants. Furthermore, careful control of the environment makes it possible to examine context-dependent effects of quantitative trait loci and genotype-by-environment interactions using genetic reference populations (Flint & Eskin, 2012). Importantly, many genetic, genomic and phenotypic data for diverse biological processes are collected on the same genetic reference population and these data might be integrated to infer more comprehensive understanding of the biological traits, their interactions with each other and the environment (Collaborative Cross Consortium, 2012).

3.8. The *Drosophila* Genetic Reference Panel

Drosophila melanogaster is a particularly suitable model organism for genome-wide association studies owing to its well characterized genome, ample amount of genetic variation and rapidly decaying local LD structure (Mackay & Anholt, 2006; Flint & Mackay, 2009; Mackay et al., 2012). The *Drosophila* Genetic Reference Panel (DGRP) was established at Trudy Mackay's initiative to serve as a community resource for association mapping studies. It is a set of 203 inbred lines derived from a single natural population on a farm market in Raleigh, North Carolina. Wild-caught females were used as founders, and their progeny were inbred to near homozygosity by full sibling mating for 20 generations (Figure 2).

The resulting 203 lines were sequenced and found to possess significant genetic variation; almost 1 in every 25 nucleotides is polymorphic in the DGRP population (Huang et al., 2014). Due to this extensive genetic diversity, DGRP lines showed significant phenotypic variations in any quantitative trait tested so far. These traits included adult starvation resistance (Mackay et al., 2012), metabolite contents and body weight (Jumbo-Lucioni et al., 2010; Unckless et al., 2015), food intake (Garlapow et al., 2015), oxidative stress resistance (Jordan et al., 2012; Weber et al., 2012), chill coma recovery (Huang et al., 2014), olfactory behavior (Arya et al., 2010), lifespan (Arya et al., 2010; Ivanov et al., 2015), organ size (Vonesch et al., 2016), startle response (Huang et al., 2014) and sleep (Harbison et al., 2013). Most of these traits were phenotyped at the adult stage, and there is still a lot to learn about larval traits. In addition to these organismal traits, whole-genome transcript abundances (Ayroles et al., 2009; Jumbo-Lucioni et al., 2010) and genotype-by-environment interactions (Unckless et al., 2015) were also studied using this population.

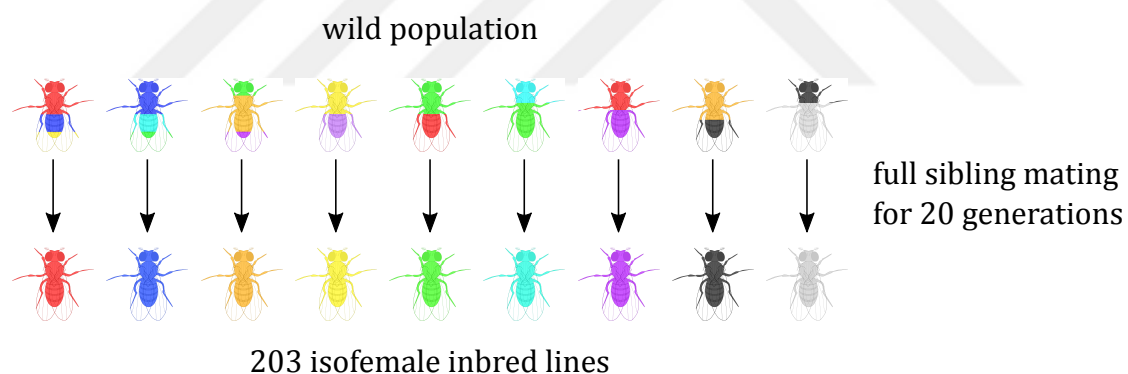


FIGURE 2. *Drosophila* Genetic Reference Panel (DGRP). DGRP is a community resource for population genomics studies; the lines and their sequences are available to all researchers. Figure adapted from Mackay Lab website, <https://mackay.gnets.ncsu.edu>.

GWASs on this population broadened our knowledge on the genetic architecture of complex quantitative traits. They revealed that hundreds of loci underlie most quantitative traits. Most of the associated variants lie within intergenic and intronic regions that are putative regulatory sites. Synonymous SNPs that influence a trait outnumber non-synonymous SNPs (Aitman et al., 2011). Associated variants typically comprise of low-frequency alleles with large effect sizes and intermediate-frequency alleles with smaller effect sizes (Flint & Mackay, 2009). This finding has an important implication on human genetics. Human GWASs frequently identify variants with small effects at intermediate frequency. Genotyping platforms and the statistical power in human GWASs are not optimal for identifying low frequency variants. Therefore, if humans have a similar genetic architecture for quantitative traits, unidentified low-frequency variants might contribute to the missing heritability observed in human GWASs (Aitman et al., 2011). It is frequently observed that genes found to be associated with *Drosophila* traits have homologs in humans that are known to play a role in a similar trait. One example is the association of *Sirt7* transcript levels with adult triglyceride contents in the DGRP (Jumbo-Lucioni et al., 2010). *Sirt7*, and sirtuins in general, are associated with obesity in mice and humans. Given the conservation of many important genes and pathways between *Drosophila* and humans, such insights contribute to the efforts in elucidating the genetic mechanisms underlying clinically relevant traits in humans (Mackay & Anholt, 2006). Interestingly, many of the associated variants lie within or nearby unknown genes or genes with no previously anticipated functions in the traits being tested. A previous GWAS in our lab for several *Drosophila* size traits found that associated variants did not map to the canonical growth related pathways, such as the insulin signaling pathway (Vonesch et al., 2016). However, these associated variants interacted epistatically with members of the known growth pathways. It was shown in a different study that epistatic interactions are highly prevalent in the genetic architecture of complex traits in the DGRP (Huang et al., 2012). In conclusion, the DGRP provides an important tool to study the complex genetic basis of quantitative traits in *Drosophila*.

3.9. Summary and project motivation

Food shortage, or starvation, is one of the most common environmental challenges in nature, and has important organismal, ecological and evolutionary impacts. Organisms can resist starvation to the extent that their energy expenditure is met by internal energy supplies. In this regard, starvation resistance is a complex quantitative trait that relies on the efficient regulation of energetic balance. Therefore, investigating starvation resistance provides

insights into the mechanisms that control nutrient signaling, metabolism and other aspects of energy homeostasis.

Drosophila melanogaster has been a lead model organism to study different aspects of starvation resistance at multiple levels. Population-level studies demonstrated correlations and trade-offs with other traits. Genetic screens identified genes that are involved in starvation resistance, and candidate gene-based studies characterized the molecular functions of these genes. Despite these advancements, molecular mechanisms that mediate resistance to starvation are largely unknown. Furthermore, the genes identified in classical genetic studies by analyzing the effects of their knockdown and knockout might be different than the genes that affect variation in starvation resistance in natural populations. Previous studies almost exclusively focused on adult starvation resistance; however, the genetic basis of larval starvation resistance might be different than that of adults due to the anatomical and physiological differences between the two stages. Therefore, we are still lacking a comprehensive understanding of the control of starvation resistance in *Drosophila*.

In this project, we use a genome-wide association approach to dissect the complex genetic basis of starvation resistance at the larval stage. GWASs investigate the phenotypic effects of natural alleles rather than artificially induced alleles that perturb gene functions more radically as used in classical genetic studies. Therefore, we aim to find the genes that contribute to the phenotypic variation in natural populations. GWASs are also unbiased in the sense that they are not restricted with the limitations of classical genetic studies such as imprecise control or lack of tissue specificity and temporal regulation of gene function. We investigate starvation resistance at an early larval stage, which is characterized by continuous feeding, rapid growth and development. Since larval stage is metabolically more dynamic and has different nutritional demands compared to the adult stage, the control of starvation resistance might also be different than that in the adult stage. Consequently, with this project, we expect to broaden our knowledge on the genetic and physiological mechanisms underlying starvation resistance. Given the evolutionary conservation of major physiological and cellular processes, the results of this project will likely have implications beyond *Drosophila*.

3.10. Aims of the thesis

In this thesis, we addressed the following questions / aimed the following tasks:

1. Standardize rearing and experimental conditions for proper quantification of larval starvation resistances.

In order to increase the power of association mapping, environmental confounders that contribute to the phenotypic variation must be minimized. We aimed to develop a culture regime and experimental setup in which we treat all *Drosophila* lines in the same way.

2. Determine whether there is genetic variation in the trait. Investigate other sources of variation.

Does larval starvation resistance vary across the DGRP lines? What is the heritability of this trait? In other words, what percentage of the total phenotypic variation can be attributed to genetic factors? What other factors can contribute to phenotypic variation?

3. Characterize the genetic architecture of larval starvation resistance in *Drosophila melanogaster*.

If there is genetic variation in our trait, we can identify the genetic variants that influence the trait by association mapping. We can then answer the following questions: How many genes or loci affect our phenotype? What are the effect sizes and frequencies of associated variants? What are the effects of the associated variants to the genes?

4. Characterize the molecular mechanisms underlying starvation resistance. What are the functions of genes that potentially play a role in starvation resistance? How do they compare to the genes identified in other population genetics studies on other traits? How do they compare to the genes that were identified in classical genetic studies? Do these genes have human homologs that are known to function in the regulation of energy intake/expenditure and that are associated with clinical traits and metabolic diseases?

5. Determine how larval starvation resistance is genetically correlated with other traits.

DGRP lines have so far been phenotyped for many traits, mostly at the adult stage. This project will identify the correlates and trade-offs of larval starvation resistance. Hopefully, it will also stimulate more research on larval traits in the DGRP population.

4. RESULTS

4.1. Optimizing the conditions for phenotyping larval starvation resistances

We assessed starvation resistances of *Drosophila* larvae by measuring their survival time in 20% sucrose in PBS. Whereas young larvae die within 1 day in plain PBS, they can survive for many days in sucrose solution. This is long enough for the differences between different genotypes to become apparent.

In all experiments, starvation resistances of larvae were individually quantified by keeping one larva per well in 96-well plates. Mean values were calculated for each sample, and the mean of all replicate sample means on the same genotype were recorded as the line mean. For most DGRP lines, we observed a certain survival pattern. There was a considerable amount of larvae that died within the first day after transfer to starvation medium. This was followed by lack of death for some time (1-3 days depending on the line) and the number of deaths increased again and peaked (Figure 3). This pattern suggested that the deaths within the first day were not due to starvation; rather, the larvae died because they were wounded during the transfer process. Therefore, we always excluded the deaths of the first viability counts (which were done 1 day after transfer) from further analysis.

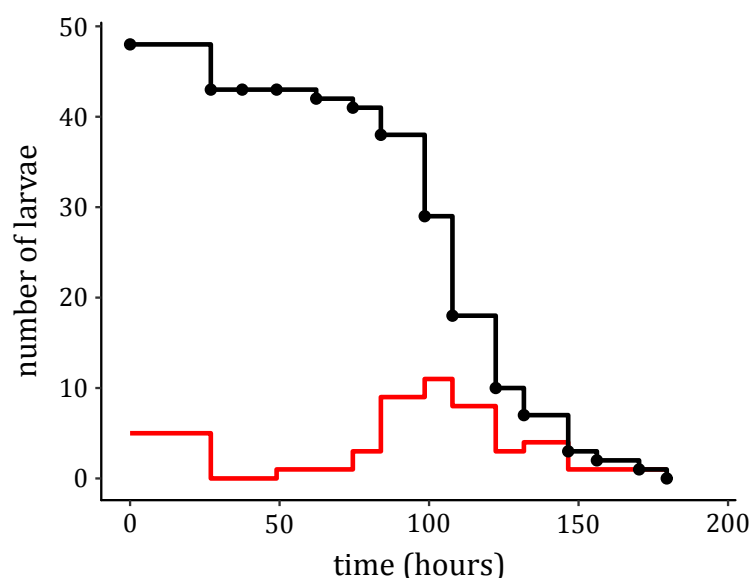


FIGURE 3. Survival curve for the DGRP line RAL_884. 36 hours old (AEL) larvae were starved in 20% sucrose solution in PBS (n=48). The number of survivors (black line) and the number of deaths (red line) were plotted as a

function of time. The deaths that occurred within the first day were most likely not due to starvation; therefore, they were omitted from further analysis, e.g., calculating the mean of the sample.

4.1.1. Starvation resistances during early larval stages

In order to determine the optimal larval age at which we would assess the starvation resistances, we randomly selected one DGRP line (RAL_379) and monitored its starvation resistance from mid-L1 to mid-L2 stage (34-60 hours after egg-laying (AEL)) (Table S1). We observed that both starvation resistance and wet weight increased exponentially as the larvae grew throughout the first instar (L1), but after development into the second instar (L2) at around 46-48 hours AEL, starvation resistance started to decrease gradually despite continued larval growth (Figure 4).

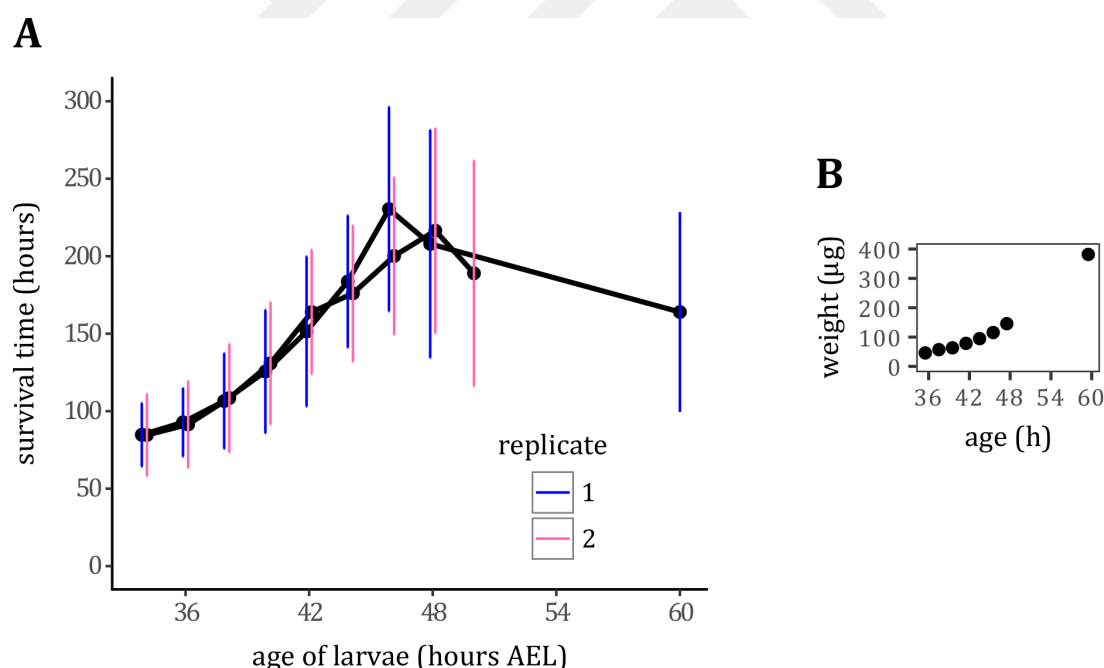


FIGURE 4. Starvation resistance as a function of larval age. (A) Larval starvation resistance of the DGRP line RAL_379 was measured between 34-60 hours AEL using larvae from the same parents in 2 successive days. Filled black circles denote means; error bars denote standard deviations. (B) RAL_379 larvae grew exponentially during the same time frame.

We noticed that as larval age increased, both intra-replicate variation and inter-replicate variation increased (Figure 4). The standard deviations of the replicate samples were tripled from mid-L1 to early L2 stage; they were roughly 3 days at around the first molt. Therefore, measuring starvation resistances on older larvae would decrease the precision in estimating the means for each sample and each line. Additionally, it is known that DGRP lines vary in their developmental times (Ellis et al., 2014; Rohde et al., 2016). Thus, the different lines might be at different developmental ages at a given time point after egg-laying, and these differences might be more profound with increased age. Therefore, quantifying larval traits at a later age would also increase the potential confounding effects arising from variation in their developmental states. To disentangle these problems and also to give larvae some time to feed and accumulate resources before starvation, we decided to start starvation at 36 hours AEL for future experiments.

4.1.2. Screening the DGRP lines

We optimized the culture conditions and experimental settings as detailed in Chapter 7 Materials and Methods, and phenotyped DGRP lines for larval starvation resistance (Figure 5). In summary, in each experimental block, we tested a total of 288 (48x2x3) larvae per line organized into 6 replicate samples (2 plates nested within every 3 days nested within one line/block). By using a large number of larvae, we increased the accuracy in estimating the mean phenotypes. Having been confined to 4 defined time slots for each day, we could phenotype at most 8 lines per block with 2 people performing the work. Hence, we had to screen the DGRP lines in many blocks. Since this might have led to unwanted phenotypic variation due to food batch heterogeneity, we tested many lines also in different blocks to detect if there is any discrepancy between blocks. More specifically, out of the 131 DGRP lines, 60 were re-phenotyped 2-4 times in a total of 60 blocks. We were able to reproduce the mean values within small deviations for the majority of the lines, despite the highly plastic nature of the larval starvation resistance trait. Only 2 lines showed a large difference between the repeats, and they were excluded from further analysis (data not shown). Other lines tested in the same blocks with these 2 lines produced highly similar mean values in other repeats, indicating that the observed deviations were specific to those 2 lines rather than the blocks themselves. To sum, our estimations for line means were overall highly reproducible, reflecting the strict environmental control in our experiments.

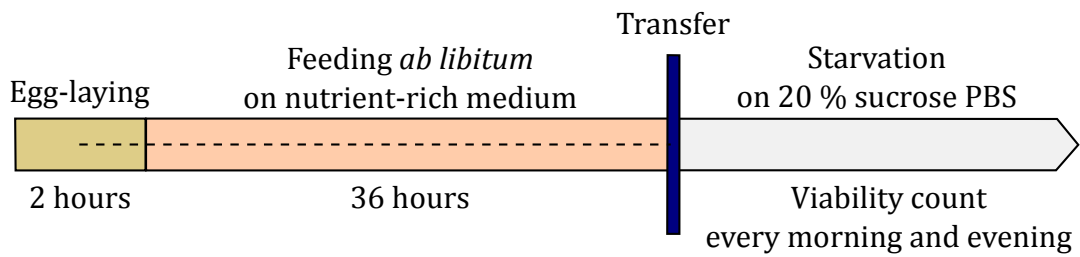
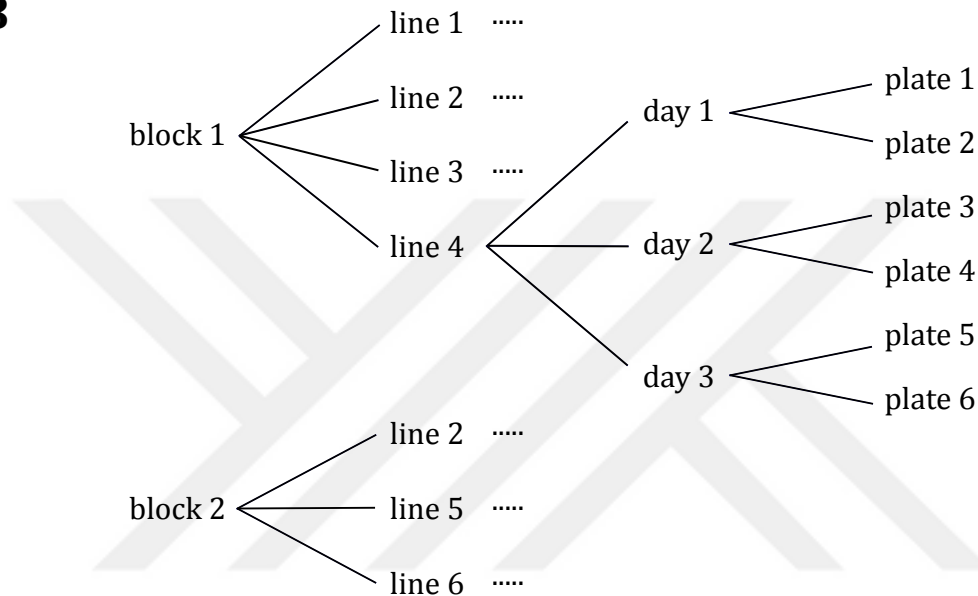
A**B**

FIGURE 5. Experimental setup for quantifying larval starvation resistances of DGRP lines. (A) Eggs were collected for 2 hours on apple agar plates with yeast paste on top. 36 hours later, L1 larvae were washed and transferred to the starvation medium in 96-well plates. Viability was counted every morning and evening starting 1 day after the transfer until all larvae died. (B) The phenotyping screen was performed in multiple blocks, with each line having at least 6 replicate samples organized into three days per block and two 96-well plates per day. Each replicate contained 48 larvae. Some lines were tested in more than one block to ensure reproducibility of the results.

4.2. Quantitative genetics analysis

We quantified the larval starvation resistance of 129 DGRP lines and found substantial genetic variation (Table S2, Figure 6). Line means varied between 72 and 167 hours (i.e., more than 2 folds of difference) with a grand mean of 110 hours, and a standard deviation of 17 hours. We analyzed the phenotypic

variation by partitioning it into its components using the linear regression model, $Y = u + \text{Line} + \text{Block} + \text{Day} + \text{Plate} + E$ (Table 1). The replicate variables block, day and plate altogether explained only less than 1.7% of the total variation, implying that we successfully controlled the macro-environmental variables that could contribute to phenotypic variation. We estimated that the broad sense heritability, i.e., the fraction of phenotypic variance that can be attributed to genetic variation, was $H^2 = 0.27$. Heritability was relatively low despite the large genetic variance among the DGRP lines, because the residual variation was considerably large: $V_E/V_P = 0.72$. To exemplify, the average standard deviation within replicate samples of the line RAL_884 is 25 hours (DGRP average is 27.5 hours) (Figure 3 and Table S12), meaning that 68% of larvae died within a range of 50 hours, or 2 days, and 95% of larvae died within a range of 100 hours, or 4 days (± 1 and ± 2 standard deviations, respectively). For comparison, the standard deviation between DGRP line means is 17 hours.

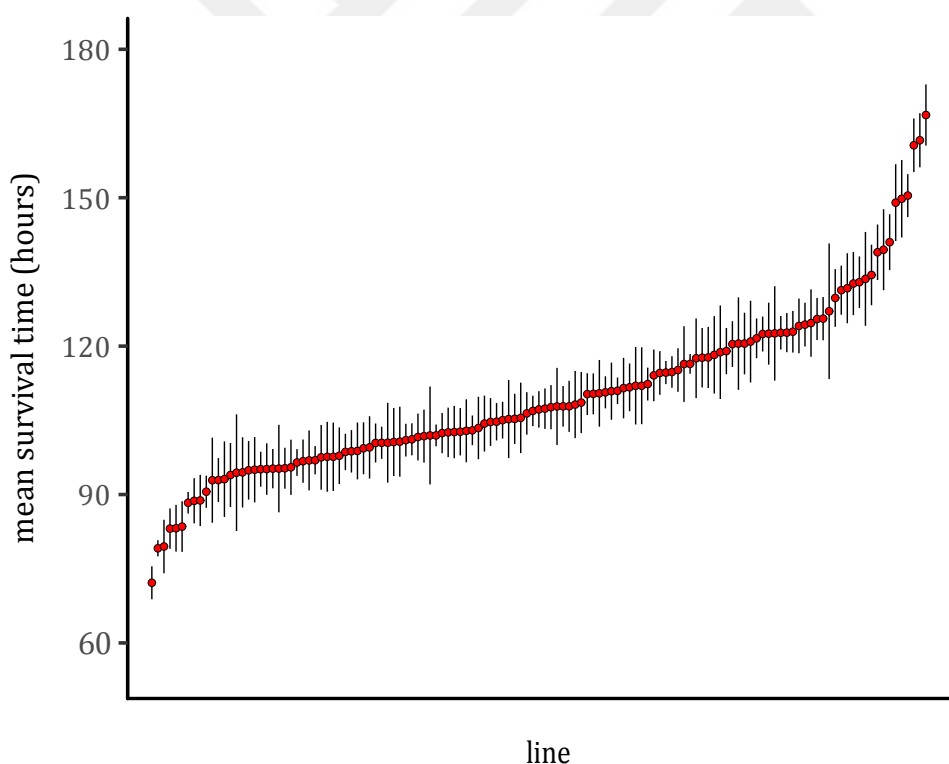


FIGURE 6. Variation in mean larval starvation resistance in the DGRP. Mean survival times of 129 DGRP lines under starvation are ranked from the lowest to the highest (red circles). Error bars represent the standard error between replicate means.

Source of variation	Variance component	Standard deviation	p value	Fraction of total phenotypic variance
line	303.892	17.432	<2e-16	0.267
block	6.007	2.451	7e-04	0.005
day	7.368	2.714	9e-11	0.007
plate	5.390	2.322	3e-09	0.005
error	816.351	28.572		0.717

TABLE 1. Quantitative genetic analysis of mean larval starvation resistance. Variances estimated by linear regression modeling using the dataset that contained 129 DGRP lines.

The large residual variation implies a high micro-environmental plasticity for larval starvation resistance (see also Chapter 4.7). Individuals with the same genotype, from the same parents, vials, and experiments could still display variable phenotypic values due to subtle differences in their microenvironments, and this is called micro-environmental plasticity (Morgante et al., 2015). For example, the location of eggs and larvae on apple-agar plates relative to the yeast, the edges, crowded regions, etc., might give rise to variation in the larval starvation resistances. Alternatively, the large residual variation might be the result of other factors that we could not account for or minimize in our setup. Given the high sensitivity of starvation resistance to larval age (Chapter 4.1.1), age might be one such variable. Another variable might be the sex of the larvae, a possibility that was tested later (see Chapter 4.5).

Prior to performing the GWAS, we tested whether *Wolbachia pipiensis* infection and five polymorphic chromosomal inversions had effects on larval starvation resistance. *Wolbachia* is an intracellular bacterium (an endosymbiont) that is maternally inherited. Many insect populations are infected with *Wolbachia*. 53% of all DGRP lines, and 55% of the (71 of the 129) DGRP lines that we phenotyped are infected. *Wolbachia* infection is known to influence some quantitative traits such as fecundity and viral resistance (Fry et al., 2004; Teixeira et al., 2008). However, we did not find any effect of *Wolbachia* on our trait (Table S3). Furthermore, we tested whether the presence of chromosomal inversions that were reported to be present in the DGRP lines would affect the trait. These inversions are very large chromosomal rearrangements that are polymorphic within a population and are found in many natural populations worldwide (Langley et al., 2012). In individuals heterozygous for an inversion, recombination is largely suppressed, especially immediately proximal to the breakpoints, preserving a haplotype structure. Therefore, inversions are the major sources of the genetic structure within populations (Corbett-Detig & Hartl, 2012). We detected a significant effect only for the In3R_Mo inversion

homozygosity (p-value=0.022) (Table S3). We also searched the effects of the interactions between *Wolbachia* infection status and inversions, but could not detect any effect. Therefore, we corrected the relevant line means by subtracting the effect of In3_Mo inversion homozygosity before proceeding with GWAS.

4.3. Genome-wide association mapping

In order to identify the genetic variants that affect the phenotypic variation in larval starvation resistance in the DGRP population, we performed a GWAS using In3R_Mo chromosomal inversion-corrected phenotypes. We used FaST-LMM, a linear mixed model that accounts for cryptic relatedness among genotypes (Lippert et al., 2011). Cryptic relatedness refers to the presence of close relatives in a study population that is not known to the investigator; it can considerably inflate false positive rates in small populations if not corrected for (Voight & Pritchard, 2005; Astle & Balding, 2009). Linear mixed models include a fixed effect for each SNP and a random effect for modeling the genetic kinship between each individual; this covariate structure in the regression model effectively resolves cryptic relatedness (Eu-ahsunthornwattana et al., 2014).

We tested a total of 1,408,196 SNPs with a minor allele frequency above 0.05 and identified 94 associated SNPs at a nominal significance threshold of $p=1.0e-05$ (Table S4). The associated SNPs are roughly spread over the entire genome as seen in the Manhattan plot (Figure 7A). However, some of them are very close to each other; they segregate together since recombination events have not taken place enough to break down the linkage between them. These SNPs are said to be in linkage disequilibrium (LD) (Figure 7B). As a result of the local LD structure between some of the associated SNPs, the 94 associated SNPs indeed tag 50 genomic loci in total. None of the associations reached the genome-wide significance level, i.e., the significance threshold at $3.6e-08$ after Bonferroni correction. Bonferroni correction is a very stringent method to correct for multiple testing and is based on the assumption that all tests are independent of each other, which is not true. Despite the fact that LD decays very fast with distance in *Drosophila*, very closely located SNPs might still be in linkage with each other (examples in Figure 7B). In addition, the presence of large chromosomal inversions also reduces independent segregation of alleles over long distances, especially around the inversion breakpoints. Therefore, not all tests in our GWAS are independent from each other, and, hence, the independence assumption for Bonferroni correction does not hold true. In this case, it leads to type II errors (i.e., false negative results). For this reason, we selected a nominal significance threshold of $p=1.e-05$, which has frequently been used in previous GWASs on this population.

The effect sizes of associated variants, i.e. the magnitude of the phenotypic difference between the two alleles of associated SNPs, correlated negatively with their allele frequencies. Low-frequency alleles had larger effects, and common alleles had smaller effects, producing a J-shaped distribution in the effect size vs. frequency plot (Figure 7C). This distribution was expected, since less frequent alleles need to have larger effects to be detected in association tests.

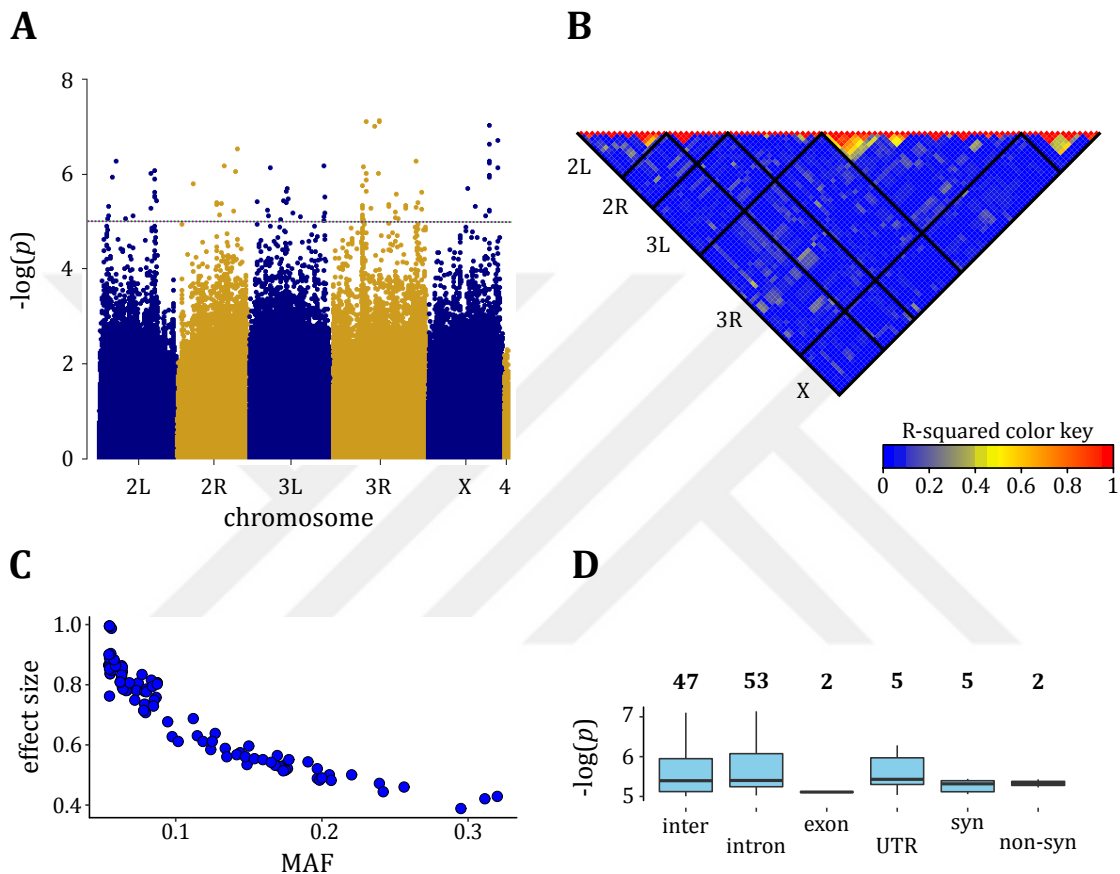


FIGURE 7. Genome-wide association analysis. (A) Manhattan plot shows the negative log₁₀ transformation of p values for all SNP tests across the genome. Dashed line denote the nominal significance threshold of $p=1.0e-05$. (B) Heat map shows the linkage disequilibrium (LD) between pairs of associated SNPs ($p < 1.0e-05$) as measured by r^2 , with the chromosome arms demarcated with black lines. The color code indicates the magnitude of LD, with blue corresponding to absence of LD and red corresponding to complete LD. (C) The scatter plot shows the effect sizes of associated SNPs plotted against their minor allele frequencies (MAF). Effect sizes (a/σ_p) were estimated as half of the difference between means of minor and major alleles relative to the population standard deviation. (D) Functional classes of associated SNPs are shown with the numbers of SNPs in each class. See also Chapter 4.8.4 for the Q-Q plot of the p values from this GWAS.

