*DISSECTING THE GENETIC BASIS OF LEARNING, MEMORY, AND
THERMAL TOLERANCE IN A MULTI-PARENTAL POPULATION OF FRUIT
FLIES*

A Dissertation

Presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by Patricka Williams-Simon

Dr. Elizabeth G. King and Dr. Troy Zars (Deceased), Dissertation Supervisors

July 2021

The undersigned, appointed by the Associate Vice Chancellor of the Office of
Research

and Graduate Studies, have examined the dissertation entitled

DISSECTING THE GENETIC BASIS OF LEARNING, MEMORY, AND
THERMAL TOLERANCE IN A MULTI-PARENTAL POPULATION OF FRUIT
FLY

presented by Patricka Williams-Simon,
a candidate for the degree of doctor of philosophy of Biology,
and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Elizabeth G. King

Dr. Troy Zars (deceased)

Dr. David Schulz

Dr. Elizabeth Bryda

**To