The majority of the associated SNPs were mapped to non-coding regions of the genome, mostly in the introns and intergenic regions. These loci are potentially involved in gene expression regulatory functions, implying that regulatory elements are largely responsible for the phenotypic variation in a population. We identified 7 SNPs that are located in the protein-coding sequences of genes. 5 of these polymorphisms result in codon changes for the same amino acids, and 2 of them result in codon changes for a different amino acid (Figure 7D). Therefore, missense SNPs, which cause changes in protein identities rather than abundances, also contribute to the phenotypic variation of traits.

Next, we searched for the genes located within 1 kb distance from the associated SNPs. We found 62 candidate genes, 54 of which are protein-coding (Table S4, Table S5). The majority of these protein-coding genes have predicted homologs in humans (Table S5). Some of these candidate genes are known to undertake important roles in the control of metabolism, feeding behavior and starvation responses, and some of them have human homologs associated with metabolic traits. Others have functions that render them plausible candidates for a role in starvation resistance, or functions that have no clear connection to starvation resistance. These candidate genes will be discussed in more detail in Chapter 5.9.

We performed a gene ontology (GO) enrichment analysis on the 54 protein-coding candidate genes using the STRING (Szklarczyk et al., 2015) and DAVID databases (Huang et al., 2009) to find whether a biological pathway was enriched among our candidate genes. STRING did not identify any enrichment whereas DAVID identified some nominally associated GO terms such as adult locomotory behavior ($p=1.4e-02$), neurogenesis ($p=1.5e-02$) and mRNA binding ($p=3.9e-02$). However, none of these terms passed the threshold of 0.05 when p values were corrected for multiple testing using the Benjamini-Hochberg method. This lack of enrichment for any biological entity is not unexpected, since starvation resistance has a very complex etiology and many biological processes are suspected to be involved in the expression of this trait.

In order to collect more information on the genetic architecture of larval starvation resistance, we selected the top 400 SNP associations and searched for the genes they are associated with. We found 213 protein-coding candidate genes, and examined them for genetic interactions and GO enrichment. 105 out of 213 genes participated in genetic interactions with each other (Figure 8). This network was enriched with GO terms such as organ development ($FDR=5.4e-04$), neuron development ($FDR=2.2e-03$) and taxis ($FDR=2.2e-03$) (Table S6). These interaction networks place the candidate genes in a context, which help us to gain insights into the genetic basis of larval starvation resistance.

4.4. SNP-based validation

As a proof-of-principle, we validated one of the most significant associations detected in the GWAS by replicating the association in an outbred genetic background. This association is *3R_12356551_SNP*, a biallelic SNP located in the second intron of *CG42342*, which encodes a transmembrane collagen (p-value = $9.9\text{e-}08$; Table S4, Figure 9A). We generated two outbred populations for the two alleles by round robin mating of the 12 DGRP lines carrying the minor (less frequent) allele (T) and randomly chosen 16 DGRP lines carrying the major (more frequent) allele (C), followed by 5 generations of advanced intercrossing. The resulting outbred populations contained the desired variant of the SNP, as confirmed by sequencing, but have randomized genetic backgrounds elsewhere (Table S7A, Figure 9B). We measured the larval starvation resistance of more than 350 individuals from each population and found a highly significant difference between these two populations (p-value= $7.6\text{e-}09$, Table S7B, Figure 9C). The minor allele population survived starvation in average 13.1 hours more than the major allele population. In other words, the minor allele showed association with starvation resistance in the same direction as predicted in the GWAS. Therefore, we were able to validate the association of this SNP in another genetic background.

This target SNP is located in the second intron of *CG42342*. Hence, we examined if the target SNP affected larval starvation resistance by changing the transcript abundances of *CG42342*. We performed quantitative RT-PCR using whole body RNA isolates from the two populations at different times during starvation in sucrose solution. We could not detect a significant difference in the expression levels of *CG42342* relative to *β -actin* between the two populations up to 10 hours of starvation (Table S7C, Figure 9D). However, we observed that expression levels in both populations almost tripled within this time frame, suggesting that the expression of *CG42342* is responsive to starvation in sucrose. Whether amino acid scarcity or sucrose abundance triggers this change has to be elucidated, but the sensitivity of the gene expression of *CG42342* to the medium we used to test the organismal phenotype strengthens the involvement of this gene in the expression of the trait.

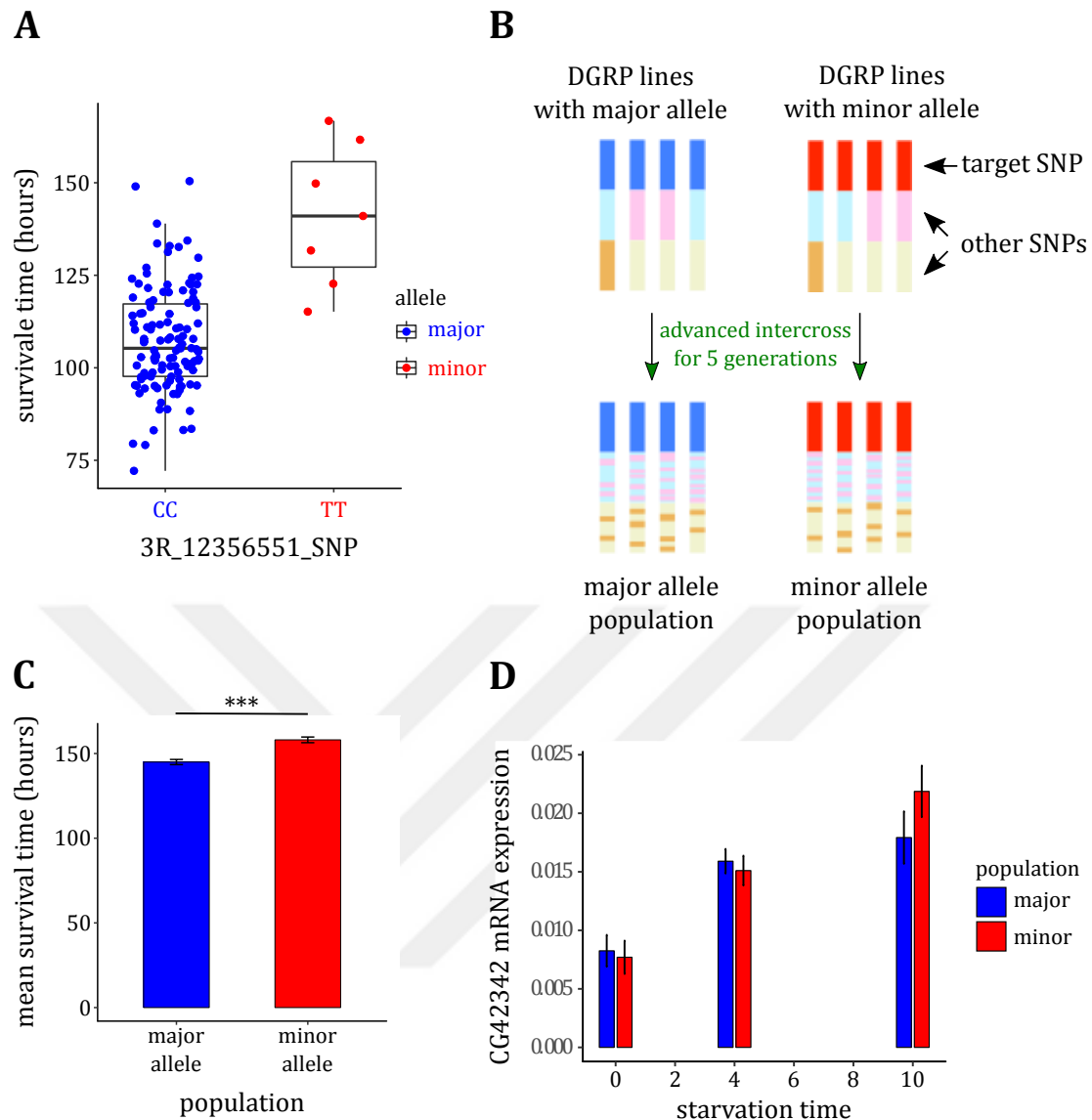


FIGURE 9. Replication of the association of 3R_12356551_SNP. (A) The minor allele of 3R_12356551_SNP is associated with higher larval starvation resistance in the DGRP population. (B) Generation of outbreeding populations homozygous for the major (CC) and minor alleles (TT) of 3R_12356551_SNP. Only 4 lines are shown for simplicity. (C) 3R_12356551_SNP is also associated in outbred background. Bar heights and error bars represent mean larval starvation resistance and standard error, respectively (145.0 ± 1.5 hours for the major allele; 158.1 ± 1.7 hours for the minor allele). ***: p -value= $7.6e-09$. (D) Expression of *CG42342* mRNA increases relative to β -actin (*Act5C*) mRNA during starvation. Bar heights and error bars represent mean relative expression level and standard error, respectively. No significant difference was detected between minor and major allele populations in all three time points.

4.5. Sexual dimorphism in larval starvation resistance

Using L1 larvae prevented us from quantifying the starvation resistances of males and females separately since there is no distinguishable morphological difference at this stage. However, the two sexes might still have different starvation resistances. In order to determine whether larval starvation resistance shows sexual dimorphism, that is, whether males and females tolerate starvation for significantly different durations, we re-tested 10 DGRP lines by determining the sex of dead larvae using multiplex PCR. Simultaneous amplification of specific autosomal and Y-chromosome regions allowed us to determine the sex of larvae based on whether they produced one or two bands. We found that 4 out of 10 DGRP lines showed significant sexual dimorphism whereby males were more resistant than females (Table S8, Figure 10). Taking into account all lines tested, males survived around 6.8 hours longer than females in average. The difference between sexes was as high as 14.4 hours in one DGRP line (RAL_787).

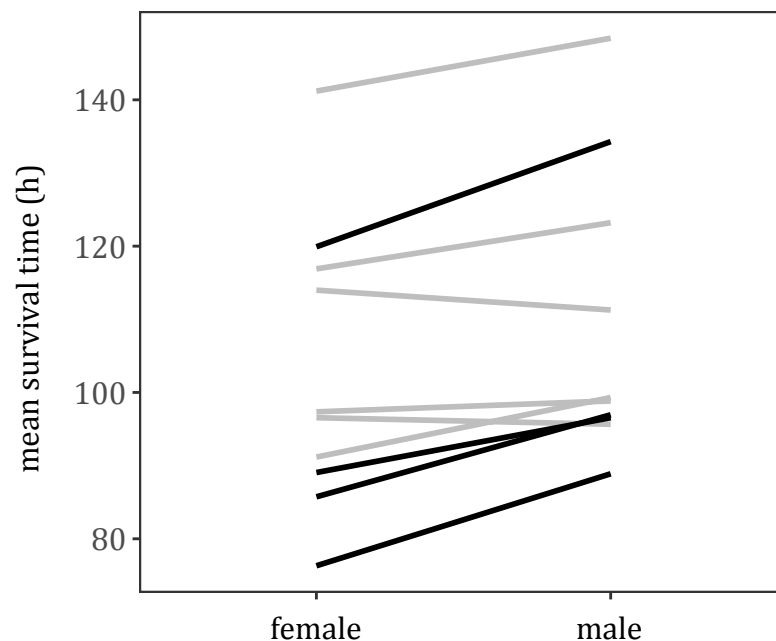


FIGURE 10. Sexual dimorphism in larval starvation resistance. Reaction norms between the mean survival times of males and females in 10 DGRP lines are depicted. Black lines represent significant differences, whereas gray lines represent lack of significant differences.

We partitioned the total phenotypic variance in this small dataset into its components using the linear mixed model $Y = \mu + S + L + L \times S + R + \varepsilon$, where Y is the phenotype, S is the fixed effect of sex, L is the random main effect of line, $L \times S$ is the random effect of line-by-sex interaction, and R is the random effect of replicate (Table 2). We detected significant effects of sex (6.8 hours, $p=0.0052$) and line ($p=5e-06$), but did not identify genotype-by-sex interaction as a significant contributor ($p=0.26$) to total variance in this dataset. The cross-sex genetic correlation, i.e., the correlation between males and females across DGRP lines, was very high and close to unity ($r_{GMF}=0.958$, $p=1.3e-05$). Hence, genotype-by-sex interaction, which is defined as the variation in the magnitude of sexual dimorphism across DGRP lines and is manifested as the deviation of cross-sex genetic correlation from unity, was very small and negligible for our dataset. These findings are also evident by the relatively low complexity of the starvation resistance vs. sex graph: The reaction norms that show the change in starvation resistance from females to males for each line do not cross over extensively (Figure 10). However, it should be noted that these results derived from only 10 DGRP lines, and cannot be generalized to the whole DGRP population.

Source of variation	Variance component	Standard deviation	p-value	Fraction of total phenotypic variation
Sex	fixed		0.0052 **	
Line	364.387	19.089	5e-06 ***	0.2850
Line x Sex	6.666	2.582	0.26	0.0052
Replicate	13.415	3.663	0.03 *	0.0105
Error	894.022	29.900		0.6993

TABLE 2. Quantitative genetic analysis on the sexed dataset. Variances are estimated by linear mixed model $Y = \mu + S + L + L \times S + R + \varepsilon$. See Table S9 for a comparison to the analyses of the same data with alternative models.

In conclusion, we showed evidence for sexual dimorphism in larval starvation resistance. However, we could not take the effects of sex into account for our GWAS since the phenotypes were from pooled sexes. Due to our small sample size, we also cannot infer to what extent sexual dimorphism and genotype-by-sex interaction exists in the whole DGRP population.

4.6. Correlations with other traits

In order to get insights into the underlying genetic mechanisms of larval starvation resistance, we examined correlations between larval starvation resistance and other traits that were previously measured in the DGRP (Table S11, Figure 11). Correlations between two quantitative traits suggest that they are partially controlled by common genes and pathways. The most significant correlation was a positive correlation with the triglyceride content of male flies reared on high glucose diet ($r=0.282$, $p\text{-value}=0.0057$) (Unckless et al., 2015). There were also a modest positive correlation with male alcohol sensitivity (Morozova et al., 2015), and a modest negative correlation with negative geotaxis activity of females exposed to chronic mild oxidative stress (MSB) (Jordan et al., 2012). We also found positive correlations with male and female waking activities (Harbison et al., 2013), and negative correlations with male and female hexanal avoidance scores (Arya et al., 2015). However, these correlations were significant due to a single outlier line, and are lost when this outlier is omitted. Therefore, we regard these relationships as not significant. Remarkably, adult starvation resistance (Mackay et al., 2012) did not show any significant correlation with the larval starvation resistance. This held true also when we examined the 10 DGRP lines that were sexed for larval starvation resistance (Table S10, Figure 11). The biological significance of these relationships will be discussed in Chapter 5.3.

FIGURE 11. Correlations of larval starvation resistance with other traits in the DGRP population. (page 55). The first 10 scatterplots show how larval starvation resistance (LSR) correlates with some of the adult traits among the 129 DGRP lines we phenotyped as pooled sexes, whereas the last 2 plots show how larval and adult starvation resistances correlate in males (m) and females (f) among the 10 DGRP lines we phenotyped as separate sexes. From the top left to bottom right (row, column) the traits compared to LSR are triglyceride (TG) content in flies reared on high glucose medium (1,1), waking activity (1,2)-(1,3), negative geotaxis activity under chronic oxidative stress (MSB treatment) (2,1),...

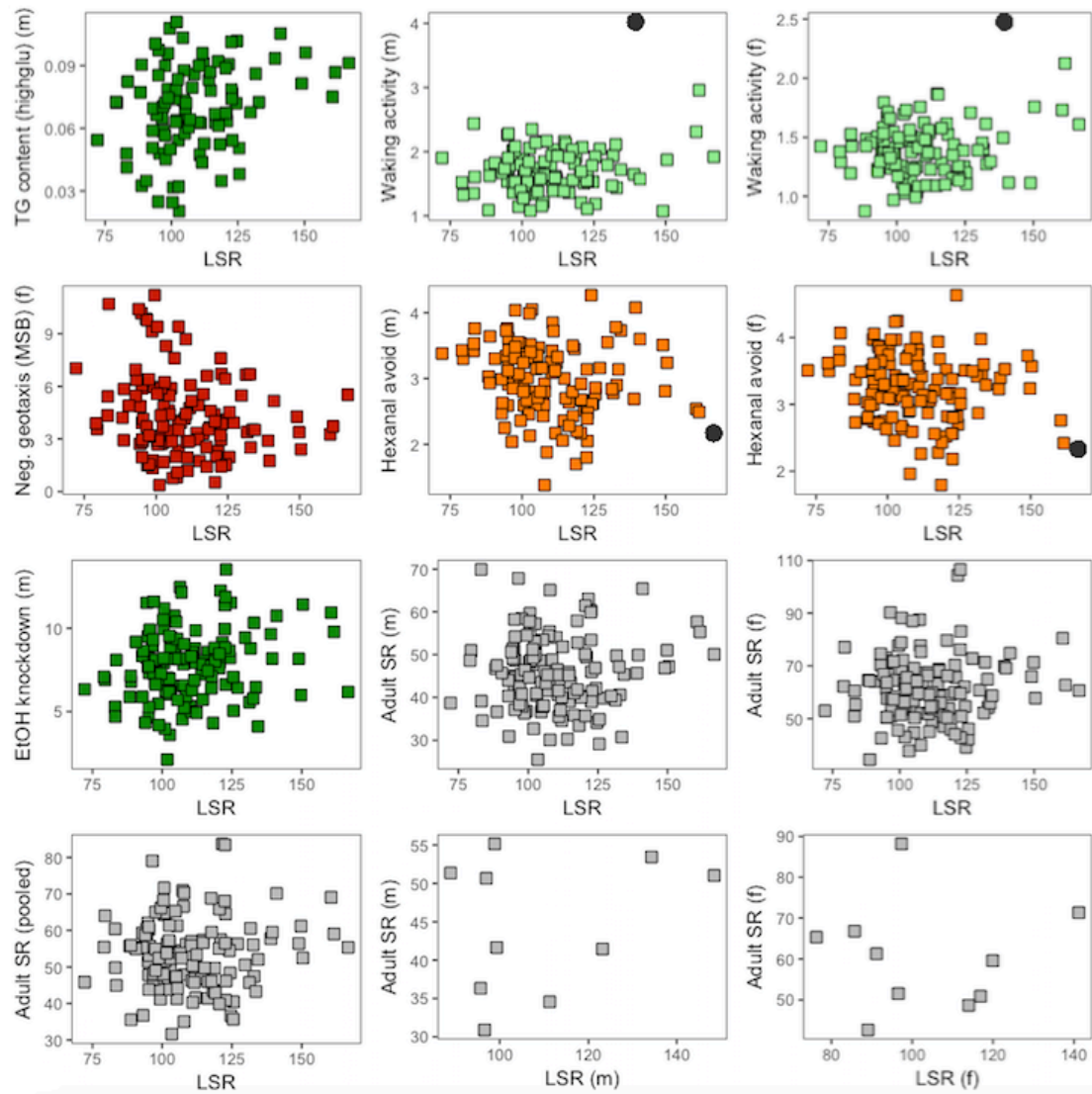


FIGURE 11. ... hexanal avoidance score (2,2)-(2,3), ethanol knockdown time (3,1), and adult starvation resistance (3,2)-(4,4). Significant positive and negative correlations are denoted in dark green and dark red, respectively, whereas lacks of correlations are denoted in gray. The positive correlation of waking activity (light green) and the negative correlation of hexanal avoidance (orange) were due to a single outlier line (marked with a dark gray filled circle). We regard them as no correlation since the correlations disappear when the outliers are removed.

4.7. Micro-environmental plasticity of larval starvation resistance

We observed overall a large residual variation, or micro-environmental plasticity, for larval starvation resistance. The magnitude of this variation was not similar between DGRP lines; individuals of some lines had survival times closer to each other, whereas individuals of others had widely varying survival times. Therefore, there is genetic heterogeneity in the micro-environmental plasticity of larval starvation resistance.

In order to study the micro-environmental plasticity in more detail, we used different metrics to quantify inter-individual spread. These metrics are within-line, within-replicate estimates of standard deviation (σ_E), coefficient of variation (CV_E), median absolute deviation (MAD), and their natural logarithms ($\ln(\sigma_E)$, $\ln(CV_E)$, $\ln(MAD)$). Each of these metrics was first estimated for each replicate sample, and then these replicate values were averaged to find the line means (Table S12). We found that there is a large genetic variation in micro-environmental plasticity, independent of the metric used (Table 4, Figure 12). For example, the mean σ_E of DGRP lines varied between 16 and 44 hours, which is almost a 3-fold difference. Using the natural logarithmic transformation for the metric of inter-individual variation generated normal distributions for micro-environmental plasticity of larval starvation resistance. Hence we decided to proceed with logarithmic metrics.

Metric	Summary Statistics							
	Min	1st Q.	Median	3rd Q.	Max	Mean	Std	$H_0: \sim N(\mu, \sigma^2)$ p value
σ_E	16.05	22.86	26.5	31.29	44.13	27.49	5.88	0.002 **
$\ln(\sigma_E)$	2.77	3.12	3.26	3.44	3.78	3.28	0.21	0.237
CV	16.06	21.93	24.70	28.06	39.37	25.11	4.47	0.010 *
$\ln(CV_E)$	2.77	3.08	3.20	3.33	3.67	3.20	0.18	0.890
MAD	9.42	14.79	17.12	21.08	29.75	18.07	4.26	0.014 *
$\ln(MAD)$	2.25	2.67	2.80	3.019	3.38	2.84	0.23	0.286

TABLE 4. Summary statistics for micro-environmental plasticity. All values and estimates are derived from line means. p values are given for Shapiro-Wilk normality tests. Std: standard deviation of the metrics for the DGRP population.

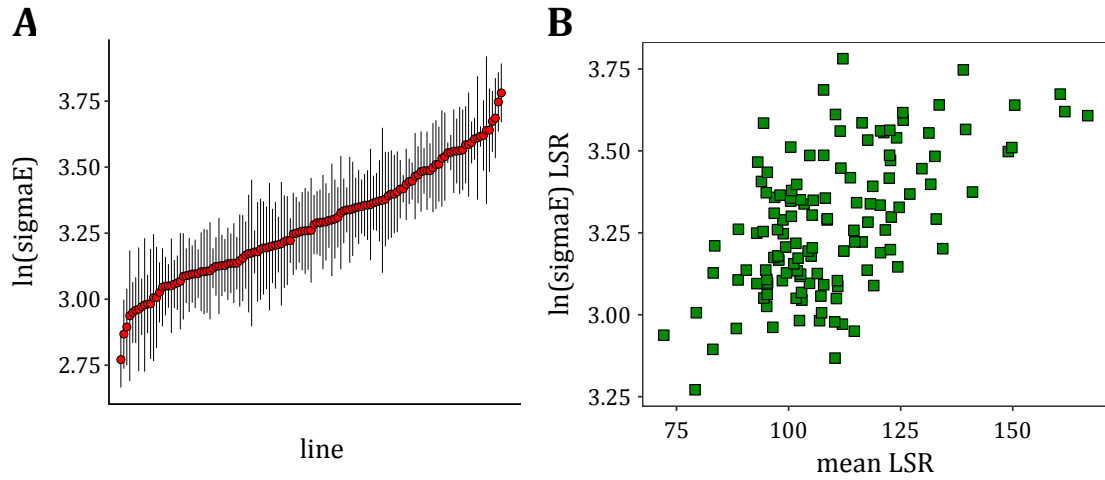


FIGURE 12. Variation in micro-environmental plasticity of larval starvation resistance in the DGRP. (A) Micro-environmental plasticity for larval starvation resistance, estimated by $\ln(\sigma_E)$, shows substantial genetic variation in the DGRP. red circles denote line means, error bars represent standard deviations among $\ln(\sigma_E)$ estimations from the replicates. (B) Micro-environmental plasticity of LSR shows a positive correlation with mean LSR ($r=0.573$, $p=1.2e-12$).

In order to estimate the heritability of micro-environmental plasticity, we partitioned the total phenotypic variance using the linear regression model $Y = \mu + L + \epsilon$ where Y is the phenotype, μ is the grand mean for DGRP, L is the random effect of line, and ϵ is the error, or residual variation between estimates of plasticity among replicate samples. We did not include the replicate factors block and day in our model since the number of replicate samples was very small (2-8 lines per block, 6 samples per line per block; 2 plates for every 3 days), meaning that the number of levels for each factor was very small. We estimated a high broad sense heritability for the metrics based on standard deviation and coefficient of variation, but not on median absolute deviation; $H^2 = 0.63$ for $\ln(\sigma_E)$, $H^2 = 0.56$ for $\ln(CV_E)$ and $H^2 = 0.37$ for $\ln(MAD)$ (Table S13). The reason why $\ln(MAD)$ resulted in a lower heritability is that the use of median values canalized most of the replicate values around two points that reflect the intervals between times of viability count (data not shown).

Next, we searched whether within-line variation was due to residual heterozygosity that remained in the DGRP lines after 20 generations of inbreeding. If this were the case, then we would expect that within-line variation increased with the number of segregating variants in each line. However, we could not detect any correlation between $\ln(\sigma_E)$ and the percentage of remaining segregating variants (Table S14). Likewise, micro-environmental plasticity did

not correlate with the number of polymorphic chromosomal inversions in the heterozygous state, which are the major sources of heterozygosity, and with the genome size. We also found that micro-environmental plasticity correlated positively with the mean larval starvation resistance, when either $\ln(\sigma_E)$ or $\ln(\text{MAD})$ was used as the metric, however this relationship disappeared when $\ln(\text{CV}_E)$ was used to correct for the means (Table S14).

Since there is substantial genetic variation for the micro-environmental plasticity of larval starvation resistance, we decided to determine the genetic loci affecting this genetic variation in this trait. We performed a GWAS using $\ln(\sigma_E)$ as the metric of choice. We reasoned that, compared to the calculation of median absolute deviation, standard deviation calculation gives more weight to values far from the center, hence, is more sensitive to the spread of data. Moreover, $\ln(\sigma_E)$ was normally distributed and hence superior to σ_E . Lastly, for GWAS, we do not need metrics that correct for the means. Therefore, we proceeded using $\ln(\sigma_E)$.

Next, we used a linear regression analysis to investigate whether *Wolbachia* infection or chromosomal inversions affect micro-environmental plasticity. Neither of them did (data not shown), and we proceeded with GWAS using the raw estimates of $\ln(\sigma_E)$ for each line. We performed the GWAS with all common variants ($\text{MAF} > 0.05$) using FaST-LMM, a linear mixed model that accounts for the effects of cryptic polygenic relatedness (Lippert et al., 2011). We found 13 SNPs that reached the same significance threshold used in the mean GWAS (1.0×10^{-5}). For a better examination of the genetic basis of our trait, we selected the top 100 hits in the GWAS (Table S15), and searched for the genes that are located within 1kb distance to them. A total of 63 candidate genes was found, 58 of which were protein coding. Some of these genes are plausible candidates for plasticity of larval starvation resistance, as they are involved in starvation responses (triglyceride lipases, autophagy proteins, etc.) and protein homeostasis (translation initiation, tRNA synthetases, degradation, etc.). We searched for enrichment of GO terms by analyzing the known or predicted interactions of these candidate genes using the DAVID and STRING databases; however, we could not identify any GO enrichment at a significant level (data not shown).

A positive correlation between mean and variance larval starvation resistance indicates that underlying mechanisms partially overlap. Therefore, we expected that candidate genes from both mean and variance GWASs participate in the same networks. We tested this expectation by GO enrichment analysis; however, using the combination of genes from both mean and variance GWASs did not generate any significant GO term (data not shown). Despite this, we found one quantitative trait locus that affects both traits. 4 SNPs (2R_11506182_SNP,

2R_11506370_SNP, 2R_11508395_SNP and 2R_11508889_SNP) located in the intronic region of a predicted gene, *CG42524*, were associated with mean larval starvation resistance. Another SNP (2R_11520690_SNP), farther away from these 4 SNPs, but still in LD with them, and also located in the intron of *CG42524*, was associated with micro-environmental variance of larval starvation resistance. These associations prioritize *CG42524* for a functional validation in larval starvation resistance.

4.8. Supplementary results

4.8.1. Sex ratio distortion as a minor confounder for phenotyping pooled sexes

As detailed in Chapter 4.5, we identified the sex of larvae in 10 phenotyped DGRP lines to assess whether sex is an important variable that contributes to the phenotypic variation. We also observed that the numbers of males and females were significantly different for some lines (Table S16). This difference was too large to be explained by chance under the assumption that female to male ratio is 1. For instance, in some samples, female larvae were twice as many as males and vice versa. There are two possible explanations for the observed sex ratio distortion. Firstly, some DGRP line produced progeny of one sex more than of the other. Secondly, we selected one sex over the other while transferring them to the starvation medium. We tested both hypotheses using the line RAL_502, which had more females than males in all replicates. To test the first hypothesis, we examined the numbers of both sexes at different stages of development. The results were inconsistent; we detected a statistically significant bias for more females only at late L3 stage, but not at egg, pupal and adult stages (Table S17A). Therefore, our results were not conclusive to prove or disprove that RAL_502 produced more females. To test the second hypothesis, we selected two groups of 36 hours old (AEL) larvae, one group containing individuals that we would normally include in the larval starvation assay and the other group containing individuals that we would not include as judged by their smaller sizes. We compared the numbers of both sexes in each group. The results showed a significant bias towards preference of females over males in the former group but not in the later (Table S17B). Therefore, the sex ratio distortion in our assay was partially due to sampling bias introduced by us. Our criterion to select larvae for the starvation assay was based on judging the average size of all larvae in the sample, ignoring unusually small and unusually large larvae, and selecting the healthy-looking (i.e., mobile) larvae. Individuals with smaller sizes might have hatched later, or might have been developmentally delayed or sick. Therefore, by

excluding smaller larvae, we aimed at avoiding a bigger sampling error. However, our data shows that size differences exist already at the mid-L1 stage, at least for some of the lines. These differences might result in distortion of sex ratio for some lines and might also lead to confounding variation when these lines are also sexually dimorphic. However, this source of variation is negligible, since line means from pooled sexes and from averaged sexes did not differ considerably (Table S8).

4.8.2. Bimodal death curves are not due to sexual dimorphism

We observed that a few DGRP lines had a bimodal distribution of the number of deaths over time. These DGRP lines were re-phenotyped; however, their bimodal death pattern was evident across replicate samples within and between blocks indicating that this pattern was specific to those lines. We included these lines in our test of sexual dimorphism to find if their bimodal distribution stems from large differences in male and female survival times. We observed that this was not the case; males and females had both bimodal death curves with similar mean values (Figure 13). In conclusion, despite sexual dimorphism in larval starvation resistance in the DGRP population (see Chapter 4.5), our data also show that sexual dimorphism cannot explain the bimodal distributions of the larval survival times.

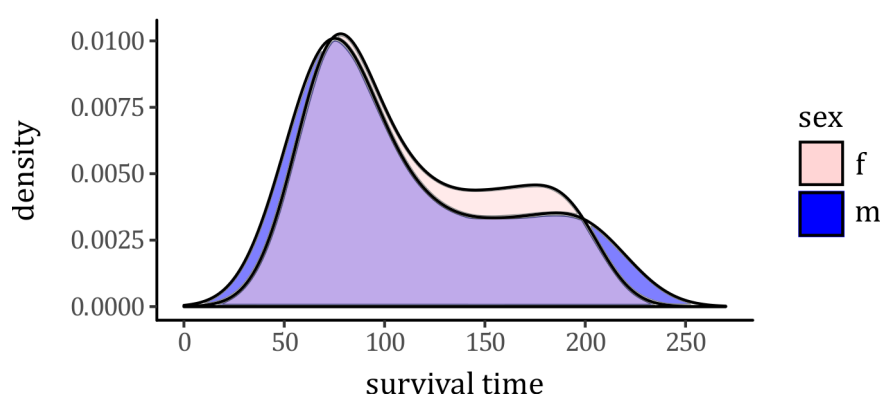


FIGURE 13. Bimodal death curve for line RAL_57. Females (f) and males (m) have highly similar bimodal distributions, means and standard deviations (females: 114 ± 45 hours, males: 111 ± 50 hours).

4.8.3. Sexual dimorphism in adult starvation resistance

Adult starvation resistances of the DGRP lines were quantified previously, and the line means are publicly available on the DGRP2 website (Mackay et al., 2012). The trait is sexually dimorphic; females are more resistant than males in the majority of the lines (Table S10). The cross-sex genetic correlation, which indicates to which extent the same variants affect a trait in males and females (Weber et al., 2012), is high, but still significantly different from unity in the whole adult dataset that contained 203 lines: $r_{\text{GMF}}=0.685$ ($p<2.2\text{e-}16$) (Table S10, Figure 14A). On the other hand, the same 10 lines that we sexed in the larval starvation assay described in Chapter 4.5 produce a higher estimate: $r_{\text{GMF}}=0.854$ ($p=1.7\text{e-}03$) (Figure 14B). A significant genotype-by-sex interaction effect was detected for adult starvation resistance in the DGRP population (Mackay et al., 2012). In accordance with these results, we ran a GWAS using the adult data available online and detected both sex-specific and common variants associated with adult starvation resistance (Figure 14C). Despite that there are only 4 associations common to both sexes, many sex-specific associations also have relatively low p values in the other sex albeit being above the significance threshold. Such variants might also be identified in association tests that use the averaged, or pooled, phenotypes. Therefore, cross-sex genetic correlation is an indicator for the extent to which a GWAS that uses sexes-averaged line means can identify sex-specific variants.

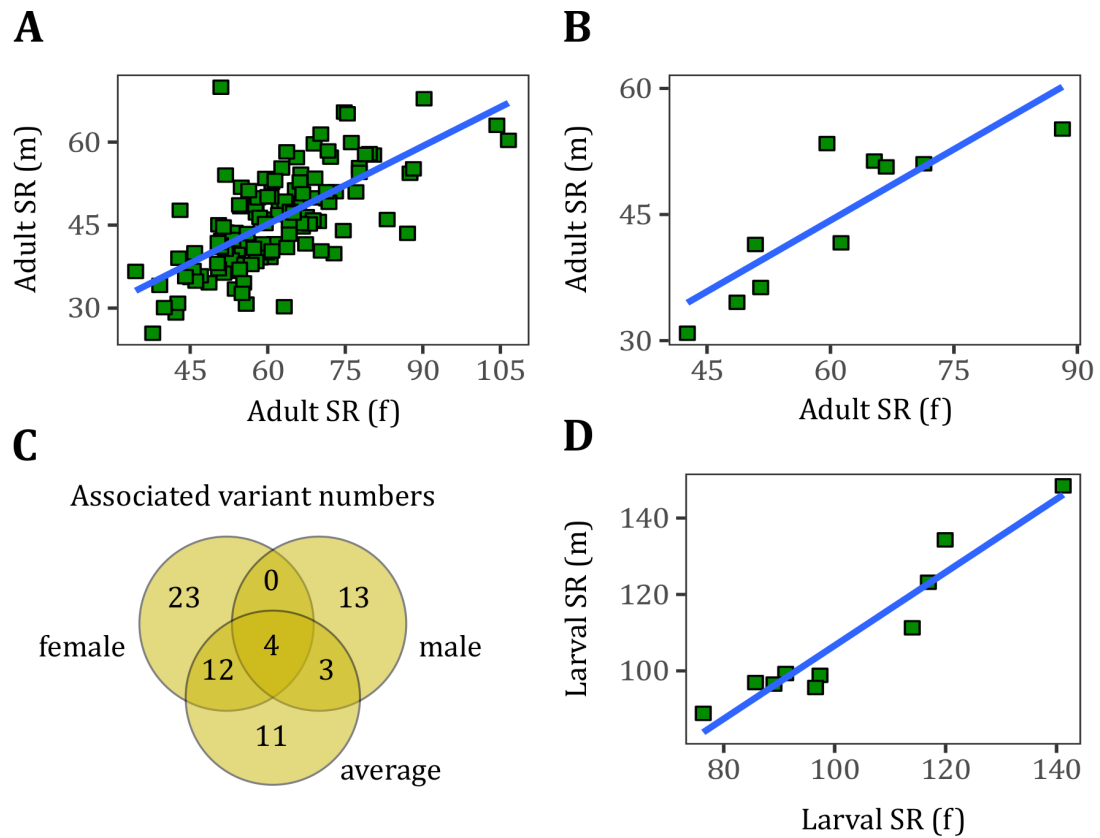


FIGURE 14. Cross-sex correlations in adult and larval starvation resistance. (A, B) Male (m) and female (f) starvation resistances (SR) are compared to each other at the adult stage using all of the 203 DGRP lines (A) and using the 10 DGRP lines that were sexed for larval SR (B). (C) Venn diagram shows the numbers of associated variants ($p < 1.0 \times 10^{-5}$) identified in three genome-wide association tests for adult SR using female line means, male line means and the their averages (tests were performed using the DGRP2 webtool, <http://dgrp2.gnets.ncsu.edu>). Some variants show sex-specific effects whereas others are common to both sexes. (D) The correlation between female and male larval SR across the 10 DGRP lines is also shown for comparison.

4.8.4. Q-Q plot for GWAS on inversion-corrected mean larval starvation resistances

To check for any systemic bias in our GWASs, we inspected the distribution of p values of our association tests in a Q-Q plot. Q-Q plots show the observed vs. expected distribution of the negative logarithm of p values, and early deviations from the diagonal line might imply a systemic bias, such as population structure (Bush & Moore, 2012). For the GWAS that used In3R_Mo inversion-corrected mean larval starvation resistance phenotypes of DGRP lines, we observed a relatively early, but not sharp, deviation from the diagonal line around $p \sim 1 \times 10^{-3}$.

(Figure 15). We think this deviation did not stem from genetic heterogeneity, or population structure. It was rather due to the non-normal distribution of the inversion-corrected phenotypes. As seen in Figure 6, the distribution is slightly right-skewed due to highly resistant lines. Considering the small size of the DGRP population (e.g., 129 lines in our case), early inflation of p values is not unexpected for non-normal phenotypic distributions, and has also been observed in other GWASs in this population. In contrast, the Q-Q plot for variance GWAS (i.e., $\ln(\sigma_E)$) was aligned along the diagonal line since its phenotypic distribution followed a normal distribution (data not shown).

We performed GWASs also using the normalized and winsorized phenotypes to compare the resulting associations. Winsorization is a data transformation method in which extreme observations are re-assigned with less extreme values, rather than being removed from the data. Both manipulations were done on inversion-corrected mean larval starvation resistance phenotypes. Winsorization was done in two different ways; firstly, the highest 7 lines that were above the 95th percentile were given the value of the 95th percentile (WIN1), and, secondly, the highest 3 lines were given the value of the fourth highest line (WIN2). We observed that winsorization resulted in weaker deviations in the Q-Q plot, whereas normalization completely eliminated the deviation. We compared the top associations in our standard GWAS (only corrected for inversions) and the GWASs for NORM, WIN1 and WIN2 phenotypes, and found large overlap between them, even though there were also associations specific to a single GWAS. Therefore, we decided to use the GWAS results derived from inversion-corrected phenotypes for further analysis, since other statistical manipulations on the phenotypic distribution were devoid of a biological underpinning. Lastly, we also tried a GWAS after trimming the most resistant 3 DGRP lines. We found that the associations were very different than our standard GWAS since eliminating 3 genotypes from our already-small study population considerably changed the allele frequencies, affecting especially the rare alleles (data not shown).

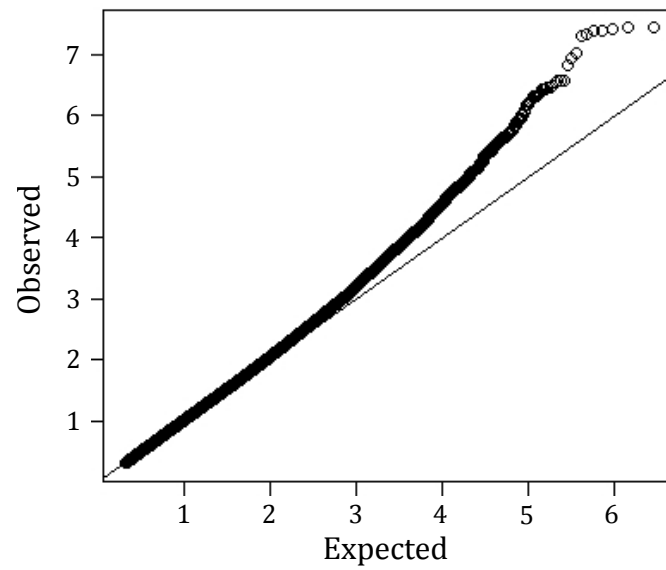


FIGURE 15. Q-Q plot for GWAS on mean larval starvation resistance. p -values were derived from GWAS that used In3R_Mo chromosomal inversion-corrected line means. The $-\log_{10}$ transformed observed p -values (y-axis) are plotted against the $-\log_{10}$ transformed expected p -values (x-axis) under the assumption of no association.

5. DISCUSSION

5.1. Phenotypic and genotypic variation in larval starvation resistance

We assessed starvation resistances of DGRP larvae by measuring their survival time under starvation in 20% sucrose in PBS. It is critical to eliminate or minimize environmental sources of variation in order to increase the statistical power of a GWAS. We standardized our culture and experimental conditions for this purpose, so that all DGRP lines were treated equally. To begin with, we performed our experiments in incubators in which temperature, humidity and lightning were tightly controlled. Likewise, we reared *Drosophila* lines for at least 3 generations before the experiment in incubators. We shuffled the vials inside the incubators every day to randomize location effects. In order to control for population density, we arranged the number of parents and their mating time individually for each line when starting new culture vials, since DGRP lines vary in their fecundity (Durham et al., 2014). We ensured all flies had maximal size for at least 2 generations before the starvation assay. We phenotyped DGRP lines using large numbers and multiple replicate samples. Almost half of the lines were tested more than once in multiple blocks, and each line was tested for 3 successive days in a single block. We used a new batch of starvation medium each day so that every DGRP line was tested in at least 3 different batches. We also distributed the larvae into two different 96-well plates in each day so that they would experience different locations inside the incubator. In conclusion, environmental variables were either minimized or randomized across genotypes and replicates.

The tight control for tangible environmental factors and having large sample sizes allowed us to obtain accurate estimates of DGRP line means. Furthermore, repeated measurements of lines in multiple blocks showed that our estimates were highly reproducible, despite the fact that larval starvation resistance is highly plastic. We used 129 line means in our GWAS, since we were not able to phenotype all of the 203 DGRP lines. The reason for this is that we needed many larvae of the same age for the larval starvation experiment, and not all DGRP lines are suitable to provide this. More specifically, we had to collect the eggs within a time window of 2 hours, and the number of well-fed, average-sized larvae 36 hours after egg-laying (AEL) had to be much larger than 96 since we selectively transferred them. However, some DGRP lines did not lay enough eggs even though there were many parents in many egg-laying cages. Some lines did not propagate in large numbers to produce enough parents, whereas others did not propagate in well-fed states as judged by their small sizes as adults. For these reasons, we could not increase our sample size above 129 lines.

In total, we phenotyped 57351 larvae for their starvation resistance, and, after omitting the deaths at the first viability counts, we ended up with 52762 larvae. We used a linear regression model to partition the total phenotypic variation in this dataset into its components, i.e., genetic variation, replicate factors and residual variation. Replicate factors (block, day and plate) had statistically significant effects, but their absolute effects were very small, altogether explaining only less than 1.7% of total variance. We did not correct for the effects of these replicate factors since each factor was comprised of many levels, and each level had a small sample size. For instance, we had 60 levels for the block factor, each level including 2-8 lines. This means that blocks were not represented by enough lines to provide any reliable estimate of their effects. As a specific example, one of the blocks contained three lines, two of which were, by chance, the most resistant lines. In this case, we re-tested these lines in other blocks to ensure that their extreme phenotypes were not due to a block effect. In addition, replicate effects were also not always in the same direction for different lines. For instance, we often observed that in a single block, one line had a lower mean value than what was observed in another block, whereas another line had a similar or higher mean value compared to its mean in another block. Such inconsistencies implied that most of the inter-replicate deviations were stochastic rather than systematic, hence, should not be corrected. However, we also eliminated some of the replicates if their deviation was likely to result from a biological reason. For instance, we sometimes observed that the number of eggs a particular DGRP line laid was considerably smaller in the first or last day of the experiment. In such instances, we noted these deviations, continued the starvation assay and used these replicates only if they were not significantly different from the others. In other instances, we observed that larvae of a particular line were significantly smaller at the start of starvation in one day than in the other days. In such cases, we eliminated either the replicates on those days or the particular line completely from the block. Therefore, we employed data trimming on the basis of biological reasoning and kept all the deviations if they did not result from a biological detectable reason.

As detailed above, the phenotyping screen was not always straightforward; we had to eliminate many lines at various steps during the experiment and repeat the phenotyping to have reliable quantification of our trait. We indeed tested almost half of the lines multiple times across blocks, especially those with extreme phenotypes, or those with broad within-line variations. Even though we initially considered using a common control line in each block, we could not continue with goal due to limitations in screening capacities and time considerations. However, the repeated measurements of multiple lines showed a high reproducibility, which reassured us that our phenotyping is rather robust.

We observed a very high genetic variation for larval starvation resistance, as evidenced by the broad range distribution of the DGRP line means (between 72 and 167 hours). This large genetic variation prompted us to perform a genome-wide association analysis to identify the underlying genetic variants. Despite the large genetic variation, we also estimated a relatively low heritability ($H^2 = 0.27$), since the contribution of between-line variance to the total phenotypic variance remained relatively small due to a larger within-line variance. It is known that traits closely related to fitness tend to have lower heritabilities than traits more distantly related to fitness (Merila & Sheldon, 1999). This is not because they have lower genetic variances. Rather, fitness traits tend to have higher levels of genetic variance, higher numbers of underlying loci and more complex etiologies, as a result of which they have higher environmental sensitivity and plasticity (Price & Schluter, 1991; Houle, 1992). Our observation of a relatively low heritability for larval starvation resistance might a good example for this phenomenon.

The majority of the total phenotypic variance ($V_E/V_P = 0.72$) is residual, or unexplained. This is the variance component that cannot be attributed to any tangible factor; therefore, it is thought to result from differential interactions of individuals with their micro-environments. The ability of isogenic individuals to assume varying phenotypic outcomes when reared or treated under the same macro-environmental conditions (e.g. temperature, batch, etc.) is called micro-environmental plasticity, and we will discuss it in more detail in Chapter 5.8. However, a large micro-environmental plasticity might not be the whole picture. Part of the residual variance might also be caused from other factors that we could not account for in our experimental setup, such as age and sex. Each of these will be discussed in more detail in the next chapters.

5.2. Starvation resistance during early larval stages

We examined starvation resistances at different larval ages during early development and observed that starvation resistance increased during the first instar but decreased after molting into the second instar. The change in the direction of the starvation resistance vs. age curve was very sharp around the time of the first molt. Furthermore, mid L2 larvae were even less resistant than early L2 larvae. Hence, the decline in starvation resistance was a general trend during the second instar rather than being confined to molting time itself. This suggests that after the first molt, there might be a shift in the physiological or metabolic state of larvae that lasts during the second instar. More detailed enzymatic, transcriptomic, and metabolic analyses are needed to elucidate the potentially distinct metabolic states between L1 and L2 instars.

Of particular interest to us is how fast larval starvation resistance changes with age around the time we start starvation. From 36 to 38 hours AEL, RAL_379 larvae survived roughly 16 hours longer under starvation. For another line, RAL_357, the difference from 36 to 38 hours AEL was 23 hours (data not shown). Hence, 2 hours of extra feeding provided larvae enough supplies to survive starvation an average of 20 hours more. Based on these results, we restricted the egg-laying time to 2 hours in order to prevent age variability among larvae and to collect sufficient amounts of eggs at the same time. Moreover, in the starvation assay, we selected the larvae that are of the size of the majority, and did not include very small or very large larvae (possibly of different age). Despite these, there must still be age differences among larvae, which contribute to the large residual variance we observed. In conclusion, age is an important confounding factor for larval traits.

5.3. Sexual dimorphism in larval starvation resistance

Sex is very often an important factor that contributes to the phenotypic variation in quantitative traits in sexually reproducing organisms (Flint & Mackay, 2009). Adult traits studied in the DGRP showed significant sexual dimorphism (i.e., males and females had different phenotypic means) and genotype-by-sex interaction (i.e., the magnitude of the difference between males and females was different for each genotype). Whereas males and females are morphologically different from each other at the adult stage, they are indistinguishable till the third larval instar. Due to this, we were not able to quantify their starvation resistance separately. However, we reasoned that phenotyping the larvae at an early larval stage of development would minimize the effects of possible physiological differences between the sexes.

We suspected sexual dimorphism for being partially responsible for the large within-line variations in our trait. More interestingly, a few DGRP lines had a bimodal distribution of survival times, raising our concern that each mode might have represented either sex. If this were the case, then there was a very large sexual dimorphism for those lines. These observations motivated us to examine whether sexual dimorphism existed in our trait. We re-phenotyped and determined the larval sexes for 10 DGRP lines, and found that in 4 lines male larvae survived significantly longer than female larvae. This is in contrast to the observations on the adult flies in these lines, where male adults were more sensitive than female adults (Mackay et al., 2012).

We observed a high genetic correlation between males and females ($r_{GMF} = 0.958$). Resistant lines were comprised of resistant males and females, and sensitive lines were comprised of sensitive males and females. Therefore, sex-by-genotype interactions were not extensive. A high cross-sex genetic correlation indicates that the majority of the variants are common to both sexes, and a minor fraction are sex-specific. Therefore, running a GWAS with sexes-pooled phenotypic data would be more reasonable when there is a high correlation across sexes; because associations with sex-specific effects might be overlooked when sexes are pooled. Even though our sexed data was comprised of only 10 lines, and, hence, cannot accurately predict the r_{GMF} in the larger dataset, it is evident that a high cross-sex genetic correlation exists for larval starvation resistance.

Another important finding from this investigation was that even though sexual dimorphism was shown to exist, it is not prevalent to the extent to explain the large intra-line variations in our trait. Indeed, males and females had very similar intra-line variances with sexes-pooled data. More importantly, the bimodal death pattern observed in some lines was preserved when sexes were plotted separately indicating that it did not result from sexual dimorphism as initially suspected. In accordance with the above observations, we did not observe an improved heritability estimate when sex was incorporated as a covariate in the regression model, or when sexes were analyzed separately.

In conclusion, even though the sex of larvae is an unknown factor in our trait, and sexual dimorphism potentially contributes to residual variation when sexes are pooled, the absolute size of its effect is not very extensive. Moreover, a high cross-sex genetic correlation suggests that using sexes-pooled phenotypes for GWAS will satisfactorily identify most of the associations.

5.4. Genetic correlations with other traits in the DGRP

Genetic reference populations offer the advantage of studying diverse quantitative traits on the same genotypes and investigating the genetic correlations among them. Here, we examined the correlations between larval starvation resistance and other traits that have been studied previously in the DGRP. Phenotypic correlations are fair estimates of genetic correlations between traits, especially when sample sizes are large (Cheverud, 1988; Roff, 1995). Therefore, we used Pearson's product moment correlation on line means to study cross-trait genetic correlations.

We found a significant positive correlation with triglyceride content of flies that were developed on high glucose diet for at least 2 generations before the measurement (Unckless et al., 2015). Even though triglyceride contents were measured at the adult stage, it is known that larval feeding also influences adult lipid levels (Chippindale et al., 1996), hence, the trait has both larval and adult stage underpinnings. Triglycerides are energy-dense storage molecules that are vital to survive prolonged periods of food limitation. Experimental evolution studies established a causal relationship between lipid levels and starvation resistance at the adult stage, i.e., selection for increased starvation resistance resulted in higher lipid levels (Rion & Kawecki, 2007). However, this relationship was missing in studies that compared these two traits in isofemale lines derived from a natural population (Hoffmann et al., 2001; Jumbo-Lucioni et al., 2010; Unckless et al., 2015). Here, we showed that there is evidence for a positive correlation between lipid levels and starvation resistance also in isofemale lines.

Strikingly, we did not observe a genetic correlation between larval and adult starvation resistances in the DGRP. This observation might make biological sense in the light of the facts that larvae and adults are anatomically very different, their organs originate from different precursors, they undertake different physiological functions, have different nutritional needs, and experience different niches in nature. Alternatively, the lack of correlation in starvation resistances might also be due to the different starvation media they were tested in. Adults were starved on non-nutritive agar medium (Mackay et al., 2012), whereas we starved our larvae in sucrose solution. Starving L1 larvae in plain PBS solution results in their immediate deaths within 1 day, hence does not allow detecting phenotypic variation between DGRP lines. However, adult starvation resistance can be quantified also on sucrose-agar medium, or in vials containing agar media and capillary tubes with sucrose solution inserted through foam plugs. It would be interesting to identify whether this trait correlates with larval starvation resistance and acute adult starvation resistance.

Starvation resistance showed a modest positive correlation with ethanol-induced knockdown time, i.e., sensitivity to alcohol exposure (Morozova et al., 2015). *Drosophila* encounter ethanol in their natural habitats (i.e., fermented food sources), and efficiently utilize it for fatty acid synthesis (Geer et al., 1985). Therefore, this positive relationship might derive from alleles that confer more efficient ethanol intoxication and increased lipid synthesis. Moreover, many genes that coordinate the cellular and organismal energy homeostasis pathways, such as malic enzyme, Neuropeptide F, etc., are also known to play key roles in regulating the metabolic and behavioral responses to ethanol (Wen et al., 2005; Morozova et al., 2014). Therefore, the relationship between these two traits might also involve other aspects of metabolic regulation.

Lastly, we also identified a modest negative correlation between larval starvation resistances and negative geotaxis activity of flies exposed to chronic oxidative stress (MSB) (Jordan et al., 2012). Flies selected for starvation resistance display reduced locomotor activities and increased sleep compared to control flies (Williams et al., 2004; Masek et al., 2014). These adaptations increase starvation survival times by reducing energy expenditure. Hence, a negative correlation between locomotor activities and starvation resistance is plausible. However, negative geotaxis activities of flies were assessed following chronic exposure to MSB, hence reflect their locomotor performance under chronic oxidative stress, i.e., their resistance to oxidative stress. Given that oxidative stress resistance has been mostly associated positively with starvation resistance (Harshman, et al., 1999b; Baldal et al., 2006), it could also be expected that a positive relationship would exist. Therefore, the biological meaning of this correlation is not very intuitive, and should be investigated more carefully in future studies.

5.5. GWAS identifies many loci and genes involved in larval starvation resistance

We have shown that there is substantial genetic variation for larval starvation resistance in the DGRP, as evidenced by more than 2 fold differences in line means (i.e., 72-167 hours). In order to understand which genetic variants underlie this variation, we performed a GWAS in which we tested 1,408,196 SNPs with a prevalence of at least 5% among the 129 lines we phenotyped. We identified 94 significantly associated SNPs ($p < 1.0e-05$) with varying sizes of effect on the phenotypic variation and varying allele frequencies in the DGRP population.

We found that the majority of the associations were located in intronic and intergenic regions. This finding highlights the importance of regulatory regions in shaping the phenotypic landscape of quantitative traits, and was also observed in other GWASs in the DGRP (Jordan et al., 2012; Mackay et al., 2012; Dembeck et al., 2015; Vonesch et al., 2016). Half of the intergenic SNPs were located within 1 kb of annotated genes. These SNPs most probably affect the organismal phenotype through their direct influences on the expression of nearby genes. Therefore, the genes nearby associated SNPs must be prioritized for follow-up studies. Identifying the function of these genes will provide insights into the molecular mechanisms underlying starvation resistances. We identified 62 genes that contain associated SNPs or lie within 1kb around them. 54 of these candidate genes are protein coding. The known or predicted functions of these

proteins are very diverse, reflecting the polygenic nature of larval starvation resistance.

The effect sizes of associations, i.e. the magnitude of the phenotypic difference between the two alleles, were negatively correlated with their (minor) allele frequencies. Associated rare alleles had larger effect sizes than associated common alleles, producing a J-shaped distribution in the effect size vs. frequency plot. This J-shaped distribution has also been observed for other traits in the DGRP population (Flint & Mackay, 2009; Jordan et al., 2012; Mackay et al., 2012; Weber et al., 2012). Mackay et al. and Jordan et al. argue that if human quantitative traits and complex diseases have a similar genetic architecture, the missing heritability problem in human GWASs can partially be explained by their inability to detect rare variant associations (Aitman et al., 2011; Jordan et al., 2012). Human GWASs could not identify very rare variants to date; better genotyping arrays and greater sample sizes are needed for this (Robinson et al., 2014). Additionally, the lack of common variants with a large effect can be explained by the mutation-selection model (Keller et al., 2012). This model predicts that alleles with large effects on starvation resistance might have adverse effects on other traits; therefore, they never reach to high frequencies. Indeed, starvation resistance is known to have a negative relationship with fecundity (Wayne et al., 2006; Lee & Jang, 2014). Hence, alleles that increase the resistance of individuals to starvation might decrease their reproductive success. Such alleles are eliminated from the population by selection. However, new alleles with large effects are also created by mutations, balancing their elimination from the population. As a result of the mutation-selection balance, alleles with larger effects are maintained at low frequencies in the population.

The 94 associated SNPs identified in the GWAS regulate the phenotypic variation in the DGRP. GWAS findings are specific to the population: if tested in other populations, some SNPs might not show a significant effect due to different allele frequencies. For example, an allele found associated in the DGRP might be present in a much lower frequency in other populations preventing detection in GWAS. Likewise, there must be many more SNPs that affect variation in other settings but not in the DGRP. DGRP lines represent, to a large degree, the genetic variation in the original wild population they were derived from. However, some rare alleles in the original population might have been lost during the inbreeding process. Rare alleles that are found in less than 7 lines ($MAF < 0.05$) were excluded from our GWAS to prevent false positive associations. Since we were not able to test sexes separately, it is possible that some alleles that show sex-specific effects were not detected. Furthermore, the small size of our DGRP population prevents detecting common alleles with very small effect sizes. Small effect sizes become statistically significant in larger populations. Large sample sizes increase the power of GWAS in two ways; first by allowing more rare SNPs

to reach the minor allele frequency threshold and second by decreasing the standard errors and hence, leading to more significant p values. Therefore, the 94 associations identified here represent only a fraction of the all variants that affect larval starvation resistance. We propose methods to identify additional variants in Chapters 6.2.3-6.2.5.

5.6. Epistatic interactions between loci

In addition to investigating the association of individual SNPs, we can also investigate the effects of loci in pairs. This can be done by fitting a model that includes the marginal effects of each SNP along with their interaction effect (Schüpbach et al., 2010). Therefore, we can identify the significant epistatic interactions that contribute to the larval starvation resistance trait. Epistasis describes the genotype-by-genotype interactions between two loci; it occurs when the effect of polymorphism at one locus is suppressed or enhanced by the polymorphism at another (interacting) locus (Mackay, 2014). Epistatic interactions are indeed an important feature of the genetic basis of quantitative traits (Huang et al., 2012).

Detection of epistasis is a more challenging task than the detection of single locus effects. Genome-wide testing of interactions between pairs of loci squares the number of tests to be performed; therefore, it is computationally highly intensive (Struchalin et al., 2010). For this reason, epistatic interactions are usually investigated using a smaller subset of SNPs. Certain regions or SNPs that are suspected to be involved in the trait can be chosen as focal SNPs and the pairwise interactions between these focal SNPs and all other common SNPs in the genome might be tested. Since SNPs involved in interactions are likely to show heterogeneity in genotypic variance of the trait, selecting the SNPs that affect the variances among the line means might be a reasonable choice for focal SNPs (Swarup et al., 2013). These are the SNPs for which the variances among line means, rather than the means of line means as in the mean GWAS, differ significantly between alleles. Such SNPs can be detected by genome-wide testing of heterogeneity using Levene's test for homogeneity of variances (Struchalin et al., 2010). The top hits from this genetic variance GWAS might then be used as focal SNPs in a two-locus GWAS to investigate the epistatic interactions contributing to the natural variation in larval starvation resistance. With this analysis, we may identify novel genes that were not detected in mean or environmental variance GWAS, extending the list of candidate genes possibly involved in the trait. Furthermore, using the STRING database of known gene interactions, we can search whether pairwise interactions between the SNPs are also observed between the genes associated with those SNPs.

The power to detect epistatic interactions is limited, even in an inbred genetic background. One reason for this is that most genotype-by-genotype interactions are weaker than additive, marginal effects of SNPs. Another reason is that the numbers of individuals that represent the homozygous double-minor allele population might be too small to be tested. Therefore, a small population size, as we have in this study, affects epistasis analysis more severely than one-locus GWAS. In addition to decreasing the significance of test statistics, a small sample size can also lead to inflated false positives (Mackay, 2014). In order to increase the power of detection and decrease the number of false positives, the inversion-corrected phenotypes must be normalized prior to epistasis GWAS. Furthermore, a higher minor allele frequency threshold (e.g., $MAF > 0.1$) must be chosen. The next step after a two-locus GWAS might be to determine which statistically significant interactions are biologically relevant. In this regard, genes tagged by each interacting SNP might be searched for a common biological process since this is a good indication for a biologically meaningful interaction. Another effective method to confirm the epistatic interactions between loci is based on the use of nearly isogenic lines (NILs) derived from the DGRP lines, which will be discussed in detail in Chapter 6.2.3.

5.7. Functional validation of associations

To demonstrate that our GWAS had enough power to identify true associations, we wanted to replicate the association of one of the top GWAS hits (*3R_12356551_SNP*) in outbred cohorts. We generated two advanced intercross populations each of which was homozygous at the target SNP for either allele, but outbred elsewhere in the genome. To maximize genetic divergence, we used all the DGRP lines (12) that possessed the minor allele at this SNP to generate the minor allele population. We also randomly selected 16 DGRP lines that possessed the major allele to generate the major allele population. Next, we crossed them in a round robin mating scheme in order to ensure that alleles from all parental lines were equally represented in the final populations. We maintained these populations with large effective sizes (400 males and 400 females) by randomly mating them for 5 generations to break down LD before phenotyping a large sample size from each population in the larval starvation assay. The two outbred populations showed a highly significant and large difference in their larval starvation resistance; hence, the association of the target SNP to the trait was replicated.

The ability to generate allele-specific outbred populations using the genotyped DGRP lines as founders eliminated the requisite to genotype all of the individuals

in the replication cohorts. Only to ensure that we produced the desired populations, i.e. without any cross-contamination from lines of the other genotype, we pool sequenced 10 flies from the last generation of each population. Despite this advantage, using DGRP-derived replication cohorts for SNP validation also has a drawback; because we were restricted with the number of lines that we could use for the minor allele. Among the 12 lines carrying the minor allele lines within the whole DGRP population, we had already phenotyped 7 of them for the GWAS. Since these 7 lines were in average more resistant than a randomly chosen DGRP line, this might have brought a bias to the replication phase. Ideally, lines that were not phenotyped must have been chosen as founders; but, we did not have a large enough number of unphenotyped lines for the minor allele of this target SNP, and hence we had to use all minor allele-carrying lines. Nevertheless, we reasoned that an outbred population would not have the same genetic make-up as an inbred line population; therefore, it is suitable for replication. For instance, in an outbred background, segregating alleles could produce new dominant/recessive and epistatic interactions that did not exist in the original DGRP population. Moreover, allele frequencies and local LD structures must have been different as well due to a small number of founder genotypes. These genetic differences between outbred and inbred backgrounds have the potential to change the phenotypic landscape radically. Indeed, this is proven by the heterosis of replication cohorts; both populations had mean values much higher than the DGRP average.

Lastly, it should be noted that even though the minor allele of our target SNP was associated with higher starvation resistances both in the GWAS and in the replication phase, the size of this effect was much smaller in the latter (33.7 vs. 13.1 hours). This discrepancy might be explained by the Beavis effect, which describes the bias of GWASs to overestimate the effect sizes of associated loci, especially when the sample sizes are small (Xu, 2003). The bias stems from the fact that associated loci actually represent a truncated distribution as they have to reach a critical significance threshold for detection, and this leads to an upward bias in predicting the effects of detected loci. Alternatively, the discrepancy might have resulted from the differences between the genetic make-ups of the DGRP and the replication population. For instance, there might be another polymorphism closely linked to our target SNP showing an opposite effect on the trait (i.e., decreasing starvation resistance). This allele might be more prevalent in the resulting minor allele population even though it only existed in 2 of the 12 minor allele DGRP lines. Or, epistatic interactions might be more prevalent in the outbreeding population, masking the additive effects of our target SNP (also known as genetic canalization). Such scenarios can explain the smaller effect size observed in the replication phase.

In addition to functional validation of associated SNPs, candidate genes identified in the GWAS must also be validated for a functional role in the trait using classical genetic methods. RNAi lines, which are available for many *Drosophila* genes, have been used in previous GWASs on this population. However, we were concerned that RNA interference might not be as convenient for larval starvation resistance as it is for adult stage traits. RNAi constructs might lead to lethality or reduced fitness when expressed ubiquitously. Additionally, tissue specific driver lines for RNAi expression might not be available for all tissues at all developmental times, or might not be very specific. More importantly, maternal effects might still be present at 36 hours AEL preventing to observe the effects of RNA knockdown in this early developmental stage. Therefore, we decided to screen transposable-element (TE)-induced mutants and assay their starvation resistance. We selected homozygous viable mutant lines from the MB (Metaxakis et al., 2005) and EY (Bellen et al., 2004) libraries in which mutant lines were generated by mobilizing a TE element and causing it to be inserted inside or very close to genes to disrupt its function. Since the mutations were homozygous, we were able to test them without prior crossing. Furthermore, since they were viable, we were not concerned with lethality issues, and any effect detected in the larval starvation assay must have arisen from the mutation of the gene function rather than the overall weakness of the line. However, the drawback of this strategy was that we could not use all the candidate genes since some of them did not have any TE-induced mutant alleles, or since they had essential functions and were not homozygous viable. Finally, as the control group, we selected 3-4 lines with TEs inserted in non-coding regions or gene deserts. Unfortunately, we observed that control lines were very variable in their starvation resistance, preventing any chance to detect the effect of mutants in candidate genes. Therefore, we terminated this screen after phenotyping around 10 lines (data not shown). The variability between control lines likely resulted from the different positional effects of TE insertions; TEs might be located in the regulatory regions of the genome, disrupting their function and ultimately affecting starvation resistance. Furthermore, the genetic background was not isogenized in the EY collection, and this might have also added to the phenotypic variation. In conclusion, we could not validate the functions of our candidate genes using this strategy. It became obvious that more tailored experiments using more controlled genetic backgrounds have to be conducted. We discuss these in Chapter 6.2.1.

5.8. Micro-environmental plasticity for larval starvation resistance

As mentioned earlier, DGRP lines showed very high within-line environmental variances for larval starvation resistance. The magnitude of this micro-

environmental plasticity varied significantly across genotypes, implying that there is a genetic control for the micro-environmental plasticity of starvation resistance. Therefore, identifying the genetic loci that contribute to this genetic variation might provide insights about the regulation of both the mean and the micro-environmental variance of larval starvation resistance.

Micro-environmental plasticity of several traits, namely, adult starvation resistance, chill coma recovery, food intake, startle response, and sleep, have been previously studied in the DGRP population (Harbison et al., 2013; Garlapow et al., 2015; Morgante et al., 2015). The DGRP population offers the advantage of reliably estimating the magnitude of micro-environmental plasticity, since the variance between individuals of an isogenic inbred line must be solely due to micro-environmental plasticity. However, this expectation is not valid for larval traits, because larval age, a variable that cannot be minimized completely, and sex, a variable that cannot be accounted for, also contribute to the inter-individual phenotypic variations, as we have shown earlier. Therefore, estimates of micro-environmental plasticity in larval traits are confounded with variations in these two factors. For instance, there might be variation among DGRP lines for acclimatization to new egg-laying plates to start egg-laying, which in turn could lead to variation in the variability of larval ages across DGRP lines. To specify, those lines that rapidly adapt to new plates will lay eggs in a broader time window within the 2 hours they are allowed to, and therefore, will show a higher variability in age. As we have shown, even minor age differences can have profound phenotypic effects for the developing and rapidly growing larvae. Therefore, variation in the variability of larval age across the DGRP lines might contribute to variation in residual variances. Likewise, variations in sex ratios accompanied with significant sex-by-genotype effects might be a potential source of variation for the phenotypic variation in residual variances across DGRP lines. Therefore, it should be noted that associated loci identified in our variance GWAS might also include some loci that indeed affect age variability (e.g., egg-laying behavior, developmental rates, etc.) or sex ratios in addition to the loci that affect micro-environmental sensitivity.

An additional confounder for the phenotypic variation in micro-environmental plasticity might be the variation in the residual heterozygosity of the DGRP lines (Morgante et al., 2015). We tested this by searching for a correlation between the micro-environmental plasticity and the percentage of segregating alleles in the DGRP lines, but could not detect any correlation. However, it is possible that, rather than the cumulative heterozygosity, heterozygosity in a smaller number of causal genetic loci might underlie the inter-individual variation within an inbred line. In our GWAS, we only used biallelic SNPs that are found in the homozygous state. Therefore, we searched for the alleles that contribute to micro-environmental plasticity or canalization in the homozygous state. Examples of

such variants have been identified in *heat shock protein 83* for startle response (Morgante et al., 2015). Heat shock proteins, which are chaperone proteins that assist proper folding of other proteins, were shown to buffer the genetic diversity for both morphological and behavioral traits (Rutherford & Lindquist, 1998; Hung et al., 2009). For instance, reduction in Hsp83 function was shown to result in more variable phenotypes, with the normally cryptic deleterious mutations giving rise to developmental abnormalities. Therefore, Hsp83 functions by buffering the destabilizing effects of genetic variants and environmental processes on proteins (Rutherford & Lindquist, 1998). Our GWAS for micro-environmental plasticity did not detect any heat shock proteins or chaperones. However, we identified other candidate genes that regulate protein homeostasis in other ways, such as *pix* (translation initiation), *ValRS-m* and *Aats-arg* (tRNA synthetases), and *CG4968* (deubiquitinase), and we speculate that these genes might be plausible candidates for micro-environmental plasticity.

Lastly, micro-environmental plasticity showed a high correlation with mean starvation resistances. Therefore, there might be common mechanisms that regulate both traits. However, we could not detect enrichment of a specific pathway when we analyzed the candidate genes from both GWASs. Despite this, we identified candidate genes that are known, or predicted, to be involved in starvation responses, such as *CG34447* and *CG3448* (triglyceride lipases) and *Atg1* and *Atg16* (autophagy proteins). The involvement of these genes in the micro-environmental plasticity can be investigated in future studies using the genetic tools available in *Drosophila*.

5.9. Insights into molecular and cellular mechanisms underlying larval starvation resistance

Organisms can tolerate starvation to the extent their energy demand is met by their energy supplies. Mechanisms that shift the balance between energy supply and demand are likely to affect starvation survival times. Therefore, identifying the genes that underlie starvation resistance might give important insights into the genetic networks that regulate energy homeostasis. In accordance with this, previous candidate gene-based studies evidenced the effects of canonical regulators of growth and metabolism on starvation resistance. Examples include *chico*, the fly homolog of the insulin receptor substrate IRS (Clancy et al., 2001), *Akh*, *adipokinetic hormone* (Lee & Park, 2004), *foxo* (Kramer et al., 2008), methuselah (Lin et al., 1998), *Myc* and *dTOR* (Tettweiler et al., 2005). In addition, a P-element mutagenesis screen identified many genes involved in oogenesis, metabolism, feeding behavior and cell fate determination, extending the list of genes affecting starvation resistance in adults (Harbison et al., 2004). Our GWAS

for larval starvation resistance complements these studies identifying many novel candidate genes.

In our GWAS, we have not identified associations of genes that function in the canonical growth and metabolism pathways such as insulin-like, *AKH* and *dTOR* signaling pathways, even after we increased the significance threshold and considered the top 400 SNP associations. This observation is consistent with previous GWASs, which also could not detect such genes in traits like food intake and size, despite the fact that they have a clear role in these traits (Garlapow et al., 2015; Vonesch et al., 2016). A possible explanation is the fact that these genes undertake pivotal roles in growth and metabolism, and are major hubs for complex genetic networks. Therefore, polymorphisms in these genes are under strong purifying selection, and are rapidly purged from the population. Hence, variants in these genes most likely do not reach high frequencies to be detected by GWAS. Nevertheless, we identified many other genes that have well-established roles in regulating different aspects of energy homeostasis.

Regulators of energy metabolism and food intake

Our candidate gene list includes several genes that control feeding behavior. One of them is *klumpfuss (klu)*, a zinc finger transcription factor (Melcher & Pankratz, 2005). *klu* is known to regulate food intake in larvae, and this function involves the regulation of a neural circuit comprised of neurons that express *hugin*, a neuropeptide homologous to mammalian neuromedin U. This neural circuit transmits signals from gustatory neurons to organs that control feeding behavior, such as pharyngeal muscles and the central neuroendocrine organ. *klu* mutants have significantly altered *hugin* expression, and also show feeding defects (Melcher & Pankratz, 2005).

Another protein that regulates feeding behavior is *allatostatin A receptor 2 (AR-2)*, a G protein-coupled receptor expressed in the gut and the brain (Bowser & Tobe, 2005). Allatostatins reduce food intake in *Drosophila* (Hergarden et al., 2012). In relation to this, mutants that lack *Allatostatin A (AstA)* accumulate higher lipid levels (Hentze et al., 2015). *AstA*-expressing neurons are also involved in the regulation of food preferences. Activation of *AstA* neurons increases the preference for protein-rich diets whereas loss of *AstA* increases sugar consumption (Hentze et al., 2015). Moreover, *AstA* and *AR-2* are regulated differentially by dietary carbohydrates and proteins. Lastly, the mammalian homolog of *AR-2*, galanin receptor, also functions in the control of feeding behavior and energy metabolism (Fang et al., 2012).

tiwaz, (*twz*) is another candidate gene with roles in feeding behavior. It is the *Drosophila* homolog of K⁺ channel tetramerization domain-containing 15 (KCTD15), which is associated with obesity in humans (Xi et al., 2013; Lv et al., 2015). In *Drosophila*, *twz* is found in octopaminergic neurons. Octopamine is the invertebrate homolog of norepinephrine, and mediates starvation-induced hyperactivity in adults (Yang et al., 2015). *Twz* is thought to regulate feeding behavior by controlling the expression of the genes necessary for production and secretion of octopamine (Williams et al., 2014a). Loss of *twz* activity in octopaminergic neurons results in increased size of individual meals (Williams et al., 2014b).

An interesting candidate gene we identified is *pointed* (*pnt*). *pnt* is a transcription factor protein previously shown to affect adult starvation resistance in a mutational analysis screen (Harbison et al., 2004). Additionally, polymorphisms in *pnt* were associated with food intake in the DGRP, and this connection was later confirmed by RNA interference (Garlapow et al., 2015).

Genes involved in resource allocation and development

In addition to the genes that control food intake, we also identified one gene that is known to affect nutrient accumulation: *Sirtuin7*. *Sirt7* is a NAD-dependent histone deacetylase that functions in the nucleus to regulate gene expression. It is a member of a highly conserved family of sirtuins that play important roles in the regulation of energy metabolism and lifespan (Houtkooper et al., 2012). The dependence of sirtuin activities on NAD couples their activities directly to the energy status of the cell. Strikingly, a systems genetics study in the DGRP found that *Sirt7* expression levels were highly correlated with triglyceride levels in flies (Jumbo-Lucioni et al., 2010). They confirmed this correlation using transposable element (TE)-insertional mutants; mutants for *Sirt7* had significantly reduced triglyceride contents (Jumbo-Lucioni et al., 2010). The association of *Sirt7* transcript abundances both to triglyceride contents and to larval starvation resistance is meaningful, and might underlie the genetic correlation between these two organismal traits.

Resource allocation is of critical importance for starvation resistance (McCue, 2010). In our gene ontology (GO) enrichment analysis, we identified significant enrichments for GO terms such as organ development, neurogenesis, system development, etc. A previous P-element mutagenesis screen addressed to find genes affecting adult starvation resistance also found a similar result; they identified several genes with known roles in cell fate determination and cell proliferation (Harbison et al., 2004). These results suggest that apart from the mechanisms that underlie the metabolic and physiological responses to

starvation, the developmental processes that form the organs involved in starvation responses are important for the manifestation of starvation resistance. Therefore, resource allocation during development and/or before starvation has significant effects on this trait (Harbison et al., 2004).

Genes involved in taxis behavior

Interestingly, another GO term enrichment for larval starvation resistance was taxis. We identified several genes that are involved in locomotory behavior. As mentioned earlier, the nutritional status of organisms have large impacts on their physical activities and also in sleep behavior (Knoppien et al., 2000; Lee & Park, 2004; McDonald & Keene, 2010). Moreover, starvation-selected flies reduce their physical activities and increase their sleep as a mechanism to conserve energy (Williams et al., 2004; Masek et al., 2014). We think that enrichment of taxis genes reflects the importance of energy conservation mechanisms. Below we discuss the genes that have roles in locomotory behavior.

Cpo, *couch potato*, is an mRNA binding protein highly expressed in the central nervous system of *Drosophila* (Bellen et al., 1992b). Loss-of-function mutations of *cpo* show reduced flight ability, phototaxis, geotaxis, and ether recovery (Bellen et al., 1992a). An additional function of *cpo* is the regulation of the reproductive diapause phenotype, which is a state of reproductive dormancy of adult flies that occurs as a climatic adaptation to increase the lifespan and resist environmental challenges. Reproductive diapause is observed more frequently in natural populations that overwinter in more temperate climates. An amino acid polymorphism (Lys/Ile substitution) was found to be responsible for variations in the diapause phenotype in natural populations (Schmidt et al., 2008). The prevalence of this polymorphism in natural populations correlated with the latitudinal cline of the diapause phenotype. Given these pleiotropic roles of *cpo*, its association to larval starvation resistance implies a more general role for *cpo* in regulating dormancy processes as an adaptation to minimize energy expenditure.

Stambha A (*stmA*) is another gene that regulates locomotor activities. It encodes a protein required for synaptic vesicle exocytosis. Its suppression results in flight defective, hypoactive flies. Hence, its function is required for movement (Kumar et al., 2001; Huang et al., 2006). Its human homolog *EFR3* is associated with blood lipids, body mass index and type 1 diabetes (Talmud et al., 2009; Bradfield et al., 2011; Graff et al., 2013). However, it is noteworthy that the associated LD block also contains other genes including *POMC*, which is a well-established obesity-associated gene.

alan shepherd (*shep*) is a RNA-binding protein that regulates alternative splicing. It regulates neuronal remodeling during metamorphosis, adult locomotory behavior and sleep (Chen et al., 2014; Thimman et al., 2015). Its human homolog, *RBMS1*, is associated with type 2 diabetes (Qi et al., 2010). Lastly, *wake* (*wide awake*) is a protein that regulates the sleep onset based on the circadian rhythm. *wake* mutants are sleep defective (Liu et al., 2014). These genes are plausible candidates to study the genetic correlation between starvation resistance and sleep traits.

Human homologs involved in metabolic traits

Most of the candidate genes implicated in our GWAS have predicted homologs in humans. Given the high conservation of fundamental biological processes between *Drosophila* and humans, characterizing the function of these genes in either species will contribute to the efforts in elucidating the function in the other. Hence, we searched for the functions of human homologs of our candidate genes. Of particular interest are those that play important roles in complex metabolic diseases and clinically important traits. In addition to the aforementioned genes *AR-2*, *twz*, *Sirt7*, *stmA*, *shep*, we found other candidate genes with predicted human homologs associated to such traits. One of these genes is *CG30389*, which has no known function in *Drosophila*. Its human homolog *TMEM57* (*transmembrane protein 57*) was shown to be associated with blood lipid levels and coronary heart disease (Aulchenko et al., 2009; Teslovich et al., 2010). Its expression pattern is conserved across species; high expression in the central nervous system and localization to rough ERs (Arellano-Carbajal et al., 2011). Mutants of this gene in *C. elegans* display certain behavioral abnormalities, which can be rescued by human *TMEM57*, suggesting also a high functional conservation (Miyara et al., 2011).

Mub, *mushroom body expressed*, is a poly(C) RNA binding protein that regulates alternative splicing. A P-element mutagenesis screen identified that *mub* mutants have altered starvation resistances (Magwire et al., 2010). Its human homolog *PCBP3* is associated with insulin resistance (Irvin et al., 2011).

CG4168 has a human homolog named *IGFALS*, *Insulin-like growth factor binding protein acid labile subunit*. *IGFALS* is a serum protein that binds and regulates insulin-like growth factor action (Domené et al., 2005). *Drosophila* has another ALS ortholog named *dALS* (or, *convoluted*), expression of which has been shown to be sensitive to starvation conditions (Colombani et al., 2003). It antagonizes insulin function to control growth and regulate carbohydrate and fat metabolism. *dALS* overexpression in the fat body mimics the effects of starvation, such as mobilization and release of lipids from fat body (Arquier et al., 2008).

Association of *CG4168* to starvation resistance suggests that this uncharacterized gene might also undertake a similar role as *dALS* in the regulation of metabolism.

Adam-TS gene has a human homolog named *ADAMTS9*. It is a metalloproteinase, and is associated with type2 diabetes, insulin resistance, abdominal adipose depots and waist-hip ratio (Zeggini et al., 2008; Fox et al., 2012).

Lastly, *CG4434* is a glutamate dehydrogenase. Its human homolog, *GLUD1*, functions in regulating amino acid-induced insulin secretion. Polymorphisms in *GLUD1* were associated with hyperinsulinism (Tran et al., 2015).

The genes mentioned above are associated to traits related to metabolic regulation both in *Drosophila* and in humans. This reinforces their putative involvement in the relevant traits. Therefore, these genes must be prioritized for follow-up studies to characterize their function in the control of energy metabolism.

Novel insights into larval starvation resistance

We identified several genes that have no annotated molecular function. This might stimulate future candidate-based studies to characterize their functions. Besides this, we also identified many genes with known functions that have no clear connection to starvation resistance and have not been considered to have a role in starvation resistance. One such gene is *CG42342*, the association of which was replicated using the outbred populations. *CG42342* encodes a transmembrane collagen protein. Collagens are the most abundant extracellular matrix proteins that keep cells together in metazoans (Kadler et al., 2007). They comprise a large family of proteins all containing collagen triple helix motifs, assembling into different macromolecular structures (fibrils or networks, etc.), and showing tissue-specific distributions. The biological relevance of transmembrane collagens is not yet solved. We here provided a surprising evidence for the potential involvement of this protein in starvation resistance. Collagens are capable of initiating signal transduction via transmembrane collagen receptors such as integrins and discoidin domain receptors regulating a wide range of behaviors (Leitinger, 2011). Therefore, it is likely that collagens have pleiotropic roles that are beyond their roles as scaffold proteins. In conjunction with this, collagen remodeling was shown to underlie lifespan prolongation of *C.elegans* with reduced insulin signaling (Ewald et al., 2015). We demonstrated that the expression of *CG42342* is induced upon starvation in sucrose solution. This induction might play an important signal function in starvation responses. Alternatively, it might be induced upon amino acid starvation to allocate the resources into extracellular matrix to reduce the

metabolic rates in the cytoplasm and/or to deposit the resources outside the cells for later use. These possibilities need to be investigated in future studies.

Despite the fact that adult and larval starvation resistance did not correlate, we identified one gene that was associated with both traits: *CG6123*. 8 SNPs between *X_18223599* and *X_18224543* were associated with larval starvation resistance; 7 of these SNPs were in the intronic region and 1 was in the 5' UTR of *CG6123*. 7 other SNPs between *X_18235206* and *X_18245828*, all of which were located in the intron of *CG6123*, were associated with adult starvation resistance, as identified in a GWAS using the data from the DGRP2 website (Mackay et al., 2012). This gene has no annotated function, but its association to starvation resistance trait in two different studies places it at the top of the list for a functional validation.

Many other genes are plausible candidates for playing a role in starvation resistance. To name a few, *cropped* (*crp*) is the *Drosophila* AP-4 transcription factor. It acts downstream of *dMyc*, like its mammalian homolog *AP-4*. Its ectopic expression leads to increased cell and organ sizes (Wong et al., 2015). *dawdle* (*daw*) is an activin ligand that controls the rate of autophagy and insulin secretion (Bai et al., 2013). *klignon* (*klg*), is a transmembrane protein that function in the behavioral response to ethanol; P-element mutations in this gene result in increased ethanol sensitivity (Berger et al., 2008). Hence, it is a plausible target that might underlie the genetic correlation between larval starvation resistance and alcohol sensitivity. A full list for the associated genes is given in Supplementary Tables. In conclusion, the diversity of the molecular functions of the candidate genes demonstrates the multifactorial nature of the starvation resistance trait.

6. CONCLUSION AND OUTLOOK

6.1. Conclusion

Here, we studied starvation resistance in *Drosophila melanogaster* at an early larval stage by undertaking a GWAS approach and using a population of inbred *Drosophila* lines (DGRP).

First, we established the experimental setup for quantifying larval starvation resistances reliably. Minimizing the potential confounding effects of environmental factors is a prerequisite for association studies. We standardized culture conditions and experimental setup to phenotype the DGRP lines with highly reproducible estimates of line means.

The DGRP population harbors ample amount of genotypic diversity and shows phenotypic variation in any trait tested so far. In accordance with these, we also found substantial genetic variation in larval starvation. We exploited this natural variation to perform a GWAS, and identified the loci that affect starvation resistance. The majority of the associated loci were in the intergenic and intronic regions of the genome, highlighting the importance of regulatory elements in creating the phenotypic variability in a population. Associated loci marked many genes that have diverse molecular functions and are involved in diverse biological processes, reflecting the complex etiology underlying starvation resistance.

Starvation resistance is determined by the balance between energy supplies and energy expenditure. Therefore, genes that regulate energy homeostasis can also play a role in starvation resistance. In accordance with this, we identified many candidate genes that are known to regulate different aspects of energy homeostasis, such as regulation of food intake, nutrient accumulation and energy consumption. Many candidate genes have known functions that can be plausible candidates for starvation responses. Moreover, we also identified genes with no known function and genes with functions that have no clear connection to starvation resistance, highlighting the complementarity of GWAS approach to the classical genetics studies. Importantly, some of the associated genes have human homologs that are also associated with metabolic traits, such as obesity, insulin resistance, blood lipid levels, etc., reflecting the similarities between energy homeostasis pathways across species.

DGRP lines were previously quantified for many other traits, allowing us to examine the genetic correlations between these traits and the larval starvation

resistance. We identified positive correlations with triglyceride levels and resistance to alcohol exposure, and a negative correlation with locomotor performance under chronic mild oxidative stress. Some of our candidate genes are plausible candidates to contribute to the genetic correlations between these traits. Importantly, we did not detect any correlation between larval and adult starvation resistances, indicating that different genetic networks underlie the natural variation in these traits.

We observed that larval starvation resistance has a high residual variance, because of which we estimated a relatively low heritability. We investigated the potential factors that could contribute to this, and found that both larval age and sex contributed to phenotypic variation. Moreover, we also studied the micro-environmental plasticity of larval starvation resistance; a GWAS on this trait provided plausible candidate genes that need to be investigated further.

In conclusion, our study has been the first attempt to dissect the genetic basis of starvation resistance at the larval stage. Given the evolutionary conservation of major physiological and metabolic pathways, the results of this project will likely have implications beyond *Drosophila*.

6.2. Outlook

6.2.1. Functional characterization of associated genes using classical genetics

Genome-wide association studies provide the basis for future genetic and functional studies. The next step following a GWAS is to validate the causality of associations by characterizing the functions of associated variants and genes, and their biological role in the trait. We identified many novel and known genes that play a role in starvation resistance. Some of these genes have homologs in humans that are associated with important clinical traits, and it is mostly not clear how they are functionally linked to the trait. *Drosophila* provides many useful genetic tools to study the function of these genes (St Johnston, 2002). Mutant, misexpression and RNAi lines can be used to reveal the function of many genes. Some of these lines contain constructs that allow visualization of spatiotemporal expression patterns of genes. However, it should also be noted that classical genetic methods have their own weaknesses. RNAi constructs have drawbacks such as lethality, lack of tissue specificity, and ineffective knockdown due to maternal effects, etc. As mentioned before, we used transposable element-induced gene disruption lines to validate our candidate genes by comparing the

candidate gene mutants to a set of control lines in which TEs are inserted in non-coding regions of the genome. This strategy suffered from unwanted positional effects of TEs and genetic diversity found in outbred lines. Therefore, more tailored, i.e., gene-based, approaches should be taken. For instance, the control line could be generated for each gene itself by removing the TE from the coding region and restoring the gene function in exactly the same genetic background. However, TE-removal should be precise and must not result in loss of neighboring sequences. An emerging alternative method for generating specific genetic manipulations is CRISPR/Cas9 technology. Genetic tools for this system are now also available in *Drosophila*, allowing to more flexible genetic manipulations to study gene functions on the same genetic background (Bassett et al., 2013). Thus, the method used for functional characterization has to be carefully evaluated and will need optimization before application.

6.2.2. Systems genetics

Systems genetics is very instrumental to functionally validate the associated variants to the trait. Model organism genetic reference populations are suitable to employ this approach with high power (Ayroles et al., 2009; Civelek & Lusis, 2014). The DGRP lines can be profiled for their transcriptome, proteome and metabolome. Variations in these intermediate phenotypes (i.e. endophenotypes) are then correlated with variations in the genome and variations in the larval starvation resistance. Therefore associations are identified also between genotypes and endophenotypes, and between endophenotypes and organismal phenotypes filling in the gap between genotypes and organismal phenotypes. This establishes a chain of causation and explains how associated genetic variants function in the larval starvation resistance trait. Furthermore, this approach identifies networks of genes that are co-regulated and associations of these networks to the trait.

For endophenotype profiling, the more sensitive and more resistant lines (lines on the tails of the distribution, rather than average lines) can be selected to reduce costs and workload. At least three samples should be used: non-starved control at the time starvation is started, treatment group after starvation for a certain time, and non-starved developmental control at the same time with the treatment group. However, the inability to identify the sex of early stage larvae will not allow to test them separately, thus, sex-specific effects might be diluted. Any “omic” profiling might be selected based on the infrastructure and expertise of the lab. The transcriptomic data for adult flies reared under standard laboratory conditions is available on the website of DGRP (<http://dgrp2.gnets.ncsu.edu>); this dataset might be useful for some traits.

Usually, transcript abundances vary more closely with the genetic variations, whereas protein abundance vary more closely with phenotypic variation, reflecting the increased complexity from genes to phenotypes. A few studies that employed systems genetics approach using the DGRP population showed promise for the technique, as mentioned before (Ayroles et al., 2009; Jumbo-Lucioni et al., 2010).

6.2.3. Linkage mapping analysis

GWASs are underpowered to detect rare variants. However, such variants can be investigated with the complementary linkage mapping approach. We observed a broad distribution of DGRP line means for larval starvation resistance, ranging from 72 hours to 167 hours. The lines at the tails of this distribution are very interesting; their extreme phenotypes might be underpinned by rare alleles that have large effect sizes. Therefore, these lines are good candidates for linkage mapping analysis. Linkage mapping populations can be established by crossing two or more of these phenotypically divergent lines to generate F_2 recombinants that are then inbred to isogenicity to generate a set of recombinant inbred lines (RILs). The genomes of RILs will be random mosaics of the parental genomes due to meiotic recombination events. With this crossing scheme, alleles that are not common in the DGRP will be at intermediate frequencies in the final mapping population. RILs can be phenotyped for larval starvation resistance and then genotyped for certain markers across the genome. Phenotypes that vary across the genotypes are then mapped to the loci tagged with markers (Pollard, 2012). Mapping in RILs will have a lower resolution than in GWAS due to only few recombination events. However, mapping resolution can be increased by implementing advanced intercrossing after the F_2 generation to accumulate more recombinations.

An alternative QTL mapping strategy to map the variants found in extreme DGRP lines involves introgression, i.e., incorporation of a small genomic region from one line into the genome of an isogenic reference line by repeated backcrossing (Mackay, 2014). Chromosomes of an isogenic line can be individually substituted with those of an extreme DGRP line using standard techniques that involve balancer chromosomes (i.e., genetically modified chromosomes that prevent meiotic recombinations). These chromosome substitution lines can then be repeatedly backcrossed to the same isogenic line to produce nearly isogenic lines (NILs), in which much smaller genomic segments of DGRP origin have been introduced into the isogenic background (Weber et al., 2001). NILs are most useful when there is a candidate locus to zoom in and fine-map. Therefore, it might follow a RIL-based linkage mapping. Since there is no confounding

background genetic variation between the co-isogenic NILs, the statistical power of detection is much higher than RILs. However, NILs do not allow studying all genome at once. Finally, comparative analyses between different (co-isogenic) NILs might reveal the epistatic interactions among DGRP variants.

6.2.4. Experimental evolution

Another alternative method to find additional associated variants and loci is based on experimental evolution, i.e., artificial selection. Experimental evolution is a powerful tool to study adaptive physiological mechanisms underlying quantitative traits. Previous studies on adult starvation resistance using this approach provided important insights into the mechanisms underlying adult starvation resistance. Interestingly, some of these mechanisms were specific to the selection experiments and were not observed in studies using other approaches. For instance, genetic correlation between lipid levels and starvation resistances was well-established in artificial selection experiments, whereas it was missing in the population comparison studies that used isofemale lines derived from the same population (Hoffmann et al., 2001; Rion & Kawecki, 2007). Furthermore, the transcriptomic changes observed upon selection for starvation resistance differed from the transcriptomic changes observed upon acute starvation (Sorensen et al., 2007). These results demonstrated that acute responses to starvation and long-term adaptations to starvation might have different genetic bases. Consequently, association mapping in artificially selected populations might potentially identify a different suite of associations than GWAS.

To study larval starvation resistance with this approach, an outbreeding base population that will be subjected to artificial selection is generated in the first place. Advanced intercrossing of 40-100 DGRP lines for many generations produces a population with a large genetic diversity. From this base population, larvae are selected by starvation at each generation at a 50% selection pressure. Once 50% of the larvae die, the survivors are transferred to normal culture food to let them develop into adults and produce the next generation. Selection is applied for many generations until a direct phenotypic response becomes evident, i.e. when selected populations have significantly higher resistances. Selected and unselected control populations are then pool sequenced to identify the alleles that segregate at different frequencies in these populations, as in a case-control setup. Selected and control populations can also be tested for whole-genome transcriptome and proteome differences. It should be noted that very large population sizes are needed in artificial selection experiments to ensure that allele frequency differences are not due to genetic drift. Up to a

thousand flies are needed at each generation. Thus, the number of larvae used in the starvation assay must be far more than this, since half of them will be lost during selection and many others will be lost during their development into adults. Therefore, the starvation experiments must be performed in many 6-well plates with large numbers of larvae, and only a few representative 96-well plates must be used to detect the point of 50% survival.

As an alternative to starvation in 20% sucrose PBS, dietary restriction might be used as a source of selection. In fact, a previous study undertook this approach; they selected larvae on poor food to examine the correlated responses in life history traits and adult starvation resistance (Kolss et al., 2009). Larvae can still grow on poor food, although at a slower rate. Hence, to apply a selection pressure, larvae that developed within a limited time were selected. In this regard, their selection regime had also a developmental rate component besides the starvation resistance component. Even though this is a more natural setting than selection under acute starvation in sucrose solution, it is not as pure as the later.

6.2.5. Association mapping in other populations

Finally, in order to identify additional associated variants and loci for larval starvation resistance, other genetic reference populations might be used to perform a GWAS. In addition to identifying more variants specific to these new populations, these new GWASs also bear the potential to validate the associations identified in this study (if they are replicated). Besides the DGRP, there are two other reference populations with different origins: *Drosophila* Population Genomics Project (DPGP) (Pool et al., 2012) and *Drosophila* Synthetic Population Resource (DSPR) (King et al., 2012). However, since assessing larval starvation resistances manually is a highly time-consuming task, high-throughput methods that involve camera recordings of movements of larvae to assess viability should be considered to undertake these large-scale projects. An example of such a system is the FlyCatwalk that measures *Drosophila* adult size traits as the flies are photographed while they pass through a narrow tube (Medici et al., 2015).

7. MATERIALS AND METHODS

***Drosophila* culture**

All *Drosophila* lines were reared on cornmeal fly food (100 g fresh yeast, 55 g cornmeal, 10 g wheat flour, 75 g sugar, 8 g bacto-agar and 1 l tap water). They were maintained at controlled density, 25°C, 60% humidity and 14:10-hour light:dark cycle (08:30-22:30) in an incubator for at least 3 generations before phenotyping assays. All vials in the incubator were shuffled randomly every day.

Larval starvation resistance assay

1-3 days old flies were placed in egg-laying cages; after acclimatization for 2 days, they were allowed to lay eggs for 2 hours on agar plates with a yeast drop on top. After completion of egg-laying, agar plates were added more yeast paste. Exactly 36 hours after egg laying, L1 larvae were collected from plates by immersing them on tap water and filtering through a sieve. After washing them briefly under tap water (25°C) to get rid of any yeast, they were spread on Kim wipes soaked with the starvation medium (20% sucrose in phosphate buffered saline (PBS)). They were collected individually with forceps under a stereomicroscope and transferred into the wells of a 96-well plate that contain 80 µl starvation medium (25°C). Larvae that had sizes smaller or larger than the majority were not included in the assay. 96-well plates were tightly covered with an adhesive film to prevent escapes.

We recorded the survival of larvae twice a day from 1 day after transferring into starvation medium till all larvae died. The viability of the larvae was assessed by their ability to move. They were first only visually inspected, then tapped gently with forceps to move. If no movement was detected, they were briefly exposed to banana odor by placing a droplet of isoamyl acetate diluted in paraffin oil (1M) on the center of the wells of a 96-well lid. If this did not induce any movement, we first gently poked and then squeezed them with the tips of the forceps to ensure their death. At each count, the number of deaths and the time of counting were recorded. The deaths in the first count were omitted from calculation of the mean starvation survival time since they were most likely the result of getting harmed during transfer rather than starvation.

Phenotyping larval starvation resistances in the DGRP

In each experiment block, we had 4 defined time slots at intervals of 15 min devoted to different lines. For the first one, egg-laying started at 19:30, ended at 21:30; larvae from this slot were transferred into starvation medium at 08:30. Other slots followed 15 min after the previous one. We completed each transfer within 15 min and proceeded with the other. In each block, all lines were tested in 3 successive days and on each day larvae were distributed into 2 halves of 2 different 96-well plates. Each day, a new batch of starvation medium was used. Consequently, there were 6 replicate samples (2 plates x 3 days) per line per block. We calculated the mean larval starvation resistance of a single line by averaging the means of these replicates. We tested many lines in more than one block.

Quantitative genetics analysis

We partitioned the phenotypic variation into its components using the linear regression model $Y = \mu + L + B + D + P + E$, where Y is the phenotype, L , B , D and P are the random effects of line (genotype), block, day, and plates, respectively, and E is the error term, or residual variation. We fitted this model using the `lmer()` function of the `lme4` package in R. We assessed the significance of the random effects using the `rand()` function of the `lmerTest` package. Broad sense heritabilities were estimated as $H^2 = \sigma^2_L / (\sigma^2_L + \sigma^2_B + \sigma^2_D + \sigma^2_P + \sigma^2_E)$. In addition, we used the line means to analyze the effects of *Wolbachia* infection status and 5 chromosomal inversions using the `lm()` function. Before proceeding with GWAS, we corrected the line means by subtracting the effects from the line means.

For the small dataset that was used to analyze the effect of sex, we partitioned the phenotypic variation using the mixed model $Y = \mu + L + S + L \times S + R + E$, where L is the random effect of lines, S is the fixed effect of sex, $L \times S$ is the random effect of line-by-sex interaction, and R is the random effect of replicate samples. We estimated the cross-sex genetic correlation as $r_{GMF} = \text{cov}_{MF} / \sigma_M \sigma_F$ where cov_{MF} is the covariance of line means for between males and females, σ_M and σ_F are the standard deviations within males and females, respectively.

Genome-wide association analysis

We performed GWAS using means values of 129 DGRP lines. Genotypes were obtained from the DGRP Freeze 2 website (<http://dgrp2.gnets.ncsu.edu>). We filtered out the SNPs missing in a maximum of 10 lines. In total, 1,408,196 SNPs were tested using FaST-LMM, a linear regression modeling that accounts for

cryptic relatedness (Lippert et al., 2011). We performed gene ontology (GO) enrichment on the list of candidate genes associated with the nominally associated variants ($p < 1.e-05$) using the DAVID (Huang et al., 2009) and the STRING databases (Szklarczyk et al., 2015). Lastly, we predicted the human homologs of these candidate genes using DIOPT (Hu et al., 2011).

Correlations of larval starvation resistance with other traits

We computed Pearson's product-moment correlation coefficient (Pearson's r) for correlations between lines means of larval starvation resistance and other quantitative traits measured in the DGRP using the `cor()` and `cor.test()` functions in R.

Detection of the sex of dead larvae

In order to identify the sex of dead larvae, we sliced them finely with sterile needles in 2 μ l of Squishing Buffer (10 mM Tris pH 8.2, 1 mM EDTA, 25 mM NaCl, 200 μ g/ml proteinase K). After incubation at 37°C for 30 min and heat inactivation of proteinase K at 95°C for 5 min, we made a multiplex PCR in which we simultaneously amplified an autosomal region and a region on the Y chromosome with different sets of primers (Table S18) using High-Fidelity Phusion Polymerase (NEB M0530L). We ran the PCR products on agarose gel and checked the bands on an UV-illuminator to identify the sex of the larvae.

SNP-based validation

To validate the association of 3R_12356551_SNP, we established two outbred populations using the DGRP lines as founders. The 12 DGRP lines homozygous for the minor allele (T) and randomly selected 16 DGRP lines homozygous for the major allele (C) were used as founder lines to establish each population. These DGRP lines were crossed in a round-robin mating scheme in the first generation; line1 females were crossed with line2 males, line2 females were crossed with line3 males, and so on till the loop is closed. Subsequent populations were maintained in large numbers (400 females and 400 males in a total of 10 vials) as outbred populations with advanced intercrossing, i.e. mating between siblings is avoided. We phenotyped these populations for their larval starvation resistance and used non-parametric Mann-Whitney-Wilcoxon test to determine whether there was a significant difference. `Wilcox.test()` Wilcoxon rank sum test

Quantitative RT-PCR

Larvae were quick-frozen in liquid nitrogen at indicated time points during starvation in sucrose solution. 30-50 mg larvae were homogenized in 1 ml Quiazol by sonication (5 times 20 seconds, 100% amplitude, 80% cycling time) (Hielscher® up200ST-G). RNA was extracted using the Qiagen RNeasy Plus mini kit. cDNA was synthesized from 2 µg total RNA using oligo-dT primers and Superscript III reverse transcriptase at 50°C for 45 min. qPCR was performed in Roche light cycler 480 SYBR green I master on 0.5 µg cDNA. The CT values were normalized against *Act5C*. All primers are listed in Table S18.

Micro-environmental plasticity of larval starvation resistance

We considered micro-environmental plasticity as a trait of its own, hence quantified in the same manner we quantified mean starvation resistances. First, the magnitude of phenotypic variation within each replicate sample, and then the average of replicate samples were calculated to estimate the line means for the micro-environmental plasticity. We used different metrics: standard deviation (σ), coefficient of variation (CV), median absolute deviation (MAD), and their natural logarithms for quantifying the within-replicate phenotypic variations of isogenic individuals. CV was calculated as $CV_E = \sigma_E / \bar{y}$ where $\bar{y}$ is the mean of the replicate.

To estimate the heritability of the trait, we used a regression model of the form $Y = \mu + L + E$, where Y is the $\ln(\sigma_E)$ measure of micro-environmental plasticity, μ is the overall DGRP mean for the trait, L is the random effect of line, and E is the error. Broad sense heritability was estimated as $H^2 = \sigma_L^2 / (\sigma_L^2 + \sigma_E^2)$. Lastly, we computed Pearson product moment correlations to study the interrelationships between metrics used to quantify micro-environmental plasticity and between the micro-environmental plasticity and other traits.

8. SUPPLEMENTARY DATA

TABLE S1

TABLE S1.A

Larval age (hours AEL)	Day 1 samples				Day 2 samples			
	Mean (h)	Std (h)	CV (%)	N	Mean (h)	Std (h)	CV (%)	N
34	84.8	20.6	24.3	96	84.6	26.4	31.3	96
36	92.8	22.1	23.8	93	91.4	28.1	30.8	96
38	106.6	30.8	28.9	94	108.5	35.1	32.3	95
40	125.6	39.8	31.7	95	130.9	39.4	30.1	94
42	151.4	48.4	32.0	96	164.0	40.2	24.5	96
44	183.7	42.6	23.2	96	176.0	44.0	25.0	95
46	230.4	66.0	28.6	96	200.1	50.7	25.3	95
48	207.9	73.5	35.4	96	216.5	66.1	30.5	95
50					188.9	72.5	38.4	95
60	164.0	64.5	39.4	93				

TABLE S1.B

Larval age (hours AEL)	Mean (μ g)	Std (μ g)	CV (%)	n (x10)
35.5	45.8	2.1	4.5	5
37.5	57.1	1.5	2.7	6
39.5	63.2	1.8	2.9	6
41.5	78.5	1.7	2.2	7
43.5	94.6	2.9	3.1	6
45.5	115.4	4.6	4.0	6
47.5	145.5	4.7	3.2	6
59.5	381.7	14.1	3.7	6

TABLE S1. Starvation resistance as a function of larval age. (A) Starvation resistances were measured in larvae from 34 to 60 hours after egg-laying (AEL) in 2 successive days using the same parents. Means and standard deviations (both in hours) were calculated for each replicate sample of N larvae. (B) Wet weight of age-matched larvae (μ g per larva) was measured using n groups of 10 larvae. Std: standard deviation, CV: coefficient of variation.

TABLE S2

Line ID		Mean survival time (h)	Number of larvae	Number of replicates	SE (h)
Bloomington	RAL				
28176	350	72.16	497	11	3.33
28166	309	79.13	260	6	1.63
28147	158	79.49	627	17	5.41
28135	88	83.11	291	8	4.08
28185	374	83.19	518	12	4.72
25185	358	83.52	504	12	5.12
28165	287	88.33	278	6	2.18
25208	820	88.73	528	12	4.57
25182	324	88.81	538	12	5.18
25206	786	90.56	563	12	3.25
25745	714	92.89	447	13	8.60
28197	440	92.92	272	6	4.45
28204	502	93.11	1092	24	7.64
25209	852	93.94	962	19	6.48
37525	306	94.42	525	9	11.80
28191	385	94.48	781	18	7.10
28139	105	94.89	260	6	5.92
28173	338	95.02	679	15	6.61
28260	897	95.12	440	12	3.53
25445	365	95.14	594	18	5.20
29658	439	95.23	318	8	3.98
25175	301	95.25	824	15	8.85
28174	340	95.32	567	12	4.14
25188	375	95.50	488	11	5.58
28132	75	96.47	269	6	2.70
28278	409	96.74	270	6	4.38
28211	563	96.90	253	6	6.05
25193	427	96.93	610	14	2.86
25194	437	97.52	574	13	6.55
28232	790	97.63	272	7	7.06
28254	879	97.63	281	6	6.92
25192	399	97.83	599	14	5.73
28194	392	98.61	500	11	3.68
28162	256	98.73	358	8	4.29
28234	801	98.81	507	12	5.75
28177	352	99.35	268	6	5.26
28244	822	99.53	499	12	6.30
28251	855	100.42	555	12	3.94
25176	303	100.46	272	6	3.25
28265	913	100.47	244	6	8.06
25177	304	100.62	427	9	6.88

TABLE S2 (continued)

Line ID		Mean survival time (h)	Number of larvae	Number of replicates	SE (h)
Bloomington	RAL				
25189	379	100.68	1065	24	7.07
55027	559	100.99	303	7	3.31
28184	373	101.16	336	8	3.24
28250	853	101.62	278	6	4.74
28220	721	101.80	470	12	5.36
28274	85	101.94	567	12	9.90
28182	370	101.95	539	12	2.19
28239	810	102.42	801	18	4.07
28243	821	102.56	533	12	5.00
28181	367	102.63	279	6	5.30
28199	443	102.70	555	12	4.78
28168	318	102.88	758	18	6.36
28224	748	102.97	256	6	3.00
28192	386	103.45	486	12	6.31
28151	181	104.34	245	6	5.65
28160	237	104.66	526	12	4.84
28219	716	104.72	277	6	3.79
25199	639	105.03	281	6	3.72
28125	38	105.25	303	7	7.89
28198	441	105.30	262	6	5.07
28257	890	105.49	264	6	7.12
28259	894	106.40	268	6	4.29
28262	907	106.90	280	6	3.08
29651	40	107.21	495	12	3.68
29654	320	107.36	324	8	4.09
28126	41	107.65	274	6	4.47
29655	321	107.80	500	12	7.80
25203	732	107.83	276	6	4.10
28247	843	107.83	288	6	5.16
28161	239	108.19	441	10	6.79
25184	357	108.59	271	6	6.19
25181	315	110.29	501	12	4.23
28222	737	110.35	283	6	4.13
28189	382	110.46	1090	22	6.75
28179	359	110.66	326	6	3.24
29657	406	110.89	267	6	5.78
28129	59	110.97	277	6	2.68
28236	804	111.52	262	6	6.12
25186	360	111.67	277	6	4.82
29652	57	111.99	676	18	7.83
55022	395	111.99	286	6	7.77

TABLE S2 (continued)

Line ID		Mean survival time (h)	Number of larvae	Number of replicates	SE (h)
Bloomington	RAL				
28227	761	112.33	264	6	3.32
25191	391	114.09	237	7	5.22
25195	486	114.57	394	9	4.41
55031	630	114.65	284	6	2.33
28190	383	114.76	275	6	3.18
28207	531	115.17	278	6	4.40
28145	149	116.37	505	11	7.65
28256	884	116.41	266	6	2.00
28229	776	117.54	260	6	8.05
28261	900	117.65	263	6	6.01
28235	802	117.71	266	6	6.18
28150	177	118.24	269	6	7.83
28146	153	118.76	519	12	9.46
28249	850	118.99	280	6	4.67
28230	783	120.41	559	13	4.63
28203	492	120.52	512	12	9.34
28264	911	120.52	278	6	6.23
28153	195	120.92	495	12	8.26
28142	136	121.57	267	6	4.02
28248	849	122.45	269	6	3.53
28154	217	122.52	264	6	6.26
28231	787	122.57	761	13	9.55
28238	808	122.70	278	7	3.40
28178	356	122.73	254	6	3.96
25198	555	122.91	520	9	4.21
28212	584	124.08	365	7	5.53
28196	426	124.35	250	6	4.42
28123	26	124.66	398	9	6.82
25190	380	125.48	259	6	4.25
28223	738	125.55	210	6	4.37
25180	313	127.05	267	6	13.71
28208	535	129.74	257	6	5.83
28156	227	131.30	294	6	4.96
28136	91	131.70	278	6	7.08
25210	859	132.65	243	6	6.38
28149	176	132.93	224	5	5.21
28180	361	133.59	261	6	9.49
29656	405	134.38	552	12	6.12
28164	280	138.97	455	10	5.61
25183	335	139.49	262	6	8.18
25187	362	141.02	282	6	5.65

TABLE S2 (continued)

Line ID		Mean survival time (h)	Number of larvae	Number of replicates	SE (h)
Bloomington	RAL				
28157	228	149.01	266	6	7.77
29660	530	149.79	807	15	7.83
28138	101	150.43	274	6	4.33
25200	707	160.59	257	6	5.42
28233	796	161.62	551	12	5.47
25205	774	166.72	276	6	6.18

TABLE S2. Larval starvation resistance in the DGRP. Line IDs are given by their Bloomington stock numbers and DGRP names (RAL). SE is the standard error of the replicate sample means.

TABLE S3

Analysis	n	Estimate	Standard error	t value	p value	
(Intercept)	58	107.68	2.22	48.50	<2e-16	***
<i>Wolbachia</i> (+)	71	4.19	2.99	1.40	0.164	
(Intercept)	101	110.40	1.70	64.93	<2e-16	***
In2L_t (1)	17	-1.65	4.48	-0.37	0.713	
In2L_t (2)	11	-2.33	5.43	-0.43	0.669	
(Intercept)	120	110.61	1.55	71.57	<2e-16	***
In2R_NS (1)	6	-6.62	7.08	-0.93	0.352	
In2R_NS (2)	3	-13.51	9.90	-1.37	0.175	
(Intercept)	123	109.87	1.54	71.47	<2e-16	***
In3R_P (1)	4	6.58	8.66	0.76	0.449	
In3R_P (2)	2	-5.82	12.15	-0.48	0.633	
(Intercept)	120	110.56	1.55	71.42	<2e-16	***
In3R_K (1)	8	-7.33	6.19	-1.18	0.239	
In3R_K (2)	1	-15.67	17.03	-0.92	0.359	
(Intercept)	112	110.83	1.58	70.28	<2e-16	***
In3R_Mo (1)	5	6.57	7.63	0.86	0.391	
In3R_Mo (2)	12	-11.76	5.07	-2.32	0.022	*

TABLE S3. Regression analysis for the effects of *Wolbachia* infection and common polymorphic inversions on mean starvation resistances. Reduced linear models were used to test the effect of each factor, separately. Estimate indicates the grand mean for the lines in the intercept group (lines that are not infected or lines that do not have inversions) and the effect of the factors *Wolbachia* infection or inversions (1) for heterozygous, (2) for homozygous. n: number of lines in each group, p value: Probability(>|t|)

TABLE S4

SNP			p value	MAF	Effect size	Genic class	Associated genes
2L	15163911	C / T	5.2e-06	0.144	9.5	intron	<i>CG4168</i>
2L	15163913	C / A	9.7e-07	0.150	9.9	intron	<i>CG4168</i>
2L	16284226	G / A	2.4e-06	0.074	13.1	intron	<i>crp</i>
2L	16284690	G / T	1.3e-06	0.074	13.4	intron	<i>crp</i>
2L	16285589	G / T	2.9e-06	0.073	13.0	intron	<i>crp</i>
2L	16286529	C / G	8.4e-07	0.063	14.3	intron	<i>crp</i>
2L	16288069	C / G	3.0e-06	0.063	13.7	inter (+221)	<i>pkaap</i>
						inter (+339)	<i>crp</i>
2L	16881596	A / G	3.7e-06	0.079	12.2	3' UTR	<i>CG34313</i>
						non-syn	<i>CG31743</i>
2L	2531414	C / G	8.7e-06	0.098	10.4	inter	
2L	2807139	G / T	4.7e-06	0.063	13.4	intron	<i>daw</i>
2L	2807196	A / G	7.5e-06	0.072	12.4	intron	<i>daw</i>
2L	3910263	C / T	1.2e-06	0.142	9.4	intron	<i>fred</i>
2L	4965854	T / A	5.3e-07	0.087	12.6	intron	<i>CG34125</i>
						inter (-424)	<i>CG34124</i>
						inter (+429)	<i>CR43824</i>
						inter (-513)	<i>Trip1</i>
2L	7635298	C / A	8.8e-06	0.055	13.9	inter (+815)	<i>CR44799</i>
2L	7635357	T / G	8.8e-06	0.055	13.9	inter (+756)	<i>CR44799</i>
2L	9938390	C / T	7.6e-06	0.067	13.0	syn	<i>Apf</i>
						exon	<i>CR45007</i>
						inter (+256)	<i>CG31712</i>
						inter (+473)	<i>CG4619</i>
2R	11506182	C / T	4.4e-06	0.055	14.3	intron	<i>CG42524</i>
2R	11506370	G / A	4.0e-06	0.054	14.4	intron	<i>CG42524</i>
2R	11508395	G / A	4.0e-06	0.054	14.4	intron	<i>CG42524</i>
2R	11508889	T / C	4.4e-06	0.055	14.3	intron	<i>CG42524</i>
2R	12390991	T / C	7.2e-06	0.102	10.2	intron	<i>Sema-2a</i>
2R	12982826	C / T	4.1e-06	0.063	13.5	intron	<i>GstS1</i>
2R	13700322	T / A	6.8e-07	0.127	10.6	intron	<i>grh</i>
2R	16517053	C / T	5.9e-06	0.197	8.1	non-syn	<i>Rcd6</i>
2R	17044324	T / G	9.0e-07	0.079	12.9	intron	<i>CG30389</i>
2R	17455235	C / T	2.9e-07	0.112	11.4	intron	<i>CG10440</i>
2R	4619430	C / T	1.6e-06	0.063	14.0	intron	<i>stmA</i>
3L	10706619	T / A	2.3e-06	0.134	9.8	inter	
3L	10706838	T / C	4.0e-06	0.295	6.5	inter	
3L	10994504	A / G	2.0e-06	0.149	8.9	intron	<i>klu</i>
						inter (-5)	<i>snoRNA:</i>
							<i>Me18S-G962</i>
3L	11341705	A / T	3.3e-06	0.063	13.4	inter (-790)	<i>CR45251</i>
3L	12672152	A / G	6.7e-06	0.056	14.1	inter	

TABLE S4 (continued)

SNP			p value	MAF	Effect size	Genic class	Associated genes
3L	15093599	C / T	7.9e-06	0.066	13.0	exon	<i>CR43972</i>
						intron	<i>CTPsyn</i>
						inter (+997)	<i>CR43973</i>
3L	21839288	G / T	6.6e-07	0.063	14.3	intron	<i>mub</i>
3L	21895803	C / T	7.8e-06	0.063	13.1	intron	<i>mub</i>
3L	21898710	T / A	9.4e-06	0.064	13.0	intron	<i>mub</i>
3L	2190527	C / A	3.8e-06	0.094	11.2	intron	<i>CG5707</i>
3L	22101842	C / A	3.0e-06	0.220	8.3	inter	
3L	22107174	G / A	6.7e-06	0.239	7.8	inter	
3L	5266330	T / A	5.8e-06	0.055	14.2	intron	<i>shep</i>
3L	5541278	A / G	7.5e-06	0.055	12.7	intron	<i>CG10625</i>
3L	6185831	A / C	7.3e-07	0.085	12.4	intron	<i>CG13298</i>
						inter (-682)	<i>CG7386</i>
						inter (+749)	<i>Trn</i>
3L	9062242	G / T	9.0e-06	0.135	9.3	inter (-507)	<i>CG4911</i>
3L	9941626	C / T	3.6e-06	0.320	7.1	syn	<i>CG34356</i>
3R	11290098	G / A	5.2e-06	0.124	9.7	intron	<i>AdamTS-A</i>
3R	12356551	C / T	9.9e-08	0.056	16.4	intron	<i>CG42342</i>
3R	13368494	A / T	9.4e-07	0.055	14.8	intron	<i>SF1</i>
3R	13773529	C / G	7.7e-08	0.055	16.5	intron	<i>cpo</i>
3R	13773530	A / G	7.3e-08	0.055	16.5	intron	<i>cpo</i>
3R	13773535	G / A	7.3e-08	0.055	16.5	intron	<i>cpo</i>
3R	16413624	G / C	4.3e-06	0.115	10.5	intron	<i>Sep2</i>
3R						inter (+809)	<i>Srp72</i>
3R	16415827	G / A	4.8e-06	0.256	7.6	syn	<i>Srp72</i>
3R						inter (-919)	<i>CG4159</i>
3R	18491213	A / G	6.4e-06	0.078	11.9	intron	<i>wake</i>
3R	18491216	C / A	8.3e-06	0.079	11.7	intron	<i>wake</i>
3R	18491226	C / T	6.4e-06	0.078	11.9	intron	<i>wake</i>
3R	18724234	G / C	2.6e-06	0.205	8.3	inter (+772)	<i>klg</i>
3R	19115512	A / G	8.7e-06	0.311	7.0	syn	<i>CG4434</i>
						inter (-441)	<i>pnt</i>
3R	21423620	C / A	4.6e-06	0.059	14.3	intron	<i>ymp</i>
						intron	<i>msi</i>
3R	21423712	T / G	5.2e-06	0.068	13.4	intron	<i>ymp</i>
						intron	<i>msi</i>
3R	24570337	A / C	5.0e-06	0.085	12.1	5' UTR	<i>AstA-R2</i>
3R	24570338	A / T	5.2e-07	0.077	13.9	5' UTR	<i>AstA-R2</i>
3R	24882319	C / T	4.0e-06	0.175	8.9	syn	<i>Sirt7</i>
						inter (-760)	<i>CG11843</i>
3R	26084411	C / T	2.4e-06	0.058	14.7	inter	
3R	26231165	T / C	5.5e-06	0.148	9.3	inter (-454)	<i>CR45570</i>

TABLE S4 (continued)

SNP			p value	MAF	Effect size	Genic class	Associated genes
3R	8651023	A / G	1.7e-06	0.169	9.4	inter	<i>snk</i> <i>CG11670</i>
3R	8679862	A / T	8.8e-06	0.171	8.8	inter	
3R	8679884	A / G	7.5e-06	0.176	8.7	inter	
3R	8680215	A / C	9.2e-06	0.175	8.6	inter	
3R	8680245	G / A	7.8e-06	0.154	9.2	inter	
3R	8680255	T / G	5.2e-06	0.160	9.2	inter	
3R	8694533	C / G	7.3e-06	0.168	8.8	inter	
3R	8695141	A / T	9.9e-06	0.174	8.5	inter	
3R	8697518	G / T	4.5e-06	0.165	9.0	inter	
3R	8867808	A / C	9.2e-06	0.119	10.2	intron 5' UTR	
3R	8991996	T / A	6.9e-07	0.080	13.0	inter	<i>cac</i> <i>CG12479</i> <i>CG5004</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i> <i>CG6123</i>
3R	9868276	T / C	7.9e-08	0.087	13.4	inter	
3R	9868284	A / G	7.9e-08	0.087	13.4	inter	
3R	9868414	G / T	9.4e-07	0.080	12.9	inter	
3R	9868557	C / A	2.3e-06	0.063	13.8	inter	
3R	9868800	C / G	1.3e-06	0.056	15.0	inter	
X	11847399	T / C	2.0e-06	0.125	10.2	intron	
X	14046112	G / A	4.8e-06	0.242	7.4	inter (+73)	
X	17024160	G / T	7.6e-06	0.198	8.0	intron	
X	18223599	G / A	6.1e-06	0.206	8.0	intron	
X	18223649	C / T	1.2e-06	0.197	8.7	intron	
X	18223652	T / A	5.7e-07	0.190	9.0	intron	
X	18223677	G / A	5.7e-06	0.200	8.1	intron	
X	18223869	C / G	2.3e-07	0.083	13.5	intron	
X	18223909	C / T	5.3e-07	0.084	13.2	intron	
X	18224084	G / A	9.4e-08	0.087	13.3	intron	
X	18224543	A / G	1.1e-06	0.177	9.1	5' UTR	
X	20501990	C / A	7.4e-07	0.062	13.4	inter	
X	20501991	C / T	7.4e-07	0.062	13.4	inter	
X	20502384	T / A	1.9e-07	0.054	15.0	inter	

TABLE S4. Genetic variants associated with larval starvation resistance.

Associated variants ($p < 1.e-05$) are listed with their chromosomal arm, location and type of polymorphism (major allele / minor allele). MAF: Minor allele frequency in the set of 129 DGRP lines. Effect sizes were estimated by half of the difference between the means of major and minor alleles: (minor-major)/2. Genic context refers to the SNP's location relative to genes. inter: intergenic, UTR: untranslated region, exon: exonic in non-coding RNA genes, syn:

synonymous SNPs, non-syn: non-synonymous SNPs. Distance of intergenic SNPs to nearby genes (within 1000 bp) is given in parenthesis. All genes located within 1000 bp of the associated SNPs are listed.



TABLE S5

Candidate gene	FlyBase ID	Chr	Molecular function	Human homolog
<i>CG4168</i>	FBgn0028888	2L	insulin-binding protein	<i>IGFALS</i>
<i>crp</i>	FBgn0001994	2L	transcription factor	<i>TFAP4</i>
<i>pkaap</i>	FBgn0040079	2L	PKA binding protein	<i>AKAP10</i>
<i>CG34313</i>	FBgn0085342	2L	small nuclear ribonucleoprotein	<i>SNRNP25</i>
<i>CG31743</i>	FBgn0032618	2L	carbohydrate sulfotransferase	<i>CHST11; CHST13; CHST12; CHST14; CHST8</i>
<i>daw</i>	FBgn0031461	2L	TGF- β ligand; cytokine	<i>TGFB1; TGFB2; INHBE; TGFB3; MSTN; GDF11</i>
<i>fred</i>	FBgn0051774	2L	membrane protein	<i>KIRREL2; KIRREL3; KIRREL</i>
<i>CG34125</i>	FBgn0083961	2L		<i>DAZAP2</i>
<i>CG34124</i>	FBgn0083960	2L		<i>CFAP44</i>
<i>CR43824</i>	FBgn0264372	2L		
<i>Trip1</i>	FBgn0015834	2L	translation initiation factor	<i>EIF3I</i>
<i>CR44799</i>	FBgn0266034	2L		
<i>CR44799</i>	FBgn0266034	2L		
<i>Apf</i>	FBgn0051713	2L	diadenosine tetraphosphate hydrolase	<i>NUDT2</i>
<i>CR45007</i>	FBgn0266359	2L		
<i>CG31712</i>	FBgn0051712	2L		<i>ARGLU1</i>
<i>CG4619</i>	FBgn0032166	2L		
<i>CG42524</i>	FBgn0260429	2R		
<i>Sema-2a</i>	FBgn0011260	2R	secreted adhesion protein	<i>SEMA3A; SEMA4D; SEMA3C; SEMA4C; SEMA3F; SEMA4A; SEMA4G; SEMA3E</i>
<i>GstS1</i>	FBgn0010226	2R	Glutathione S transferase S1	<i>HPGDS; GSTA3</i>
<i>grh</i>	FBgn0259211	2R	transcription factor	<i>GRHL1; GRHL2; GRHL3</i>
<i>Rcd6</i>	FBgn0034530	2R	membrane protein	
<i>CG30389</i>	FBgn0050389	2R	membrane protein	<i>TMEM57</i>
<i>CG10440</i>	FBgn0034636	2R	potassium channel tetramerization protein	<i>KCTD15; KCTD1</i>
<i>stmA</i>	FBgn0086784	2R	membrane protein	<i>EFR3B; EFR3A</i>

TABLE S5 (continued)

Candidate gene	FlyBase ID	Chr	Molecular function	Human homolog
<i>klu</i>	FBgn0013469	3L	DNA binding; regulation of transcription	<i>ZBTB7A; ZBTB14; ZBTB7B; ZBTB7C; ZNF837; ZNF543</i>
<i>snoRNA: Me18S-G962</i>	FBgn0086077	3L		
<i>CR45251</i>	FBgn0266786	3L		
<i>CR43972</i>	FBgn0264703	3L		
<i>CTPsyn</i>	FBgn0266452	3L	CTP synthase	<i>CTPS1; CTPS2</i>
<i>CR43973</i>	FBgn0264704	3L		
<i>mub</i>	FBgn0262737	3L	RNA binding; regulation of alternative splicing	<i>PCBP3; PCBP2; PCBP1; PCBP4</i>
<i>CG5707</i>	FBgn0026593	3L	serine hydrolase	<i>SERHL2; SERHL</i>
<i>shep</i>	FBgn0052423	3L	RNA binding; regulation of alternative splicing	<i>RBMS1; RBMS3; RBMS2</i>
<i>CG10625</i>	FBgn0035612	3L	cuticle protein	<i>COL24A1; COLEC12; COL6A2; COL1A2; COL28A1; AMACR; BSN</i>
<i>CG13298</i>	FBgn0035692	3L	splicing factor	<i>SF3B6</i>
<i>CG7386</i>	FBgn0035691	3L	DNA binding	<i>ZNF284; ZNF341; ZNF225; ZNF221; ZNF223; ZNF44; ZNF788</i>
<i>Trn</i>	FBgn0024921	3L	protein import into nucleus	<i>TNPO1; TNFO2</i>
<i>CG4911</i>	FBgn0035959	3L	regulation of ubiquitin-dependent degradation	<i>FBX033</i>
<i>CG34356</i>	FBgn0085385	3L	protein kinase	<i>SCYL2</i>
<i>AdamTS-A</i>	FBgn0038341	3R	secreted matrix metalloprotease	<i>ADAMTS9; ADAMTS20; ADAMTS1; ADAMTS8; ADAMTS4; ADAMTS15; ADAMTS5; ADAMTS18; ADAMTS16</i>
<i>CG42342</i>	FBgn0259244	3R	collagen	<i>COL13A1; COL23A1; COL25A1</i>
<i>SF1</i>	FBgn0025571	3R	splicing factor	<i>SF1</i>
<i>cpo</i>	FBgn0263995	3R	RNA binding	<i>RBPMs2; RBPMs</i>
<i>Septin2</i>	FBgn0014029	3R	GTPase	<i>SEPT6; SEPT11; SEPT8; SEPT10; SEPT14</i>

TABLE S5 (continued)

Candidate gene	FlyBase ID	Chr	Molecular function	Human homolog
<i>Srp72</i>	FBgn0038810	3R	co-translational targeting of secretory proteins into the ER	<i>SRP72</i>
<i>CG4159</i>	FBgn0038811	3R	pseudouridine synthase	<i>PUS1</i>
<i>wake</i>	FBgn0266418	3R	receptor binding; signal transduction	<i>ANKFN1</i>
<i>klg</i>	FBgn0017590	3R	membrane-associated protein	<i>MUC15; LSAMP; JAM2; NTM; OPCML; PVRL3</i>
<i>CG4434</i>	FBgn0039071	3R	glutamate dehydrogenase	<i>GLUD1; GLUD2</i>
<i>pnt</i>	FBgn0003118	3R	transcription factor	<i>ETS1; ETS2; ETV2</i>
<i>ymp</i>	FBgn0261287	3R	regulation of splicing	
<i>msi</i>	FBgn0011666	3R	RNA binding; regulation of translation	<i>MSI2; MSI1</i>
<i>AstA-R2</i>	FBgn0039595	3R	G-protein coupled receptor	<i>GALR2; GALR1; KISS1R</i>
<i>Sirt7</i>	FBgn0039631	3R	NAD ⁺ dependent deacetylase	<i>SIRT7</i>
<i>CG11843</i>	FBgn0039630	3R	serine protease	<i>TMPRSS12</i>
<i>CR45570</i>	FBgn0267130	3R		
<i>snk</i>	FBgn0003450	3R	secreted serine protease	<i>TMPRSS12; CMA1</i>
<i>CG11670</i>	FBgn0038114	3R	serine endopeptidase	<i>TMPRSS12</i>
<i>cac</i>	FBgn0263111	X	voltage-gated calcium channel	<i>CACNA1B; CACNA1E; CACNA1C; CACNA1D; CACNA1F</i>
<i>CG12479</i>	FBgn0040871	X		<i>MINOS1</i>
<i>CG5004</i>	FBgn0260748	X	regulatory protein	<i>PHLDB1; PHLDB2; PHLDB3</i>
<i>CG6123</i>	FBgn0030913	X		

TABLE S5. Candidate genes associated to larval starvation resistance.

Candidate genes are listed with their common names, FlyBase gene ID, the chromosome they are located on (Chr), their molecular function and names of their human homologs.

TABLE S6

Pathway ID	Pathway description	Observed gene count	False discovery rate
G0.0048513	organ development	43	0.000543
G0.0048731	system development	61	0.0013
G0.0007409	axonogenesis	17	0.00215
G0.0042330	taxis	18	0.00215
G0.0048666	neuron development	27	0.00215
G0.0048699	generation of neurons	32	0.00215
G0.0061564	axon development	17	0.00215
G0.0030182	neuron differentiation	29	0.00244
G0.0097485	neuron projection guidance	15	0.00402
G0.0022008	neurogenesis	42	0.00503
G0.0006935	chemotaxis	15	0.00583
G0.0007275	multicellular organismal development	67	0.00606
G0.0048856	anatomical structure development	71	0.00626
G0.0000904	cell morphogenesis involved in differentiation	20	0.00723
G0.0007399	nervous system development	44	0.00723
G0.0007411	axon guidance	14	0.00723
G0.0007444	imaginal disc development	24	0.00723
G0.0009653	anatomical structure morphogenesis	45	0.00723
G0.0009887	organ morphogenesis	27	0.00723

TABLE S6. Top20 biological process (GO) enrichment for genes associated with top 400 associated loci.

TABLE S7

Table S7.A

Major allele population		Minor allele population	
Bloomington ID	DGRP RAL no	Bloomington ID	DGRP RAL no
25191	391	25187	362
25195	486	25205	774
25197	517	28122	21
25206	786	28136	91
25207	799	28140	109
28135	88	28188	381
28144	142	28207	531
28145	149	28233	796
28157	228	28237	805
28171	332	28238	808
28189	382	29660	530
28220	721	55015	32
28234	801		
29652	57		
29653	229		
29654	320		

Table S7.B

Outbred population	Allele	N	Mean survival time (hours)	Std (h)	SE (h)	CV (%)
Major allele	CC	379	145.0	29.6	1.5	20.4
Minor allele	TT	367	158.1	31.9	1.7	20.2

Table S7.C

Starvation time (hours)	Major allele population		Minor allele population		t-test p value
	Mean	SE	Mean	SE	
0	0.00825	0.00135	0.00770	0.00141	0.7816
4	0.01590	0.00104	0.01510	0.00127	0.6324
10	0.01792	0.00223	0.02187	0.00219	0.2258

TABLE S7. Validation for 3R_12356551_SNP. (A) Founder DGRP lines used to generate the outbreeding major and minor allele populations. (B) Summary

statistics for larval starvation resistance of the outbreeding populations. p-value for the Mann-Whitney (Wilcoxon rank sum) test detected significant difference between major and minor allele population means: $p=7.555e-09$. N: sample size, Std: standard deviation, SE: standard error, CV: coefficient of variation. (C) Expression levels of *CG42342* relative to *Act5C* during starvation.



TABLE S8

Line	Female	Male	Difference (male- female)	p value	Sexes pooled	Sexes averaged
306	76.3 ±30.0 (85)	88.9 ±34.4 (92)	12.6	0.0076 **	82.8 ±32.9	82.6
852	85.7 ±23.7 (115)	96.9 ±27.3 (65)	11.2	0.0097 **	89.8 ±25.6	91.3
502	89.1 ±30.8 (126)	96.6 ±26.0 (60)	7.5	0.0054 **	91.5 ±29.5	92.8
338	91.2 ±28.8 (100)	99.3 ±32.2 (71)	8.1	0.1335	94.5 ±30.4	95.2
392	96.6 ±23.8 (99)	95.6 ±20.5 (87)	-0.9	0.7434	96.1 ±22.2	96.1
379	97.4 ±30.6 (95)	98.8 ±23.2 (74)	1.5	0.6386	98.0 ±27.6	98.1
57	114.0 ±45.0 (76)	111.3 ±50.4 (76)	-2.7	0.5881	112.6 ±47.7	112.6
555	116.9 ±24.7 (115)	123.2 ±20.7 (60)	6.3	0.0717	119.1 ±23.5	120.0
787	119.9 ±28.1 (81)	134.3 ±34.2 (103)	14.4	0.0005 ***	128.0 ±32.4	127.1
530	141.2 ±27.9 (98)	148.4 ±24.1 (87)	7.2	0.0871	144.6 ±26.3	144.8

TABLE S8. Test for sexual dimorphism in 10 DGRP lines. Line means ± standard deviations (hours) and sample size (in parenthesis) are given for females, males and sexes pooled data. Sexes pooled data represent the whole sample, whereas sexes averaged data refer to the average of female and male line means. p values show the significance levels of Wilcoxon rank sum test.

TABLE S9

Analysis	Source of variation	Variance component	Standard deviation	p-value		Fraction of total phenotypic variation
Sexes pooled ¹	Sex	fixed		0.0052	**	
	Line	364.387	19.089	5e-06	***	0.2850
	Line x Sex	6.666	2.582	0.26		0.0052
	Replicate	13.415	3.663	0.03	*	0.0105
	Error	894.022	29.900			0.6993
Sexes pooled ²	Sex	fixed		3e-06	***	
	Line	369.4	19.219	4e-06	***	0.2888
	Replicate	12.8	3.578	0.03	*	0.0100
	Error	897.0	29.951			0.7012
Sexes ignored ³	Line	376.3	19.398	2e-06	***	0.2905
	Replicate	10.8	3.286	0.07		0.0083
	Error	908.2	30.136			0.7012
Males ⁴	Line	375.246	19.371	9e-06	***	0.2824
	Replicate	4.002	2.001	0.7		0.0030
	Error	949.583	30.815			0.7146
Females ⁵	Line	360.080	18.976	6e-05	***	0.2921
	Replicate	22.880	4.783	0.02	*	0.0186
	Error	849.970	29.154			0.6894

TABLE S9. Quantitative genetic analysis of the sexed dataset with alternative models. Variances are estimated by linear regression using mixed effects models on the whole dataset that include sex as a factor ($Y = \mu + S + L + L \times S + R + \varepsilon^{(1)}$ and $Y = \mu + S + L + R + \varepsilon^{(2)}$), random effects model that ignores the sexual information ($Y = \mu + L + R + \varepsilon^{(3)}$), and reduced random effects models for each sex separately ($Y = \mu + L + R + \varepsilon^{(4)}$ ⁽⁵⁾).

TABLE S10

Dataset	Analysis	Mean	Standard deviation	r_{GMF} (p value)	r_{AL} (p value)
Adult large dataset (203 lines)	Female ^a	60.67	12.60	0.685 ($<2.2e-16$)	^{a-g} 0.086 (0.337)
	Male ^b	45.73	9.55		^{b-g} 0.048 (0.594)
	Sexes pooled ^c	53.20	10.18		^{c-g} 0.076 (0.393)
	Difference (m-f)	-14.93	9.22		
Adult small dataset (10 lines)	Female ^d	60.61	13.26	0.854 ($1.7e-03$)	^{d-h} 0.017 (0.964)
	Male ^e	44.65	8.76		^{e-i} 0.260 (0.468)
	Sexes pooled ^f	52.63	10.61		^{f-j} 0.100 (0.785)
	Difference (m-f)	-15.96	7.37		
Larvae large dataset (129 lines)	Sexes pooled ^g	109.97	16.98		
Larvae small dataset (10 lines)	Female ^h	102.82	19.65	0.958 ($1.3e-05$)	
	Male ⁱ	109.34	19.66		
	Sexes pooled ^j (averaged)	106.08	19.45		
	Difference (m-f)	6.52	5.68		

TABLE S10. Comparisons between larval and adult starvation resistance datasets in the DGRP. All estimates are calculated using line means. Mean and standard deviation are in hours. r_{GMF} and r_{AL} are Pearson's product moment correlation coefficients for the genetic correlations between males and females, and genetic correlations between adult and larval starvation resistances.

TABLE S11

Trait	Females		Males		Sexes pooled		Ref.
	<i>r</i>	p	<i>r</i>	p	<i>r</i>	p	
Adult starvation resistance	0.086	0.337	0.048	0.594	0.076	0.393	1
Chill coma recovery	0.054	0.559	0.063	0.493	0.061	0.515	
Virgin lifespan	0.013	0.880	-0.158	0.150	-0.042	0.707	2
Glucose content (high glucose diet)			0.067	0.522			3
Glucose content (low glucose diet)			0.089	0.387			
Glycerol content (high glucose diet)			-0.086	0.406			
Glycerol content (low glucose diet)			0.023	0.827			
Glycogen content (high glucose diet)			0.012	0.907			
Glycogen content (low glucose diet)			0.011	0.912			
Triglyceride content (high glucose diet)			0.282	0.006**			
Triglyceride content (low glucose diet)			-0.052	0.612			
Protein content (high glucose diet)			0.165	0.111			
Protein content (low glucose diet)			-0.078	0.453			
Weight (high glucose diet)			0.058	0.573			
Weight (low glucose diet)			0.003	0.977			
Night sleep	0.058	0.540	0.020	0.834	0.043	0.649	4
Day sleep	-0.045	0.638	0.040	0.674	-0.007	0.943	
Night sleep bout number	0.036	0.704	-0.092	0.334	-0.028	0.770	
Day sleep bout number	-0.166	0.081	-0.141	0.137	-0.179	0.059	
Night sleep bout length	0.006	0.949	0.082	0.392	0.046	0.633	
Day sleep bout length	0.088	0.357	0.134	0.158	0.132	0.164	

TABLE S11 (continued)

Trait	Females		Males		Sexes pooled		Ref.
	<i>r</i>	p	<i>r</i>	p	<i>r</i>	p	
Waking activity	0.210	0.027 * †	0.211	0.025 * †	0.221	0.019 * †	4
Hexanal avoidance	-0.201	0.024 * †	-0.184	0.040 * †	-0.204	0.022 * †	5
Food intake	-0.084	0.376	-0.076	0.426	-0.085	0.373	6
Aggression			-0.004	0.963			7
ER stress hazard ratio			0.186	0.101			8
ER stress t50			0.170	0.132			
Startle response (control)	-0.171	0.055	-0.160	0.074	-0.168	0.060	9
Startle response (MSB-treated)	-0.117	0.192	-0.075	0.402	-0.101	0.260	
Negative geotaxis (control)	-0.104	0.244	-0.011	0.906	-0.058	0.515	
Negative geotaxis (MSB-treated)	-0.200	0.025 *	-0.139	0.121	-0.178	0.046 *	
Alcohol sensitivity 1st exposure	0.072	0.419	0.198	0.024 *	0.144	0.104	10
Alcohol sensitivity 2nd exposure	0.083	0.350	0.130	0.143	0.113	0.204	
Alcohol tolerance	0.019	0.826	-0.053	0.552	-0.022	0.805	
Paraquat resistance	-0.014	0.886	0.104	0.273	0.046	0.632	11
MSB resistance	-0.094	0.326	-0.006	0.949	-0.059	0.538	
Centroid size	0.026	0.785	-0.037	0.701	-0.003	0.973	12
Interocular distance	0.031	0.749	-0.005	0.959	0.014	0.880	

TABLE S11 (continued)

Trait	Females		Males		Sexes pooled		Ref.
	<i>r</i>	p	<i>r</i>	p	<i>r</i>	p	
Genome size	0.067	0.447					13
Survival to adult (25°C)					0.169	0.373	
Survival to pupation (25°C)					0.229	0.224	
Minimum pupation time (25°C)					-0.181	0.339	
Pupal mass (25°C)	-0.107	0.574					
Eclosion time (25°C)	0.145	0.444					

TABLE S11. Correlations between larval starvation resistance and other organismal phenotypes in the DGRP. Pearson's correlation coefficient (*r*) and its p-value (p) was given for correlations between sexes-pooled larval line means and female, male and sexes-pooled (averaged) adult line means. References: 1. (Mackay et al., 2012); 2. (Arya et al., 2010); 3. (Unckless et al., 2015); 4. (Harbison et al., 2013); 5. (Arya et al., 2015); 6. (Garlapow et al., 2015); 7. (Shorter et al., 2015); 8. (Chow et al., 2013); 9. (Jordan et al., 2012); 10. (Morozova et al., 2015); 11. (Weber et al., 2012); 12. (Vonesch et al., 2016); 13. (Ellis et al., 2014). * significant associations (p<0.05). † correlations are due to a single outlier; must be ignored.

TABLE S12

Line	Mean survival time (h)	n	Mean σ_E	Std(σ_E)	ln(σ_E)	Std(ln(σ_E))
350	72.16	11	19.39	4.69	2.94	0.25
309	79.13	6	16.05	1.66	2.77	0.11
158	79.41	16	20.55	3.83	3.01	0.19
88	83.11	8	18.24	2.66	2.89	0.15
374	83.19	12	22.96	2.64	3.13	0.12
358	83.52	12	25.41	5.97	3.21	0.23
287	88.33	6	19.39	2.44	2.96	0.12
820	88.73	12	22.59	3.60	3.11	0.16
324	88.81	12	26.50	4.64	3.26	0.20
786	90.56	12	23.26	3.63	3.14	0.15
714	92.89	13	26.00	3.56	3.25	0.14
440	92.92	6	22.19	2.31	3.10	0.11
502	93.11	24	32.41	5.19	3.47	0.16
852	93.94	19	30.51	4.81	3.41	0.15
375	94.31	10	26.04	2.93	3.25	0.12
306	94.42	9	36.39	5.41	3.58	0.15
385	94.48	18	21.46	3.96	3.05	0.18
105	94.89	6	23.12	2.30	3.14	0.10
338	95.02	15	29.47	4.38	3.37	0.15
897	95.12	12	20.73	2.27	3.03	0.11
365	95.14	18	22.73	4.24	3.11	0.18
439	95.23	8	21.47	2.20	3.06	0.10
301	95.25	15	31.29	4.21	3.43	0.13
340	95.32	12	22.22	2.06	3.10	0.09
75	96.47	6	19.75	4.16	2.96	0.23
409	96.74	6	24.73	7.24	3.17	0.28
427	96.79	13	27.54	3.12	3.31	0.11
563	96.90	6	29.29	6.28	3.36	0.21
437	97.52	13	26.30	3.68	3.26	0.15
879	97.63	6	24.40	4.89	3.18	0.19
399	97.83	14	23.92	3.12	3.17	0.13
790	98.11	6	29.07	3.11	3.37	0.10
392	98.61	11	22.51	3.37	3.10	0.15
256	98.73	8	27.02	3.52	3.29	0.13
801	98.81	12	26.02	4.30	3.25	0.16
352	99.35	6	24.81	2.71	3.21	0.11
822	99.53	12	22.93	2.27	3.13	0.10
855	100.42	12	28.90	3.61	3.36	0.13
303	100.46	6	28.56	3.08	3.35	0.11
913	100.47	6	33.78	4.79	3.51	0.14
304	100.62	9	27.66	5.69	3.30	0.21
379	100.68	24	29.81	5.48	3.38	0.18

TABLE S12 (continued)

Line	Mean survival time (h)	n	Mean σ_E	Std(σ_E)	ln(σ_E)	Std(ln(σ_E))
373	101.16	8	23.73	3.77	3.16	0.16
853	101.62	6	25.31	4.60	3.22	0.18
559	101.68	6	21.18	1.75	3.05	0.08
721	101.80	12	30.28	5.02	3.40	0.17
370	101.95	12	23.17	3.16	3.13	0.14
85	102.12	11	24.38	4.88	3.17	0.22
810	102.42	18	20.15	4.40	2.98	0.21
821	102.56	12	22.64	1.71	3.12	0.08
367	102.63	6	28.65	2.94	3.35	0.10
443	102.70	12	22.86	2.53	3.12	0.11
318	102.88	18	21.83	3.89	3.07	0.18
748	102.97	6	21.21	3.42	3.04	0.15
386	103.45	12	28.26	2.37	3.34	0.09
181	104.34	6	24.70	4.09	3.19	0.18
237	104.66	12	22.28	2.88	3.10	0.13
716	104.72	6	32.76	2.87	3.49	0.09
639	105.03	6	24.09	1.92	3.18	0.08
38	105.25	7	27.50	4.42	3.30	0.15
441	105.30	6	24.91	3.92	3.20	0.16
890	105.49	6	28.73	4.29	3.35	0.14
894	106.40	6	23.09	4.14	3.13	0.18
907	106.90	6	19.83	2.29	2.98	0.11
40	107.21	12	21.35	1.97	3.06	0.09
320	107.36	8	20.73	5.15	3.01	0.24
321	107.80	12	40.30	6.05	3.69	0.15
732	107.83	6	22.19	3.09	3.09	0.14
843	107.83	6	32.98	5.04	3.49	0.15
239	108.19	10	28.85	3.35	3.36	0.12
357	108.59	6	26.97	2.98	3.29	0.12
41	108.64	5	27.14	3.95	3.29	0.15
315	110.29	12	20.21	4.82	2.98	0.25
737	110.35	6	17.73	2.40	2.87	0.13
382	110.46	22	37.91	8.03	3.61	0.23
359	110.66	6	21.21	2.21	3.05	0.11
406	110.89	6	22.53	3.52	3.10	0.16
59	110.97	6	22.15	3.82	3.09	0.16
804	111.52	6	35.52	5.07	3.56	0.16
360	111.67	6	31.48	2.20	3.45	0.07
395	111.99	6	19.61	2.10	2.97	0.11
57	112.11	17	44.13	4.98	3.78	0.11
761	112.33	6	24.63	3.55	3.19	0.16
391	113.78	6	30.55	1.62	3.42	0.05

TABLE S12 (continued)

Line	Mean survival time (h)	n	Mean σ_E	Std(σ_E)	ln(σ_E)	Std(ln(σ_E))
486	114.57	9	26.25	4.03	3.26	0.15
630	114.65	6	19.21	2.23	2.95	0.11
383	114.76	6	25.15	2.11	3.22	0.08
531	115.17	6	28.69	5.83	3.34	0.18
149	116.37	11	36.75	7.78	3.59	0.20
884	116.41	6	25.37	4.16	3.22	0.17
776	117.54	6	23.14	2.62	3.14	0.11
900	117.65	6	34.70	6.31	3.53	0.18
802	117.71	6	26.78	3.01	3.28	0.11
177	118.24	6	28.65	5.70	3.34	0.20
153	118.76	12	30.10	4.91	3.39	0.17
850	118.99	6	22.06	2.28	3.09	0.10
783	120.41	13	28.25	3.42	3.33	0.12
492	120.52	12	24.71	5.18	3.19	0.19
911	120.52	6	35.42	4.39	3.56	0.12
195	121.23	11	35.09	2.35	3.56	0.07
136	121.57	6	26.30	4.31	3.26	0.16
849	122.45	6	30.62	3.32	3.42	0.11
217	122.52	6	32.82	3.39	3.49	0.10
787	122.57	13	35.72	5.98	3.56	0.17
808	122.70	7	25.07	5.44	3.20	0.24
356	122.73	6	32.32	3.47	3.47	0.11
555	122.91	9	27.20	2.83	3.30	0.10
584	124.08	7	34.89	6.08	3.54	0.17
426	124.35	6	23.52	3.63	3.15	0.17
26	124.66	9	28.45	6.09	3.33	0.21
380	125.48	6	37.23	1.41	3.62	0.04
738	125.55	6	36.46	2.84	3.59	0.08
313	127.05	6	29.18	3.21	3.37	0.11
535	129.74	6	31.47	2.79	3.45	0.09
227	131.30	6	34.99	1.34	3.55	0.04
91	131.70	6	30.19	4.58	3.40	0.15
859	132.65	6	32.88	4.88	3.48	0.15
176	132.93	5	27.27	5.43	3.29	0.18
361	133.59	6	38.48	5.82	3.64	0.16
405	134.38	12	24.79	3.39	3.20	0.14
280	138.97	10	42.63	4.58	3.75	0.11
335	139.49	6	35.72	5.82	3.57	0.15
362	141.02	6	30.19	9.01	3.37	0.27
228	149.01	6	33.42	5.11	3.50	0.17
530	149.79	15	33.89	5.71	3.51	0.17
101	150.43	6	39.23	9.77	3.64	0.28

TABLE S12 (continued)

Line	Mean survival time (h)	n	Mean σ_E	Std(σ_E)	$\ln(\sigma_E)$	Std($\ln(\sigma_E)$)
707	160.59	6	39.59	4.40	3.67	0.12
796	161.62	12	37.63	4.96	3.62	0.13
774	166.72	6	36.97	3.00	3.61	0.08

TABLE S12. Micro-environmental plasticity of larval starvation resistance. n: number of replicates; σ_E : within-replicate standard deviation. $\ln(\sigma_E)$: natural logarithm of σ_E ; Std(σ_E) and Std($\ln(\sigma_E)$): between-replicates standard deviations of σ_E and $\ln(\sigma_E)$. Survival times are in hours.



TABLE S13

Analysis	Source of variation	Variance	Standard deviation	p value	Fraction of total phenotypic variance
σ_E	Line	32.28	5.681	<2e-16 ***	0.630
	Error	18.92	4.350		
MAD	Line	15.48	3.934	<2e-16 ***	0.396
	Error	23.57	4.855		
CV_E	Line	18.50	4.301	<2e-16 ***	0.560
	Error	14.51	3.809		
$\ln(\sigma_E)$	Line	0.0413	0.203	<2e-16 ***	0.626
	Error	0.0247	0.157		
$\ln(\text{MAD})$	Line	0.0457	0.214	<2e-16 ***	0.370
	Error	0.0779	0.279		
$\ln(CV_E)$	Line	0.0283	0.168	<2e-16 ***	0.559
	Error	0.0223	0.149		

TABLE S13. Quantitative genetic analysis for micro-environmental plasticity in larval starvation resistance. σ_E : standard deviation, CV : coefficient of variation, MAD: median absolute deviation, $\ln$: natural logarithm. All measures were based on within-line within-replicate estimates.

TABLE S14

TABLE S14.A

Metric x	Metric y	r_{xy}	p value
$\ln(\sigma_E)$	σ_E	0.992	< 2.2e-16
	CV_E	0.707	< 2.2e-16
	$\ln(CV_E)$	0.715	< 2.2e-16
	MAD	0.856	< 2.2e-16
	$\ln(MAD)$	0.851	< 2.2e-16

TABLE S14.B

Micro-environmental plasticity	Trait	r_{xy}	p value
σ_E	Larval starvation resistance (mean survival time)	0.580	6.0e-13
$\ln(\sigma_E)$		0.573	1.2e-12
CV_E		-0.161	0.068
$\ln(CV_E)$		-0.158	0.075
MAD		0.527	1.4e-10
$\ln(MAD)$		0.535	6.8e-11
$\ln(\sigma_E)$	Genome size	0.060	0.497
$\ln(\sigma_E)$	Percentage of segregating variants	-0.015	0.868
$\ln(\sigma_E)$	Number of heterozygous chromosomal inversions	-0.087	0.326

TABLE S14. Correlation analyses for micro-environmental plasticity in the DGRP. (A) Correlations between different metrics to measure micro-environmental plasticity. (B) Correlations between micro-environmental plasticity and other traits in the DGRP. r_{xy} : Pearson's *rho*.

TABLE S15

SNP			p value	Genic_class	Associated genes	Gene function
2L	1996220	C / T	3.9e-05	inter (-882)	<i>CG33543</i>	triglyceride lipase triglyceride lipase
2L	2417794	A / T	2.7e-05	intron	<i>CG34448</i>	
				inter (-628)	<i>CG34447</i>	
2L	2418353	A / T	4.0e-05	inter (-69)	<i>CG34447</i>	
				inter (+338)	<i>CG34448</i>	
2L	2418516	G / T	5.8e-05	synonym.	<i>CG34447</i>	
				inter (+551)	<i>CG34448</i>	
2L	2418533	C / T	4.7e-05	non-synonym.	<i>CG34447</i>	
				inter (+568)	<i>CG34448</i>	
2L	5861767	G / A	5.6e-05	intron	<i>rau</i>	RAS signaling synaptic junction sulfotransferase
2L	6379467	G / A	1.1e-05	intron	<i>CG42368</i>	
2L	6521190	G / A	4.0e-05	intron	<i>CG31637</i>	
2L	8742817	A / T	3.2e-05			
2L	8742871	T / G	2.9e-05			
2L	8742883	A / C	2.4e-05			
2L	8742884	T / A	4.2e-05			
2L	8742905	A / T	9.3e-06			
2L	9584417	G / A	8.3e-06	inter (-253)	<i>CR43621</i>	
2L	9982184	G / T	4.8e-05			
2L	10018496	A / C	1.7e-05			ubiquitin-specific protease Cdk activator
2L	10307135	C / T	2.3e-05	intron	<i>CG4968</i>	
				inter (+541)	<i>CG5390</i>	
				inter (-996)	<i>Cdk5alpha</i>	
2L	10307160	A / G	7.0e-06	intron	<i>CG4968</i>	
				inter (+556)	<i>CG5390</i>	
				inter (-971)	<i>Cdk5alpha</i>	
2L	11956912	C / A	5.6e-05	exon	<i>CR43314</i>	
2L	12483685	G / T	4.5e-05	intron	<i>bun</i>	
2L	12484503	G / C	1.4e-05	intron	<i>bun</i>	transcription factor transcription factor
2L	12484651	C / T	1.2e-05	intron	<i>bun</i>	
2L	12484795	T / G	1.3e-05	intron	<i>bun</i>	
2L	12686469	T / C	2.1e-05	inter (-163)	<i>pdm2</i>	
2L	13082902	C / T	4.8e-05			
2L	17717698	A / T	3.3e-05	intron	<i>CadN</i>	
2L	20430075	A / T	2.7e-05	intron	<i>CG31688</i>	

TABLE S15 (continued)

SNP			p value	Genic_class	Associated genes	Gene function
2R	1838469	C / A	4.0e-05			
2R	4710730	G / A	5.4e-05	intron	<i>sns</i>	cell adhesion
2R	6557712	C / G	5.9e-05			
2R	7393165	T / C	4.3e-05	intron	<i>inv</i>	transcription factor
2R	8182411	C / T	1.8e-05	intron	<i>CG42700</i>	
2R	9707435	C / T	5.7e-05	3'UTR inter (-359)	<i>fl(2)d</i> <i>stj</i>	mRNA splicing voltage-gated calcium channel
2R	9716049	T / A	4.7e-05	intron	<i>CG6329</i>	
2R	9716116	C / T	1.4e-05	intron	<i>CG6329</i>	
2R	9716118	A / G	1.0e-05	intron	<i>CG6329</i>	
2R	11520690	G / A	5.7e-06	intron	<i>CG42524</i>	
2R	12390991	T / C	2.2e-05	intron	<i>Sema-2a</i>	secreted adhesion protein
2R	13766091	T / G	2.0e-05	synonym. intron	<i>CG30325</i> <i>Elk</i>	voltage-gated K ⁺ channel
2R	14088982	A / T	4.4e-05			
2R	15240662	A / G	3.9e-05	intron	<i>Rgk1</i>	GTPase
2R	17682067	G / T	3.7e-05			
3L	1635476	G / A	3.7e-05	intron	<i>nSyb</i>	vesicle fusion
3L	1650358	G / T	1.3e-05	intron	<i>CG13920</i>	transmembrane protein
3L	1650371	A / G	1.4e-05	intron	<i>CG13920</i>	
3L	1650379	C / T	2.4e-05	intron	<i>CG13920</i>	
3L	1653968	A / G	3.0e-05	3'UTR	<i>CG7970</i>	peroxisome membrane protein
3L	1653998	A / G	4.1e-05	inter (-279) 3'UTR	<i>CG7971</i> <i>CG7970</i>	splicing factor peroxisome membrane protein
3L	2990700	G / T	3.0e-06	inter (-249)	<i>CG7971</i>	splicing factor
3L	4620951	C / A	5.3e-05	intron inter (+268)	<i>Src64B</i> <i>CG32246</i>	tyrosine kinase
3L	4661628	T / A	5.1e-05	intron	<i>axo</i>	transmission of nerve impulse
3L	4852778	A / G	3.4e-05	inter (+317)	<i>CG32237</i>	
3L	4852790	T / A	2.0e-05	inter (+329)	<i>CG32237</i>	

TABLE S15 (continued)

SNP			p value	Genic_class	Associated genes	Gene function
3L	4852803	C / T	5.0e-05	inter (+342)	<i>CG32237</i>	microtubule motor activity
3L	4852807	A / G	1.8e-05	inter (+346)	<i>CG32237</i>	
3L	4852819	A / C	1.5e-05	inter (+358)	<i>CG32237</i>	
3L	4852835	A / C	1.5e-05	inter (+374)	<i>CG32237</i>	
3L	4866063	G / C	3.4e-05	intron	<i>CG17150</i>	
3L	6152783	T / G	4.2e-05	intron	<i>Zasp66</i>	muscle alpha-actinin binding
3L	8626698	C / G	3.9e-05			
				non-synonym.	<i>Cdc6</i>	cell division
3L	8952030	T / G	1.4e-05	5'UTR	<i>CG13313</i>	
3L	8952880	C / A	2.4e-05	intron	<i>CG13313</i>	mitochondrial valine-tRNA synthetase oxidoreductase
3L	8953007	A / T	4.0e-06	synonym.	<i>CG13313</i>	
3L	8956411	T / A	2.7e-05	synonym.	<i>ValRS-m</i>	
3L	8958361	T / G	5.0e-06	non-synonym.	<i>CG5653</i>	
				inter (+492)	<i>ValRS-m</i>	
3L	8958367	C / T	5.0e-06	non-synonym.	<i>CG5653</i>	PI3 phosphatase translation initiation PI3 phosphatase
				inter (+498)	<i>ValRS-m</i>	
3L	8958761	A / T	3.8e-05	synonym.	<i>CG5653</i>	
				inter (+892)	<i>ValRS-m</i>	
3L	8958815	T / C	2.1e-06	synonym.	<i>CG5653</i>	
				inter (+946)	<i>ValRS-m</i>	PI3 phosphatase translation initiation PI3 phosphatase
3L	8962604	G / A	3.9e-05	synonym.	<i>CG5026</i>	
3L	8965679	G / A	1.1e-05	synonym.	<i>pix</i>	
				inter (-827)	<i>CG5026</i>	GABA-gated Cl ⁻ channel
3L	8965738	A / G	2.3e-05	intron	<i>pix</i>	
				inter (-886)	<i>CG5026</i>	
3L	9172138	C / T	3.1e-05	intron	<i>Rdl</i>	autophagy
3L	12546352	A / C	3.1e-05	intron	<i>sowah</i>	
3L	12546399	T / A	4.3e-07	intron	<i>sowah</i>	
3L	12546400	C / T	2.0e-06	intron	<i>sowah</i>	
3L	12789250	A / C	5.5e-05	intron	<i>Atg1</i>	
				inter (-66)	<i>CG42709</i>	helicase
3L	15351740	C / G	4.1e-05	5'UTR	<i>Chd3</i>	
3L	19440280	T / A	5.6e-05	inter (+736)	<i>CG33062</i>	

TABLE S15 (continued)

SNP			p value	Genic_class	Associated genes	Gene function
3R	701906	T / C	5.5e-05	non-synonym.	<i>laf</i>	synaptic vesicle exocytosis
3R	18633068	C / T	3.7e-05	inter (-56)	<i>CG7031</i>	
3R	20382380	T / A	5.5e-05	inter (+34)	<i>snRNA:U6:96Ab</i>	
				inter (-109)	<i>Esyt2</i>	
				inter (-464)	<i>snRNA:U6:96Aa</i>	TOR signaling
				inter (+557)	<i>snRNA:U6:96Ac</i>	
				inter (-710)	<i>REPTOR</i>	
3R	20593299	G / A	3.1e-05	inter (-151)	<i>CR45650</i>	
3R	23373865	A / G	5.2e-05	3'UTR	<i>Klp98A</i>	microtubule binding
				inter (-681)	<i>Gr98d</i>	gustatory receptor
3R	24380706	A / C	1.9e-05			mRNA splicing
3R	24642509	G / A	5.0e-05	synonym	<i>CG1646</i>	
				inter (-555)	<i>CG1523</i>	transmembrane transport
3R	24991020	C / T	2.9e-05	inter (-332)	<i>CG11897</i>	
				inter (+345)	<i>CG11898</i>	
3R	25194577	G / T	1.2e-05	inter (+483)	<i>CG45546</i>	autophagy
3R	25659763	A / C	4.2e-05	synonym	<i>Atg16</i>	
X	5686229	G / A	8.2e-06	5'UTR	<i>CG12239</i>	synaptic junction
				inter (+660)	<i>CR44832</i>	
X	5687408	G / A	5.9e-05	synonym	<i>CG12239</i>	
X	5687426	C / G	3.8e-05	synonym	<i>CG12239</i>	
X	5687432	G / A	5.6e-05	synonym	<i>CG12239</i>	
X	12085680	C / T	3.4e-05	intron	<i>Ten-a</i>	
X	12085681	T / A	3.4e-05	intron	<i>Ten-a</i>	
X	12085695	C / G	4.1e-05	intron	<i>Ten-a</i>	
X	12085703	G / A	3.5e-05	intron	<i>Ten-a</i>	arginine-tRNA synthetase
X	12085712	A / T	3.6e-05	intron	<i>Ten-a</i>	
X	12085716	G / A	1.9e-05	intron	<i>Ten-a</i>	
X	12085719	A / C	2.7e-05	intron	<i>Ten-a</i>	
X	15737065	C / G	1.9e-05	intron	<i>Aats-arg</i>	

TABLE S15. Top 100 genetic variants associated with micro-environmental plasticity of larval starvation resistance. Associated variants are listed with their chromosomal arm, location and type of polymorphism (major allele /

minor allele). Genic context refers to the SNP's location relative to genes. inter: intergenic, UTR: untranslated region, synonym: synonymous SNPs, non-synonym: non-synonymous SNPs. Distance of intergenic SNPs to nearby genes is given in parenthesis if within 1000 bp. All genes located within 1000 bp of the associated SNPs and their molecular functions are listed.



TABLE S16

DGRP line	Sample #1			Sample #2			Total		
	f	m	p value	f	m	p value	f	m	p value
306	49	50	1	51	50	1	100	100	1
852	76	34	7.7e-05 ***	61	37	0.02 *	137	71	5.5e-06 ***
502	67	31	3.5e-04 ***	68	31	2.6e-04 ***	135	62	2.1e-07 ***
338	52	43	0.41	61	36	0.01 *	113	79	0.02 *
392	55	40	0.15	47	50	0.84	102	90	0.43
379	47	45	0.92	58	40	0.09	105	85	0.17
57	37	34	0.81	51	52	1	88	86	0.94
555	61	26	2.2e-04 ***	59	35	0.02 *	120	61	1.4e-05 ***
787	52	47	0.69	34	65	2.4e-03 **	86	112	0.08
530	59	38	0.04 *	42	51	0.41	101	89	0.42

TABLE S16. Numbers of male and female larvae in the sexed dataset. Numbers are given including the larvae at the first count, which are omitted from analysis in calculation of starvation survival times. p values for binomial test under the assumption that the sex ratio is 1:1. f: female, m: male.

TABLE S17

TABLE S17.A

Stage of life		Females	Males	Total	p value
Egg		30	23	53	0.410
L3 larvae		65	38	103	0.010 *
Pupae		49	57	106	0.497
Adult	repl. #1	193	158	351	0.069
	repl. #2	195	198	393	0.920
	repl. #3	227	207	434	0.362
	repl. #4	194	213	407	0.372
	total	809	776	1585	0.422

TABLE S17.B

Groups of L1 larvae		Females	Males	Total	p value
to be included	repl. #1	43	29	72	0.125
	repl. #2	33	15	48	0.013 *
	total	76	44	120	0.004 **
to be excluded	repl. #1	34	36	70	0.905
	repl. #2	19	29	48	0.193
	total	53	65	118	0.311

TABLE S17. Two hypotheses for sex ratio distortion observed in RAL_502.

(A) Hypothesis 1: RAL_502 significantly produce more females than males. We monitored the numbers of both sexes during all developmental stages to find if there were consistent differences. (B) Hypothesis 2: We selected more females than males for the larval starvation assay. We compared the numbers of both sexes in two groups of larvae at 36 hours after egg-laying, one group containing larvae that we would select to include in the assay and the other group containing larvae that we would exclude. p values are given for the binomial test under the null hypothesis that sex ratio is 1.

TABLE S18

TABLE S18.A

Primer pair	Chromosome	PCR product length	Forward primer	Reverse primer
K	3R	539	TTCTCGGAGAGGAAGG CGTA	TATTCGATTCCCACACCG CA
L	3L	1413	AGTACGGCAAGTTCTG CATCGGAAA	AGAATTAGCTGTTGAG GGAATGGC
M	3L	1441	ATGGTGGTAAGCGTGG GTATGGAAG	GAGAATTAGCTGTTGAG GGGAATGGC
N	Y	1042	GATGGCTACGATGGAG GAGGGACTA	TGCCGGATGTCTATGGTT GCTGAAC
O	Y	820	TCCAGCAGAAGATAAG CGAGCAAGC	GGGCTTATCGCCAGGAGT TACAAGA
Q	Y	1233	TAACCCCTGAGCTACC TTCCAGGTC	TTGAGCATGTCCAAACCA TAAATCCCA
R	Y	1209	CCAGCCAGTCTCTCAG TGTTCTTCG	CTCTCGACCCACACCTTT GGCTTAG
S	Y	1093	ACAAACACTTACCACA GCGTCAGGT	TGTTGGAGGATCGGGAA AGCAAAGT
T	Y	743	AGTGCGTTATCCGACG AGCATACAC	TACAGCAAAAGAGTCGA GCAACCG
U	Y	821	AGCGTTATTTGATGCT GGGCGACTA	TACAGCAAAAGAGTCGA GCAACCG

TABLE S18.B

Line	Primer sets
338	KO, KS
392	LT, LO, LN
502	LU, MN
852	LO
530	KR, KQ
787	KQ
555	KQ, MO
306	KQ
57	KQ
379	KR

TABLE S18 (continued)

TABLE S18.C

Primer	Sequence	Product length (bp)
<i>Act5C</i> forward	CCCTCGTTCTTGGGAATGG	100
<i>Act5C</i> reverse	CGGTGTTGGCATACAGATCCT	
<i>CG42342</i> forward	AACAGCAAGCGTGGAAAAGG	137
<i>CG42342</i> reverse	TTCCTCGGTTCGCACTTGTC	

TABLE S18. List of all primers used in the project. (A) All primers used to determine the sex of dead larvae. (B) Specific primer sets used for each line. (C) Primers used in quantitative RT-PCR.

9. REFERENCES

- 1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. *Nature*, 467(7319), 1061–1073.
- Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., ... Venter, J. C. (2000). The genome sequence of *Drosophila melanogaster*. *Science*, 287(5461), 2185–2195.
- Aguila, J. R., Hoshizaki, D. K., & Gibbs, A. G. (2013). Contribution of larval nutrition to adult reproduction in *Drosophila melanogaster*. *Journal of Experimental Biology*, 216(3), 399–406.
- Aguila, J. R., Suszko, J., Gibbs, A. G., & Hoshizaki, D. K. (2007). The role of larval fat cells in adult *Drosophila melanogaster*. *Journal of Experimental Biology*, 210(6), 956–963.
- Aitman, T. J., Boone, C., Churchill, G. A., Hengartner, M. O., Mackay, T. F. C., & Stemple, D. L. (2011). The future of model organisms in human disease research. *Nature Reviews. Genetics*, 12(8), 575–582.
- Alpatov, W. W. (1930). Phenotypical variation in body and cell size of *Drosophila melanogaster*. *The Biological Bulletin*, 58(1), 85–103.
- Arellano-Carbajal, F., Briseño-Roa, L., Couto, A., Cheung, B. H. H., Labouesse, M., & de Bono, M. (2011). Macoilin, a conserved nervous system-specific ER membrane protein that regulates neuronal excitability. *PLoS Genetics*, 7(3), e1001341.
- Arquier, N., Géminard, C., Bourouis, M., Jarretou, G., Honegger, B., Paix, A., & Léopold, P. (2008). *Drosophila* ALS regulates growth and metabolism through functional interaction with insulin-like peptides. *Cell Metabolism*, 7(4), 333–338.
- Arya, G. H., Magwire, M. M., Huang, W., Serrano-Negron, Y. L., Mackay, T. F. C., & Anholt, R. R. H. (2015). The genetic basis for variation in olfactory behavior in *Drosophila melanogaster*. *Chemical Senses*, 40(4), 233–243.
- Arya, G. H., Weber, A. L., Wang, P., Magwire, M. M., Serrano-Negron, Y. L., Mackay, T. F. C., & Anholt, R. R. H. (2010). Natural variation, functional pleiotropy and transcriptional contexts of odorant binding protein genes in *Drosophila melanogaster*. *Genetics*, 186(4), 1475–1485.
- Astle, W., & Balding, D. (2009). Population structure and cryptic relatedness in genetic association studies. *Statistical Science*, 24(4), 451–471.
- Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., ... Nordborg, M. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, 465(7298), 627–631.
- Aulchenko, Y. S., Ripatti, S., Lindqvist, I., Boomsma, D., Heid, I. M., Pramstaller, P. P., ... Peltonen, L. (2009). Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nature Genetics*, 41(1), 47–

- Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., ... Schloss, J. A. (2015). A global reference for human genetic variation. *Nature*, 526(7571), 68–74.
- Ayroles, J. F., Carbone, M. A., Stone, E. A., Jordan, K. W., Lyman, R. F., Magwire, M. M., ... Mackay, T. F. C. (2009). Systems genetics of complex traits in *Drosophila melanogaster*. *Nature Genetics*, 41(3), 299–307.
- Baggott, G. K. (1975). Molt, flight muscle “hypertrophy” and premigratory lipid deposition of the juvenile willow warbler, *Phylloscopus trochilus*. *Journal of Zoology, London*, 175, 299–314.
- Bai, H., Kang, P., Hernandez, A. M., & Tatar, M. (2013). Activin signaling targeted by insulin/dFOXO regulates aging and muscle proteostasis in *Drosophila*. *PLoS Genetics*, 9(11), e1003941.
- Bakker, K. (1959). Feeding period, growth, and pupation in larvae of *Drosophila melanogaster*. *Entomologia Experimentalis et Applicata*, 2(3), 171–186.
- Baldal, E. A., Brakefield, P. M., & Zwaan, B. J. (2006). Multitrait evolution in lines of *Drosophila melanogaster* selected for increased starvation resistance: the role of metabolic rate and implications for the evolution of longevity. *Evolution*, 60(7), 1435–1444.
- Ballard, J. W. O., Melvin, R. G., & Simpson, S. J. (2008). Starvation resistance is positively correlated with body lipid proportion in five wild caught *Drosophila simulans* populations. *Journal of Insect Physiology*, 54(9), 1371–1376.
- Bassett, A. R., Tibbit, C., Ponting, C. P., & Liu, J.-L. (2013). Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Reports*, 4(1), 220–228.
- Beadle, G. W., Tatum, E. L., & Clancy, C. W. (1938). Food level in relation to rate of development and eye pigmentation in *Drosophila Melanogaster*. *The Biological Bulletin*, 75(3), 447–462.
- Bellen, H. J., Vaessin, H., Bier, E., Kolodkin, A., D'Evelyn, D., Kooyer, S., & Jan, Y. N. (1992a). The *Drosophila couch potato* gene: An essential gene required for normal adult behavior. *Genetics*, 131(2), 365–375.
- Bellen, H. J., Kooyer, S., D'Evelyn, D., & Pearlman, J. (1992b). The *Drosophila couch potato* protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. *Genes and Development*, 6(11), 2125–2136.
- Bellen, H. J., Levis, R. W., Liao, G., He, Y., Carlson, J. W., Tsang, G., ... Spradling, A. C. (2004). The BDGP gene disruption project: Single transposon insertions associated with 40% of *Drosophila* genes. *Genetics*, 167(2), 761–781.
- Bennett, B. J., Farber, C. R., Orozco, L., Kang, H. M., Ghazalpour, A., Siemers, N., ... Lusis, A. J. (2010). A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Research*, 20(2), 281–290.
- Berger, K. H., Kong, E. C., Dubnau, J., Tully, T., Moore, M. S., & Heberlein, U. (2008).

- Ethanol sensitivity and tolerance in long-term memory mutants of *Drosophila melanogaster*. *Alcoholism: Clinical and Experimental Research*, 32(5), 895–908.
- Bertile, F., Oudart, H., Criscuolo, F., Maho, Y. L., & Raclot, T. (2003). Hypothalamic gene expression in long-term fasted rats: Relationship with body fat. *Biochemical and Biophysical Research Communications*, 303(4), 1106–1113.
- Bharucha, K. N. (2009). The epicurean fly: Using *Drosophila melanogaster* to study metabolism. *Pediatric Research*, 65(2), 132–137.
- Bicego, K. C., Barros, R. C. H., & Branco, L. G. S. (2007). Physiology of temperature regulation: Comparative aspects. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 147(3), 616–639.
- Bodmer, W., & Bonilla, C. (2008). Common and rare variants in multifactorial susceptibility to common diseases. *Nature Genetics*, 40(6), 695–701.
- Boétius, I., & Boétius, J. (1985). Lipid and protein content in *Anguilla anguilla* during growth and starvation. *Dana*, 4, 1–17.
- Borash, D. J., Teotónio, H., Rose, M. R., & Mueller, L. D. (2000). Density-dependent natural selection in *Drosophila*: Correlations between feeding rate, development time and viability. *Journal of Evolutionary Biology*, 13(2), 181–187.
- Borash, D. J., & Ho, G. T. (2001). Patterns of selection: Stress resistance and energy storage in density-dependent populations of *Drosophila melanogaster*. *Journal of Insect Physiology*, 47(12), 1349–1356.
- Bowser, P. R. F., & Tobe, S. S. (2005). Immunocytochemical analysis of putative allatostatin receptor (DAR-2) distribution in the CNS of larval *Drosophila melanogaster*. *Peptides*, 26(1), 81–87.
- Bradfield, J. P., Qu, H. Q., Wang, K., Zhang, H., Sleiman, P. M., Kim, C. E., ... Hakonarson, H. (2011). A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genetics*, 7(9), e1002293.
- Bradley, T. J., & Simmons, F. H. (1997). An analysis of resource allocation in response to dietary yeast in *Drosophila melanogaster*. *Journal of Insect Physiology*, 43(8), 779–788.
- Britton, J. S., & Edgar, B. A. (1998). Environmental control of the cell cycle in *Drosophila*: nutrition activates mitotic and endoreplicative cells by distinct mechanisms. *Development*, 125(11), 2149–2158.
- Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., & Hafen, E. (2001). An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Current Biology*, 11(4), 213–221.
- Broman, K. W. (2005). The genomes of recombinant inbred lines. *Genetics*, 169(2), 1133–1146.
- Broughton, S. J., Piper, M. D. W., Ikeya, T., Bass, T. M., Jacobson, J., Drieger, Y., ... Partridge, L. (2005). Longer lifespan, altered metabolism, and stress resistance in *Drosophila* from ablation of cells making insulin-like ligands. *Proceedings of the National Academy of Sciences of the United States of America*, 102(12), 5078–5083.

- America*, 102(8), 3105–3110.
- Bubliy, O. A., & Loeschcke, V. (2005). Correlated responses to selection for stress resistance and longevity in a laboratory population of *Drosophila melanogaster*. *Journal of Evolutionary Biology*, 18(4), 789–803.
- Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLoS Computational Biology*, 8(12), e1002822.
- Cahill, G. F. (1976). Starvation in man. *Clinics in Endocrinology and Metabolism*, 5(2), 397–415.
- Cahill, G. F. (2006). Fuel metabolism in starvation. *Annual Review of Nutrition*, 26, 1–22.
- Carbone, M. A., Jordan, K. W., Lyman, R. F., Harbison, S. T., Leips, J., Morgan, T. J., ... Mackay, T. F. C. (2006). Phenotypic variation and natural selection at *catsup*, a pleiotropic quantitative trait gene in *Drosophila*. *Curr.Biol.*, 16(9), 912–919.
- Carvalho, G. B., Kapahi, P., & Benzer, S. (2005). Compensatory ingestion upon dietary restriction in *Drosophila melanogaster*. *Nature Methods*, 2(11), 813–815.
- Carvalho, G. B., Kapahi, P., Anderson, D. J., & Benzer, S. (2006). Allocrine modulation of feeding behavior by the sex peptide of *Drosophila*. *Current Biology*, 16(7), 692–696.
- Chen, D., Qu, C., & Hewes, R. S. (2014). Neuronal remodeling during metamorphosis is regulated by the *alan shepard (shep)* gene in *Drosophila melanogaster*. *Genetics*, 197(4), 1267–1283.
- Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. *Evolution*, 42(5), 958–968.
- Chippindale, A., Leroi, A., Kim, S., & Rose, M. (1993). Phenotypic plasticity and selection in *Drosophila* life-history evolution. I. Nutrition and the cost of reproduction. *Journal of Evolutionary ...*, 6, 171–193.
- Chippindale, A. K., Chu, T. J. F., & Rose, M. R. (1996). Complex trade-offs and the evolution of starvation resistance in *Drosophila melanogaster*. *Evolution*, 50(2), 753–766.
- Chow, C. Y., Wolfner, M. F., & Clark, A. G. (2013). Using natural variation in *Drosophila* to discover previously unknown endoplasmic reticulum stress genes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(22), 9013–9018.
- Church, R. B., & Robertson, F. W. (1966). A biochemical study of the growth of *Drosophila melanogaster*. *Journal of Experimental Zoology*, 162(3), 337–351.
- Churchill, G. A., Airey, D. C., Allayee, H., Angel, J. M., Attie, A. D., Beatty, J., ... Zou, F. (2004). The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genetics*, 36(11), 1133–1137.
- Civelek, M., & Lusi, A. J. (2014). Systems genetics approaches to understand complex traits. *Nature Reviews. Genetics*, 15(1), 34–48.
- Clancy, D. J., Gems, D., Harshman, L. G., Oldham, S., Stocker, H., Hafen, E., ... Partridge, L. (2001). Extension of life-span by loss of CHICO, a *Drosophila*

- insulin receptor substrate protein. *Science*, 292(5514), 104–106.
- Clayton, D. G., Walker, N. M., Smyth, D. J., Pask, R., Cooper, J. D., Maier, L. M., ... Todd, J. A. (2005). Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nature Genetics*, 37(11), 1243–1246.
- Collaborative Cross Consortium. (2012). The genome architecture of the Collaborative Cross mouse genetic reference population. *Genetics*, 190(2), 389–401.
- Colombani, J., Raisin, S., Pantalacci, S., Radimerski, T., Montagne, J., & Léopold, P. (2003). A nutrient sensor mechanism controls *Drosophila* growth. *Cell*, 114(6), 739–749.
- Cooper, R. S., Tayo, B., & Zhu, X. (2008). Genome-wide association studies: Implications for multiethnic samples. *Human Molecular Genetics*, 17(R2), R151–R155.
- Corbett-Detig, R. B., & Hartl, D. L. (2012). Population genomics of inversion polymorphisms in *Drosophila melanogaster*. *PLoS Genetics*, 8(12), e1003056.
- Curtiss, J., & Heilig, J. S. (1995). Establishment of *Drosophila* imaginal precursor cells is controlled by the *Arrowhead* gene. *Development*, 121(11), 3819–3828.
- Da Lage, J.-L., Capy, P., & David, J. R. (1989). Starvation and desiccation tolerance in *Drosophila melanogaster* adults: Effects of environmental temperature. *Journal of Insect Physiology*, 35(6), 453–457.
- Da Lage, J.-L., Capy, P., & David, J. R. (1990). Starvation and desiccation tolerance in *Drosophila melanogaster*: Differences between European, North African and Afrotropical populations. *Genetics Selection Evolution*, 22(4), 381–391.
- Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J., & Lander, E. S. (2001). High-resolution haplotype structure in the human genome. *Nature Genetics*, 29(2), 229–232.
- Dawn Teare, M., & Barrett, J. H. (2005). Genetic linkage studies. *The Lancet*, 366(9490), 1036–1044.
- de Bakker, P. I. W., Yelensky, R., Pe'er, I., Gabriel, S. B., Daly, M. J., & Altshuler, D. (2005). Efficiency and power in genetic association studies. *Nat Genet*, 37(11), 1217–1223.
- de Bakker, P. I. W., Ferreira, M. A. R., Jia, X., Neale, B. M., Raychaudhuri, S., & Voight, B. F. (2008). Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Human Molecular Genetics*, 17(R2), R122–R128.
- Dembeck, L. M., Huang, W., Magwire, M. M., Lawrence, F., Lyman, R. F., & Mackay, T. F. C. (2015). Genetic architecture of abdominal pigmentation in *Drosophila melanogaster*. *PLoS Genetics*, 11(5), e1005163.
- Dickson, S. P., Wang, K., Krantz, I., Hakonarson, H., & Goldstein, D. B. (2010). Rare variants create synthetic genome-wide associations. *PLoS Biology*, 8(1), e1000294.

- Djawdan, M., Rose, M. R., & Bradley, T. J. (1997). Does selection for stress resistance lower metabolic rate? *Ecology*, 78(3), 828–837.
- Djawdan, M., Chippindale, A. K., Rose, M. R., & Bradley, T. J. (1998). Metabolic reserves and evolved stress resistance in *Drosophila melanogaster*. *Physiological Zoology*, 71(5), 584–594.
- Domené, H. M., Bengolea, S. V., Jasper, H. G., & Boisclair, Y. R. (2005). Acid-labile subunit deficiency: Phenotypic similarities and differences between human and mouse. *Journal of Endocrinological Investigation*, 28(5), 43–46.
- Dunlap, K. D. (1995). Hormonal and behavioral responses to food and water deprivation in a lizard (*Sceloporus occidentalis*): Implications for assessing stress in a natural population. *Journal of Herpetology*, 29, 345–351.
- Durham, M. F., Magwire, M. M., Stone, E. A., & Leips, J. (2014). Genome-wide analysis in *Drosophila* reveals age-specific effects of SNPs on fitness traits. *Nature Communications*, 5, 1–8.
- Edgar, B. A. (2006). How flies get their size: Genetics meets physiology. *Nature Reviews. Genetics*, 7(12), 907–916.
- Ehret, G. B., Munroe, P. B., Rice, K. M., Bochud, M., Johnson, A. D., Chasman, D. I., ... Johnson, T. (2011). Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*, 478(7367), 103–109.
- Elia, M., Stubbs, R. J., & Henry, C. J. (1999). Differences in fat, carbohydrate, and protein metabolism between lean and obese subjects undergoing total starvation. *Obesity Research*, 7(6), 597–604.
- Ellis, L. L., Huang, W., Quinn, A. M., Ahuja, A., Alfrejd, B., Gomez, F. E., ... Tarone, A. M. (2014). Intrapopulation genome size variation in *D. melanogaster* reflects life history variation and plasticity. *PLoS Genetics*, 10(7), e1004522.
- Elstob, P. R., Brodu, V., & Gould, A. P. (2001). *spalt*-dependent switching between two cell fates that are induced by the *Drosophila* EGF receptor. *Development*, 128(5), 723–732.
- Eu-ahsunthornwattana, J., Miller, E. N., Fakiola, M., Jeronimo, S. M. B., Blackwell, J. M., & Cordell, H. J. (2014). Comparison of methods to account for relatedness in genome-wide association studies with family-based data. *PLoS Genetics*, 10(7), e1004445.
- Ewald, C. Y., Landis, J. N., Porter Abate, J., Murphy, C. T., & Blackwell, T. K. (2015). Dauer-independent insulin/IGF-1-signalling implicates collagen remodelling in longevity. *Nature*, 519(7541), 97–101.
- Fang, P., Yu, M., Guo, L., Bo, P., Zhang, Z., & Shi, M. (2012). Galanin and its receptors: A novel strategy for appetite control and obesity therapy. *Peptides*, 36(2), 331–339.
- Farhadian, S. F., Suárez-Fariñas, M., Cho, C. E., Pellegrino, M., & Vossall, L. B. (2012). Post-fasting olfactory, transcriptional, and feeding responses in *Drosophila*. *Physiology and Behavior*, 105(2), 544–553.
- Flint, J., & Mackay, T. F. C. (2009). Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Research*, 19, 723–733.

- Flint, J., & Eskin, E. (2012). Genome-wide association studies in mice. *Nature Reviews. Genetics*, 13(11), 807–817.
- Fox, C. S., Liu, Y., White, C. C., Feitosa, M., Smith, A. V., Heard-Costa, N., ... Borecki, I. B. (2012). Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genetics*, 8(5), e1002695.
- Fry, A. J., Palmer, M. R., & Rand, D. M. (2004). Variable fitness effects of *Wolbachia* infection in *Drosophila melanogaster*. *Heredity*, 93(4), 379–89.
- Fujikawa, K., Takahashi, A., Nishimura, A., Itoh, M., Takano-Shimizu, T., & Ozaki, M. (2009). Characteristics of genes up-regulated and down-regulated after 24 h starvation in the head of *Drosophila*. *Gene*, 446(1), 11–17.
- Gaertner, B. E., Parmenter, M. D., Rockman, M. V., Kruglyak, L., & Phillips, P. C. (2012). More than the sum of its parts: A complex epistatic network underlies natural variation in thermal preference behavior in *Caenorhabditis elegans*. *Genetics*, 192(4), 1533–1542.
- Garlapow, M. E., Huang, W., Yarboro, M. T., Peterson, K. R., & Mackay, T. F. C. (2015). Quantitative genetics of food intake in *Drosophila melanogaster*. *PloS ONE*, 10(9), e0138129.
- Geer, B. W., Langevin, M. L., & McKechnie, S. W. (1985). Dietary ethanol and lipid synthesis in *Drosophila melanogaster*. *Biochemical Genetics*, 23(7–8), 607–622.
- Geiser, F. (2004). Metabolic rate and body temperature reduction during hibernation and daily torpor. *Annual Review of Physiology*, 66, 239–274.
- Géminard, C., Rulifson, E. J., & Léopold, P. (2009). Remote control of insulin secretion by fat cells in *Drosophila*. *Cell Metabolism*, 10(3), 199–207.
- Gibbs, A. G., & Reynolds, L. A. (2012). *Drosophila* as a model for starvation: Evolution, physiology, and genetics. In M. D. McCue (Ed.), *Comparative Physiology of Fasting, Starvation, and Food Limitation*, Springer, 37–51.
- Gibson, G. (2012). Rare and common variants: twenty arguments. *Nature Reviews Genetics*, 13(2), 135–145.
- Graff, M., Ngwa, J. S., Workalemahu, T., Homuth, G., Schipf, S., Teumer, A., ... Berndt, S. I. (2013). Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course. *Human Molecular Genetics*, 22(17), 3597–3607.
- Gregg, T. G., McCrate, A., Reveal, G., Hall, S., & Rypstra, A. L. (1990). Insectivory and social digestion in *Drosophila*. *Biochemical Genetics*, 28(3–4), 197–207.
- Grewal, S. S., Li, L., Orian, A., Eisenman, R. N., & Edgar, B. A. (2005). Myc-dependent regulation of ribosomal RNA synthesis during *Drosophila* development. *Nature Cell Biology*, 7(3), 295–302.
- Grey, N. J., Karl, I., & Kipnis, D. M. (1975). Physiologic mechanisms in the development of starvation ketosis in man. *Diabetes*, 24(1), 10–16.
- Groenke, S., Mildner, A., Fellert, S., Tennagels, N., Petry, S., Mueller, G., ... Kuehnlein, R. P. (2005). Brummer lipase is an evolutionary conserved fat

- storage regulator in *Drosophila*. *Cell Metabolism*, 1(5), 323–330.
- Grosjean, Y., Rytz, R., Farine, J., Abuin, L., Cortot, J., Jefferis, G. S. X. E., & Benton, R. (2011). An olfactory receptor for food-derived odours promotes male courtship in *Drosophila*. *Nature*, 478(7368), 236–240.
- Guryev, V., Smits, B. M. G., Van De Belt, J., Verheul, M., Hubner, N., & Cuppen, E. (2006). Haplotype block structure is conserved across mammals. *PLoS Genetics*, 2(7), e121.
- Gutierrez, E., Wiggins, D., Fielding, B., & Gould, A. P. (2007). Specialized hepatocyte-like cells regulate *Drosophila* lipid metabolism. *Nature*, 445(7125), 275–280.
- Habold, C., Chevalier, C., Dunel-Erb, S., Foltzer-Jourdainne, C., Le Maho, Y., & Lignot, J.-H. (2004). Effects of fasting and refeeding on jejunal morphology and cellular activity in rats in relation to depletion of body stores. *Scandinavian Journal of Gastroenterology*, 39(6), 531–539.
- Hainsworth, F. R., & Wolf, L. L. (1970). Regulation of oxygen consumption and body temperature during torpor in a hummingbird, *Eulampis jugularis*. *Science*, 168(929), 368–369.
- Hampe, J., Franke, A., Rosenstiel, P., Till, A., Teuber, M., Huse, K., ... Schreiber, S. (2007). A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in *ATG16L1*. *Nature Genetics*, 39(2), 207–211.
- Hand, S. C., & Hardewig, I. (1996). Downregulation of cellular metabolism during environmental stress: mechanisms and implications. *Annual Review of Physiology*, 58, 539–563.
- Handke, B., Poernbacher, I., Goetze, S., Ahrens, C. H., Omasits, U., Marty, F., ... Lehner, C. F. (2013). The hemolymph proteome of fed and starved *Drosophila* larvae. *PLoS ONE*, 8(6), e67208.
- Hanson, R. W., & Owen, O. E. (2013). Gluconeogenesis. In Lennarz, W.J., & Lane, M.D. (Ed.), *Encyclopedia of Biological Chemistry*, 381–386.
- Harbison, S. T., Yamamoto, A. H., Fanara, J. J., Norga, K. K., & Mackay, T. F. C. (2004). Quantitative trait loci affecting starvation resistance in *Drosophila melanogaster*. *Genetics*, 166(4), 1807–1823.
- Harbison, S. T., Chang, S., Kamdar, K. P., & Mackay, T. F. C. (2005). Quantitative genomics of starvation stress resistance in *Drosophila*. *Genome Biology*, 6(4), R36.
- Harbison, S. T., McCoy, L. J., & Mackay, T. F. C. (2013). Genome-wide association study of sleep in *Drosophila melanogaster*. *BMC Genomics*, 14, 281.
- Harshman, L. G., & Schmid, J. L. (1998). Evolution of starvation resistance in *Drosophila melanogaster*: Aspects of metabolism and counter-impact selection. *Evolution*, 52(6), 1679–1685.
- Harshman, L. G., Hoffmann, A. A., & Clark, A. G. (1999a). Selection for starvation resistance in *Drosophila melanogaster*: Physiological correlates, enzyme activities and multiple stress responses. *Journal of Evolutionary Biology*,

- 12(2), 370–379.
- Harshman, L. G., Moore, K. M., Sty, M. A., & Magwire, M. M. (1999b). Stress resistance and longevity in selected lines of *Drosophila melanogaster*. *Neurobiology of Aging*, 20(5), 521–529.
- Harshman, L. G., & Hoffmann, A. A. (2000). Laboratory selection experiments using *Drosophila*: What do they really tell us? *Trends in Ecology and Evolution*, 15(1), 32–36.
- Hentze, J. L., Carlsson, M. A., Kondo, S., Nässel, D. R., & Rewitz, K. F. (2015). The neuropeptide Allatostatin A regulates metabolism and feeding decisions in *Drosophila*. *Scientific Reports*, 5, 11680.
- Hergarden, A. C., Tayler, T. D., & Anderson, D. J. (2012). Allatostatin-A neurons inhibit feeding behavior in adult *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 109(10), 3967–3972.
- Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, 6(2), 95–108.
- Hoang, A. (2001). Immune response to parasitism reduces resistance of *Drosophila melanogaster* to desiccation and starvation. *Evolution*, 55(11), 2353–2358.
- Hoffmann, A. A., & Harshman, L. G. (1999). Desiccation and starvation resistance in *Drosophila*: Patterns of variation at the species, population and intrapopulation levels. *Heredity* 83(6), 637–643.
- Hoffmann, A. A., Hallas, R., Sinclair, C., Partridge, L., & Mitrovski, P. (2001). Levels of variation in stress resistance in *Drosophila* among strains, local populations, and geographic regions: Patterns for desiccation, starvation, cold resistance, and associated traits. *Evolution*, 55(8), 1621–1630.
- Hoffmann, A. A., Hallas, R., Anderson, A. R., & Telonis-Scott, M. (2005). Evidence for a robust sex-specific trade-off between cold resistance and starvation resistance in *Drosophila melanogaster*. *Journal of Evolutionary Biology*, 18(4), 804–810.
- Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, 130(1), 195–204.
- Houtkooper, R. H., Pirinen, E., & Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. *Nature Reviews Molecular Cell Biology*, 13(4), 225–238.
- Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., & Mohr, S. E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics*, 12, 357.
- Huang, F.-D., Woodruff, E., Mohrmann, R., & Broadie, K. (2006). Rolling blackout is required for synaptic vesicle exocytosis. *The Journal of Neuroscience*, 26(9), 2369–2379.
- Huang, D. W., Lempicki, R. A., & Sherman, B. T. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), 44–57.

- Huang, W., Richards, S., Carbone, M. A., Zhu, D., Anholt, R. R. H., Ayroles, J. F., ... Mackay, T. F. C. (2012). Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. *Proceedings of the National Academy of Sciences of the United States of America*, 109(39), 15553–15559.
- Huang, W., Massouras, A., Inoue, Y., Peiffer, J., Ramia, M., Tarone, A. M., ... Mackay, T. F. C. (2014). Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Research*, 24(7), 1193–1208.
- Huey, R. B., Suess, J., Hamilton, H., & Gilchrist, G. W. (2004). Starvation resistance in *Drosophila melanogaster*: Testing for a possible “cannibalism” bias. *Functional Ecology*, 18(6), 952–954.
- Humphries, M. M., Thomas, D. W., & Kramer, D. L. (2003). The role of energy availability in mammalian hibernation: a cost-benefit approach. *Physiological and Biochemical Zoology*, 76(2), 165–179.
- Hung, H.-C., Kay, S. a, & Weber, F. (2009). HSP90, a capacitor of behavioral variation. *Journal of Biological Rhythms*, 24, 183–192.
- Ioannidis, J. P. A., Thomas, G., & Daly, M. J. (2009). Validating, augmenting and refining genome-wide association signals. *Nature Reviews. Genetics*, 10(5), 318–329.
- Irvin, M. R., Wineinger, N. E., Rice, T. K., Pajewski, N. M., Kabagambe, E. K., Gu, C. C., ... Arnett, D. K. (2011). Genome-wide detection of allele specific copy number variation associated with insulin resistance in African Americans from the HyperGEN study. *PLoS ONE*, 6(8), e24052.
- Ivanov, D. K., Escott-Price, V., Ziehm, M., Magwire, M. M., Mackay, T. F. C., Partridge, L., & Thornton, J. M. (2015). Longevity GWAS using the *Drosophila* Genetic Reference Panel. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 70(12), 1470-1478.
- Jones, J. E., & Wade, G. N. (2002). Acute fasting decreases sexual receptivity and neural estrogen receptor-alpha in female rats. *Physiology and Behavior*, 77(1), 19–25.
- Jordan, K. W., Craver, K. L., Magwire, M. M., Cubilla, C. E., Mackay, T. F. C., & Anholt, R. R. H. (2012). Genome-wide association for sensitivity to chronic oxidative stress in *Drosophila melanogaster*. *PLoS ONE*, 7(6), e38722.
- Jumbo-Lucioni, P., Ayroles, J. F., Chambers, M. M., Jordan, K. W., Leips, J., Mackay, T. F. C., & De Luca, M. (2010). Systems genetics analysis of body weight and energy metabolism traits in *Drosophila melanogaster*. *BMC Genomics*, 11, 297.
- Kadler, K. E., Baldock, C., Bella, J., & Boot-Handford, R. P. (2007). Collagens at a glance. *Journal of Cell Science*, 120(12), 1955–1958.
- Keller, T. E., Wilke, C. O., & Bull, J. J. (2012). Interactions between evolutionary processes at high mutation rates. *Evolution*, 66(7), 2303–2314.
- Kerndt, P. R., Naughton, J. L., Driscoll, C. E., & Loxterkamp, D. A. (1982). Fasting: the history, pathophysiology and complications. *The Western Journal of*

- Medicine*, 137(5), 379–399.
- Ketterson, E. D., & King, J. R. (1977). Metabolic and behavioral responses to fasting in the white-crowned sparrow *Zonotrichia leucophrys gambelii*. *Physiological Zoology*, 50(2), 115–129.
- King, E. G., Macdonald, S. J., & Long, A. D. (2012). Properties and power of the *Drosophila* Synthetic Population Resource for the routine dissection of complex traits. *Genetics*, 191(3), 935–949.
- Kitsios, G. D., & Zintzaras, E. (2009). Genome-wide association studies: hypothesis-“free” or “engaged”? *Translational Research*, 154(4), 161–164.
- Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J.-Y., Sackler, R. S., Haynes, C., ... Hoh, J. (2005). Complement factor H polymorphism in age-related macular degeneration. *Science*, 308(5720), 385–389.
- Knoppien, P., Van Der Pers, J. N. C., & Van Delden, W. (2000). Quantification of locomotion and the effect of food deprivation on locomotor activity in *Drosophila*. *Journal of Insect Behavior*, 13(1), 27–43.
- Kochi, Y., Suzuki, A., Yamada, R., & Yamamoto, K. (2009). Genetics of rheumatoid arthritis: Underlying evidence of ethnic differences. *Journal of Autoimmunity*, 32(3-4), 158-162.
- Kolss, M., Vijendravarma, R. K., Schwaller, G., & Kawecki, T. J. (2009). Life-history consequences of adaptation to larval nutritional stress in *Drosophila*. *Evolution*, 63(9), 2389–2401.
- Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9(1), 29.
- Koubi, H. E., Robin, J. P., Dewasmes, G., Le Maho, Y., Frutoso, J., & Minaire, Y. (1991). Fasting-induced rise in locomotor activity in rats coincides with increased protein utilization. *Physiology and Behavior*, 50(2), 337–343.
- Kramer, J. M., Slade, J. D., & Staveley, B. E. (2008). *foxo* is required for resistance to amino acid starvation in *Drosophila*. *Genome*, 51(8), 668–672.
- Kreitzman, S. N., Coxon, A. Y., & Szaz, K. F. (1992). Glycogen storage: Illusions of easy weight loss, excessive weight regain, and distortions in estimates of body composition. *American Journal of Clinical Nutrition*, 56(1), 292S-293S.
- Kumar, M., Joseph, M., & Chandrashekar, S. (2001). Effects of mutations at the *stambh A* locus of *Drosophila melanogaster*. *Journal of Genetics*, 80(2), 83–95.
- Langley, C. H., Stevens, K., Cardeno, C., Lee, Y. C. G., Schrider, D. R., Pool, J. E., ... Begun, D. J. (2012). Genomic variation in natural populations of *Drosophila melanogaster*. *Genetics*, 192(2), 533–598.
- Lebreton, S., Trona, F., Borrero-Echeverry, F., Bilz, F., Grabe, V., Becher, P. G., ... Witzgall, P. (2015). Feeding regulates sex pheromone attraction and courtship in *Drosophila* females. *Scientific Reports*, 5, 13132.
- Lee, G., & Park, J. H. (2004). Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in *Drosophila melanogaster*. *Genetics*, 167(1), 311–323.

- Lee, K. S., You, K. H., Choo, J. K., Han, Y. M., & Yu, K. (2004). *Drosophila* short neuropeptide F regulates food intake and body size. *Journal of Biological Chemistry*, 279(49), 50781–50789.
- Lee, J. C., & Parkes, M. (2011). Genome-wide association studies and Crohn's disease. *Briefings in Functional Genomics*, 10(2), 71–76.
- Lee, K. P., & Jang, T. (2014). Exploring the nutritional basis of starvation resistance in *Drosophila melanogaster*. *Functional Ecology*, 28(5), 1144–1155.
- Leitinger, B. (2011). Transmembrane collagen receptors. *Annu Rev Cell Dev Biol*, 27, 265–290.
- Lemaitre, B., & Miguel-Aliaga, I. (2013). The digestive tract of *Drosophila melanogaster*. *Annual Review of Genetics*, 47, 377–404.
- Lendvai, A. Z., Barta, Z., Liker, A., & Bókonyi, V. (2004). The effect of energy reserves on social foraging: hungry sparrows scrounge more. *Proceedings of the Royal Society B, Biological Sciences*, 271(1556), 2467–2472.
- Leopold, P., & Perrimon, N. (2007). *Drosophila* and the genetics of the internal milieu. *Nature*, 450, 186–188.
- Li, L., Edgar, B. A., & Grewal, S. S. (2010). Nutritional control of gene expression in *Drosophila* larvae via TOR, Myc and a novel cis-regulatory element. *BMC Cell Biology*, 11(1), 7.
- Lin, Y.-J., Seroude, L., & Benzer, S. (1998). Extended life-span and stress resistance in the *Drosophila* mutant *methuselah*. *Science*, 282(5390), 943–946.
- Lindstedt, S. L., & Boyce, M. S. (1985). Seasonality, fasting endurance, and body size in mammals. *The American Naturalist*, 125(6), 873–878.
- Lippert, C., Listgarten, J., Liu, Y., Kadie, C. M., Davidson, R. I., & Heckerman, D. (2011). FaST linear mixed models for genome-wide association studies. *Nature Methods*, 8(10), 833–837.
- Liu, S., Lamaze, A., Liu, Q., Tabuchi, M., Yang, Y., Fowler, M., ... Wu, M. N. (2014). Wide Awake mediates the circadian timing of sleep onset. *Neuron*, 82(1), 151–166.
- Lv, D., Zhang, D. D., Wang, H., Zhang, Y., Liang, L., Fu, J. F., ... Zhu, Y. M. (2015). Genetic variations in *SEC16B*, *MC4R*, *MAP2K5* and *KCTD15* were associated with childhood obesity and interacted with dietary behaviors in Chinese school-age population. *Gene*, 560(2), 149–155.
- Mackay, T. F. C., & Anholt, R. R. H. (2006). Of flies and man: *Drosophila* as a model for human complex traits. *Annual Review of Genomics and Human Genetics*, 7, 339–367.
- Mackay, T. F. C., Stone, E. A., & Ayroles, J. F. (2009). The genetics of quantitative traits: challenges and prospects. *Nature Reviews. Genetics*, 10(8), 565–577.
- Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., ... Gibbs, R. A. (2012). The *Drosophila melanogaster* Genetic Reference Panel. *Nature*, 482(7384), 173–178.

- Mackay, T. F. C. (2014). Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nature Reviews. Genetics*, 15(1), 22–33.
- Magwire, M. M., Yamamoto, A., Carbone, M. A., Roshina, N. V., Symonenko, A. V., Pasyukova, E. G., ... Mackay, T. F. C. (2010). Quantitative and molecular genetic analyses of mutations increasing *Drosophila* life span. *PLoS Genetics*, 6(7), 1–21.
- Manolio, T. A., Brooks, L. D., & Collins, F. S. (2008). A HapMap harvest of insights into the genetics of common disease. *Journal of Clinical Investigation*, 118(5), 1590–1605.
- Marchini, J., & Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nature Reviews. Genetics*, 11(7), 499–511.
- Masek, P., Reynolds, L. A., Bollinger, W. L., Moody, C., Mehta, A., Murakami, K., ... Keene, A. C. (2014). Altered regulation of sleep and feeding contribute to starvation resistance in *Drosophila*. *The Journal of Experimental Biology*, 217(17), 3122–3132.
- Matzkin, L. M., Watts, T. D., & Markow, T. A. (2009). Evolution of stress resistance in *Drosophila*: interspecific variation in tolerance to desiccation and starvation. *Functional Ecology*, 23(3), 521–527.
- McCarthy, M. I., & Hirschhorn, J. N. (2008). Genome-wide association studies: Potential next steps on a genetic journey. *Human Molecular Genetics*, 17(R2), R156–R165.
- McCue, M. D. (2007). Snakes survive starvation by employing supply- and demand-side economic strategies. *Zoology*, 110(4), 318–327.
- McCue, M. D. (2010). Starvation physiology: reviewing the different strategies animals use to survive a common challenge. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology*, 156(1), 1–18.
- McDonald, D. M., & Keene, A. C. (2010). The sleep-feeding conflict: Understanding behavioral integration through genetic analysis in *Drosophila*. *Aging*, 2(8), 519–522.
- McPherson, R., Pertsemlidis, A., Kavaslar, N., Stewart, A., Roberts, R., Cox, D., ... Cohen, J. (2007). A common allele on chromosome 9 associated with coronary heart disease. *Science*, 316(5830), 1488–1491.
- Medici, V., Vonesch, S. C., Fry, S. N., & Hafen, E. (2015). The FlyCatwalk: a high-throughput feature-based sorting system for artificial selection in *Drosophila*. *G3: Genes, Genomes, Genetics*, 5(3), 317–327.
- Melcher, C., & Pankratz, M. J. (2005). Candidate gustatory interneurons modulating feeding behavior in the *Drosophila* brain. *PLoS Biology*, 3(9), e305.
- Melcher, C., Bader, R., Walther, S., Simakov, O., & Pankratz, M. J. (2006). Neuromedin U and its putative *Drosophila* homolog hugin. *PLoS Biology*, 4(3), e68.
- Merila, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: Empirical patterns and development of ideas. *Heredity*, 83(2), 103–

- Merkey, A. B., Wong, C. K., Hoshizaki, D. K., & Gibbs, A. G. (2011). Energetics of metamorphosis in *Drosophila melanogaster*. *Journal of Insect Physiology*, 57(10), 1437–1445.
- Metaxakis, A., Oehler, S., Klinakis, A., & Savakis, C. (2005). Minos as a genetic and genomic tool in *Drosophila melanogaster*. *Genetics*, 171(2), 571–581.
- Millar, J., & Hickling, G. (1990). Fasting endurance and the evolution of mammalian body size. *Functional Ecology*, 4(1), 5–12.
- Mirth, C. K., & Shingleton, A. W. (2012). Integrating body and organ size in *Drosophila*: Recent advances and outstanding problems. *Frontiers in Endocrinology*, 3, 49.
- Miyara, A., Ohta, A., Okochi, Y., Tsukada, Y., Kuhara, A., & Mori, I. (2011). Novel and conserved protein macoilin is required for diverse neuronal functions in *Caenorhabditis elegans*. *PLoS Genetics*, 7(5), e1001384.
- Morgante, F., Sørensen, P., Sorensen, D. A., Maltecca, C., & Mackay, T. F. C. (2015). Genetic architecture of micro-environmental plasticity in *Drosophila melanogaster*. *Scientific Reports*, 5, 9785.
- Morin, L. P. (1986). Environment and hamster reproduction: responses to phase-specific starvation during estrous cycle. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 251(4), 663–669.
- Morozova, T. V., Mackay, T. F. C., & Anholt, R. R. H. (2014). Genetics and genomics of alcohol sensitivity. *Molecular Genetics and Genomics*, 289(3), 253–269.
- Morozova, T. V., Huang, W., Pray, V. A., Whitham, T., Anholt, R. R. H., & Mackay, T. F. C. (2015). Polymorphisms in early neurodevelopmental genes affect natural variation in alcohol sensitivity in adult *Drosophila*. *BMC Genomics*, 16(1), 865.
- Mosin, A. F. (1982). Some physiological and biochemical features of starvation and refeeding in small wild rodents (*Microtinae*). *Comparative Biochemistry and Physiology - Part A: Physiology*, 71(3), 461–464.
- Mosin, A. F. (1984). On the energy fuel in voles during their starvation. *Comparative Biochemistry and Physiology - Part A: Physiology*, 77(3), 563–565.
- Mrosovsky, N., & Sherry, D. F. (1980). Animal anorexias. *Science*, 207(6), 837–842.
- Nedergaard, J., & Cannon, B. (1990). Mammalian hibernation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 326(1237), 669–686.
- Owen, O. E., Morgan, A. P., Kemp, H. G., Sullivan, J. M., Herrera, M. G., & Cahill, G. F. (1967). Brain metabolism during fasting. *Journal of Clinical Investigation*, 46(10), 1589–1595.
- Owen, O. E., & Reichard, G. A. (1971). Human forearm metabolism during progressive starvation. *Journal of Clinical Investigation*, 50(7), 1536–1545.
- Owusu-Ansah, E., & Perrimon, N. (2014). Modeling metabolic homeostasis and

- nutrient sensing in *Drosophila*: implications for aging and metabolic diseases. *Disease Models & Mechanisms*, 7(3), 343–350.
- Ozaki, K., Ohnishi, Y., Iida, A., Sekine, A., Yamada, R., Tsunoda, T., ... Tanaka, T. (2002). Functional SNPs in the *lymphotoxin-alpha* gene that are associated with susceptibility to myocardial infarction. *Nature Genetics*, 32(4), 650–654.
- Paoli, A., Rubini, A., Volek, J., & Grimaldi, K. (2013). Beyond weight loss: a review of the therapeutic uses of very-low-carbohydrate (ketogenic) diets. *European Journal of Clinical Nutrition*, 67(10), 789–796.
- Parkash, R., & Munjal, A. K. (2000). Evidence of independent climatic selection for desiccation and starvation tolerance in Indian tropical populations of *Drosophila melanogaster*. *Evolutionary Ecology Research*, 2(5), 685–699.
- Parkes, M., Barrett, J. C., Prescott, N. J., Tremelling, M., Anderson, C. A., Fisher, S. A., ... Mathew, C. G. (2007). Sequence variants in the autophagy gene *IRGM* and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nature Genetics*, 39(7), 830–832.
- Partridge, L., Gems, D., & Withers, D. J. (2005). Sex and death: What is the connection? *Cell*, 120 (4), 461–472.
- Pollard, D. A. (2012). Design and construction of recombinant inbred lines. *Methods in Molecular Biology*, 871, 31–39.
- Pool, J. E., Corbett-Detig, R. B., Sugino, R. P., Stevens, K. A., Cardeno, C. M., Crepeau, M. W., ... Langley, C. H. (2012). Population genomics of sub-Saharan *Drosophila melanogaster*: African diversity and non-African admixture. *PLoS Genetics*, 8(12), e1003080.
- Price, T., & Schluter, D. (1991). On the low heritability of life-history traits. *Evolution*, 45(4), 853–861.
- Printz, M. P., Jirout, M., Jaworski, R., Alemayehu, A., & Kren, V. (2003). Genetic models in applied physiology. HXB/BXH rat recombinant inbred strain platform: a newly enhanced tool for cardiovascular, behavioral, and developmental genetics and genomics. *Journal of Applied Physiology*, 94, 2510–2522.
- Qi, L., Cornelis, M. C., Kraft, P., Stanya, K. J., Kao, W. H. L., Pankow, J. S., ... Hu, F. B. (2010). Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. *Human Molecular Genetics*, 19(13), 2706–2715.
- Rajan, A., & Perrimon, N. (2012). *Drosophila* cytokine unpaired 2 regulates physiological homeostasis by remotely controlling insulin secretion. *Cell*, 151(1), 123–137.
- Rion, S., & Kawecki, T. J. (2007). Evolutionary biology of starvation resistance: What we have learned from *Drosophila*. *Journal of Evolutionary Biology*, 20(5), 1655–1664.
- Robinson, S. J., Zwaan, B., & Partridge, L. (2000). Starvation resistance and adult body composition in a latitudinal cline of *Drosophila melanogaster*. *Evolution*, 54(5), 1819–1824.

- Robinson, M. R., Wray, N. R., & Visscher, P. M. (2014). Explaining additional genetic variation in complex traits. *Trends in Genetics*, 30(4), 124-132.
- Rockman, M. V., Skrovanek, S. S., & Kruglyak, L. (2010). Selection at linked sites shapes heritable phenotypic variation in *C. elegans*. *Science*, 330(6002), 372-376.
- Roff, D. A. (1995). The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture. *Heredity*, 74, 481-490.
- Rohde, P. D., Krag, K., Loeschcke, V., Overgaard, J., Sørensen, P., & Kristensen, T. N. (2016). A quantitative genomic approach for analysis of fitness and stress related traits in a *Drosophila melanogaster* model population. *International Journal of Genomics*, vol. 2016, 2157494.
- Rose, M. R., Vu, L. N., Park, S. U., & Graves, J. L. (1992). Selection on stress resistance increases longevity in *Drosophila melanogaster*. *Experimental Gerontology*, 27(2), 241-250.
- Rush, B., Sandver, S., Bruer, J., Roche, R., Wells, M., & Giebultowicz, J. (2007). Mating increases starvation resistance and decreases oxidative stress resistance in *Drosophila melanogaster* females. *Aging Cell*, 6(5), 723-726.
- Rusten, T. E., Cantera, R., Urban, J., Technau, G., Kafatos, F. C., & Barrio, R. (2001). Spalt modifies EGFR-mediated induction of chordotonal precursors in the embryonic PNS of *Drosophila* promoting the development of oenocytes. *Development*, 128(5), 711-722.
- Rutherford, S. L., & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature*, 396(6709), 336-342.
- Schmidt, P. S., Zhu, C.-T., Das, J., Batavia, M., Yang, L., & Eanes, W. F. (2008). An amino acid polymorphism in the *couch potato* gene forms the basis for climatic adaptation in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 105(42), 16207-16211.
- Schüpbach, T., Xenarios, I., Bergmann, S., & Kapur, K. (2010). FastEpistasis: A high performance computing solution for quantitative trait epistasis. *Bioinformatics*, 26(11), 1468-1469.
- Schwasinger-Schmidt, T. E., Kachman, S. D., & Harshman, L. G. (2012). Evolution of starvation resistance in *Drosophila melanogaster*: Measurement of direct and correlated responses to artificial selection. *Journal of Evolutionary Biology*, 25(2), 378-387.
- Sclafani, A., & Rendel, A. (1978). Food deprivation-induced activity in dietary obese, dietary lean, and normal-weight rats. *Behavioral Biology*, 24(2), 220-228.
- Secor, S. M., & Diamond, J. M. (2000). Evolution of regulatory responses to feeding in snakes. *Physiological and Biochemical Zoology*, 73(2), 123-141.
- Service, M. (1987). Physiological mechanisms of increased stress resistance in *Drosophila melanogaster* selected for postponed senescence. *Physiological Zoology*, 60(3), 321-326.
- Sham, P. C., & Purcell, S. M. (2014). Statistical power and significance testing in

- large-scale genetic studies. *Nature Reviews Genetics*, 15(5), 335–346.
- Shorter, J., Couch, C., Huang, W., Carbone, M. A., Peiffer, J., Anholt, R. R. H., & Mackay, T. F. C. (2015). Genetic architecture of natural variation in *Drosophila melanogaster* aggressive behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 112(27), E3555–E3563.
- Sillanpää, M. J. (2011). Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. *Heredity*, 106(4), 511–519.
- Simonis, M., Atanur, S. S., Linsen, S., Guryev, V., Ruzius, F.-P., Game, L., ... Cuppen, E. (2012). Genetic basis of transcriptome differences between the founder strains of the rat HXB/BXH recombinant inbred panel. *Genome Biology*, 13(4), r31.
- Sisodia, S., & Singh, B. N. (2012). Experimental evidence for nutrition regulated stress resistance in *Drosophila ananassae*. *PLoS ONE*, 7(10), e46131.
- Smith, A. V., & Orr-Weaver, T. L. (1991). The regulation of the cell cycle during *Drosophila* embryogenesis: the transition to polyteny. *Development*, 112(4), 997–1008.
- Sorensen, J. G., Nielsen, M. M., & Loeschcke, V. (2007). Gene expression profile analysis of *Drosophila melanogaster* selected for resistance to environmental stressors. *Journal of Evolutionary Biology*, 20(4), 1624–1636.
- Stahl, E. A., Raychaudhuri, S., Remmers, E. F., Xie, G., Eyre, S., Thomson, B. P., ... Plenge, R. M. (2010). Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nature Genetics*, 42(6), 508–514.
- St Johnston, D. (2002). The art and design of genetic screens: *Drosophila melanogaster*. *Nature Reviews. Genetics*, 3(3), 176–188.
- Stieper, B. C., Kupershtok, M., Driscoll, M. V., & Shingleton, A. W. (2008). Imaginal discs regulate developmental timing in *Drosophila melanogaster*. *Developmental Biology*, 321(1), 18–26.
- Stranger, B. E., Stahl, E. A., & Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*, 187(2), 367–383.
- Struchalin, M. V., Dehghan, A., Witteman, J. C. M., van Duijn, C., & Aulchenko, Y. S. (2010). Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations. *BMC Genetics*, 11, 92.
- Swarup, S., Huang, W., Mackay, T. F. C., & Anholt, R. R. H. (2013). Analysis of natural variation reveals neurogenetic networks for *Drosophila* olfactory behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 110(3), 1017–1022.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., ... Von Mering, C. (2015). STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, 43(D1), D447–D452.

- Talmud, P. J., Drenos, F., Shah, S., Shah, T., Palmen, J., Verzilli, C., ... Hingorani, A. D. (2009). Gene-centric association signals for lipids and apolipoproteins identified via the HumanCVD BeadChip. *American Journal of Human Genetics*, 85(5), 628–642.
- Teixeira, L., Ferreira, Á., & Ashburner, M. (2008). The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biology*, 6(12), e1000002.
- Tennessen, J. M., & Thummel, C. S. (2011). Coordinating growth and maturation - insights from *Drosophila*. *Current Biology*, 21(18), R750-R757.
- Terashima, J., & Bownes, M. (2005). A microarray analysis of genes involved in relating egg production to nutritional intake in *Drosophila melanogaster*. *Cell Death and Differentiation*, 12(5), 429–440.
- Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., ... Kathiresan, S. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, 466(7307), 707–713.
- Tettweiler, G., Miron, M., Jenkins, M., Sonenberg, N., & Lasko, P. F. (2005). Starvation and oxidative stress resistance in *Drosophila* are mediated through the eIF4E-binding protein, d4E-BP. *Genes and Development*, 19(16), 1840–1843.
- The International Hapmap Consortium. (2003). The International HapMap Project. *Nature*, 426(6968), 789–796.
- Thimman, M. S., Seugnet, L., Turk, J., & Shaw, P. J. (2015). Identification of genes associated with resilience/vulnerability to sleep deprivation and starvation in *Drosophila*. *Sleep*, 38(5), 801–814.
- Threadgill, D. W., Miller, D. R., Churchill, G. A., & de Villena, F. P.-M. (2011). The collaborative cross: a recombinant inbred mouse population for the systems genetic era. *ILAR Journal*, 52(1), 24–31.
- Thummel, C. S. (2001). Molecular Mechanisms of Developmental Timing in *C. elegans* and *Drosophila*. *Developmental Cell*, 1(4), 453–465.
- Tian, C., Gregersen, P. K., & Seldin, M. F. (2008). Accounting for ancestry: Population substructure and genome-wide association studies. *Human Molecular Genetics*, 17(R2), R143-R150.
- Tobey, J. A. (1951). The Biology of Human Starvation. *American Journal of Public Health and the Nations Health*, 41(2), 236–237.
- Tran, C., Konstantopoulou, V., Mecija, M., Perlman, K., Mercimek-Mahmutoglu, S., & Kronick, J. B. (2015). Hyperinsulinemic hypoglycemia: think of hyperinsulinism/hyperammonemia (HI/HA) syndrome caused by mutations in the *GLUD1* gene. *Journal of Pediatric Endocrinology and Metabolism*, 28(7–8), 873–876.
- Turner, S., Armstrong, L. L., Bradford, Y., Carlson, C. S., Crawford, D. C., Crenshaw, A. T., ... Ritchie, M. D. (2011). Quality control procedures for genome-wide association studies. In Haines, J. L., et al., (Ed.), *Current Protocols in Human Genetics*, Chapter 1, Unit 1.19.

- Unckless, R. L., Rottschaefer, S. M., & Lazzaro, B. P. (2015). A genome-wide association study for nutritional indices in *Drosophila*. *G3: Genes, Genomes, Genetics*, 5(3), 417–425.
- Van Herrewege, J., & David, J. R. (1997). Starvation and desiccation tolerances in *Drosophila*: Comparison of species from different climatic origins. *Ecoscience*, 4(2), 151–157.
- Vijendravarma, R. K., Narasimha, S., & Kawecki, T. J. (2013). Predatory cannibalism in *Drosophila melanogaster* larvae. *Nature Communications*, 4, 1789.
- Voight, B. F., & Pritchard, J. K. (2005). Confounding from cryptic relatedness in case-control association studies. *PLoS Genetics*, 1(3), e32.
- Vonesch, S. C., Lamparter, D., Mackay, T. F. C., Bergmann, S., & Hafen, E. (2016). Genome-wide analysis reveals novel regulators of growth in *Drosophila melanogaster*. *PLoS Genetics*, 12(1), e1005616.
- Wang, T., Hung, C. C. Y., & Randall, D. J. (2006). The comparative physiology of food deprivation: from feast to famine. *Annual Review of Physiology*, 68(3), 223–251.
- Wayne, M. L., Soundararajan, U., & Harshman, L. G. (2006). Environmental stress and reproduction in *Drosophila melanogaster*: starvation resistance, ovariole numbers and early age egg production. *BMC Evolutionary Biology*, 6, 57.
- Weber, K., Eisman, R., Higgins, S., Morey, L., Patty, A., Tausek, M., & Zeng, Z. B. (2001). An analysis of polygenes affecting wing shape on chromosome 2 in *Drosophila melanogaster*. *Genetics*, 159(3), 1045–1057.
- Weber, A. L., Khan, G. F., Magwire, M. M., Tabor, C. L., Mackay, T. F. C., & Anholt, R. R. H. (2012). Genome-wide association analysis of oxidative stress resistance in *Drosophila melanogaster*. *PloS ONE*, 7(4), e34745.
- Wen, T., Parrish, C. A., Xu, D., Wu, Q., & Shen, P. (2005). *Drosophila* neuropeptide F and its receptor, NPFR1, define a signaling pathway that acutely modulates alcohol sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), 2141–2146.
- Westerterp, K. (1977). How rats economize energy loss in starvation. *Physiological Zoology*, 50(4), 331–362.
- Williams, A. E., Rose, M. R., & Bradley, T. J. (2004). The respiratory pattern in *Drosophila melanogaster* selected for desiccation resistance is not associated with the observed evolution of decreased locomotory activity. *Physiological and Biochemical Zoology*, 77(1), 10–17.
- Williams, M. J., Goergen, P., Rajendran, J., Klockars, A., Kasagiannis, A., Fredriksson, R., & Schiöth, H. B. (2014a). Regulation of aggression by obesity-linked genes *TfAP-2* and *Twz* through octopamine signaling in *Drosophila*. *Genetics*, 196(1), 349–362.
- Williams, M. J., Goergen, P., Rajendran, J., Zheleznyakova, G., Hägglund, M. G., Perland, E., ... Schiöth, H. B. (2014b). Obesity-linked homologues *TfAP-2* and

- Twz establish meal frequency in *Drosophila melanogaster*. *PLoS Genetics*, 10(9), e1004499.
- Wong, M. M.-K., Liu, M.-F., & Chiu, S. K. (2015). *Cropped*, *Drosophila* transcription factor AP-4, controls tracheal terminal branching and cell growth. *BMC Developmental Biology*, 15, 20.
- Wu, Q., Zhang, Y., Xu, J., & Shen, P. (2005). Regulation of hunger-driven behaviors by neural ribosomal S6 kinase in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 102(37), 13289–13294.
- Xi, B., Shen, Y., Reilly, K. H., Zhao, X., Cheng, H., Hou, D., ... Mi, J. (2013). Sex-dependent associations of genetic variants identified by GWAS with indices of adiposity and obesity risk in a Chinese children population. *Clinical Endocrinology*, 79(4), 523–528.
- Xu, S. (2003). Theoretical basis of the Beavis effect. *Genetics*, 165(4), 2259–2268.
- Yang, H., & Yamashita, Y. M. (2015). The regulated elimination of transit-amplifying cells preserves tissue homeostasis during protein starvation in *Drosophila* testis. *Development*, 142(10), 1756–1766.
- Yang, Z., Yu, Y., Zhang, V., Tian, Y., Qi, W., & Wang, L. (2015). Octopamine mediates starvation-induced hyperactivity in adult *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 112(16), 5219–5224.
- Zeggini, E., Scott, L. J., Saxena, R., Voight, B. F., Marchini, J. L., Hu, T., ... Altshuler, D. (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nature Genetics*, 40(5), 638–645.
- Zinke, I., Schütz, C. S., Katzenberger, J. D., Bauer, M., & Pankratz, M. J. (2002). Nutrient control of gene expression in *Drosophila*: microarray analysis of starvation and sugar-dependent response. *The EMBO Journal*, 21(22), 6162–6173.
- Zwaan, B. J., Bijlsma, R., & Hoekstra, R. F. (1991). On the developmental theory of ageing. I. Starvation resistance and longevity in *Drosophila melanogaster* in relation to pre-adult breeding conditions. *Heredity*, 66, 29–39.

10. ACKNOWLEDGEMENTS

I would like to thank to the people who have tutored, supported or helped me during this project. First of all, I am very grateful to my PhD advisor Prof. Ernst Hafen and my direct supervisor Dr. Katja Koehler for giving me the opportunity to do this project, introducing to me the field of population genetics, and for their guidance and support during the project.

I would like to thank to my committee members Prof. Matthias Peter and Prof. Trudy Mackay for their invaluable inputs to the project.

I am thankful to Dr. Sibylle Vonesch for her advises during the project and her guidance in the statistical analysis of the data. I am thankful to Masters students Stefan Zeyen and Sophie Zahalka for their contribution to the experiments.

I would like to thank to Dr. Hugo Stocker, Dr. Ryohei Yagi and Dr. Hirokazu Okada for very fruitful discussions.

I would also like to thank to all past and present members of Hafen and Stocker labs for being helpful and supportive to me, and for creating a friendly working environment.

11. CURRICULUM VITAE

Surname: Akcay

Name: Izzet Mehmet

Date of birth: 13.06.1985

Place of birth: Istanbul, Turkey

Education:

2000-2003

Kultur Science High School, CGPA: 5.00/5.00

2003-2007

BSc in Molecular Biology and Genetics, Middle East Technical University, CGPA: 3.51/4.00

2008-2010

MSc in Biological Sciences and Bioengineering, Sabanci University, CGPA: 3.92/4.00

Thesis title: Transcriptional regulation of *IL-7R α* gene in T lymphocytes supervised by Assoc. Prof. Batu Erman

2011-2016

PhD in Institute of Molecular Systems Biology, ETH Zurich

Thesis title: Quantitative Genetics of Larval Starvation Resistance in *Drosophila melanogaster*

supervised by Prof. Ernst Hafen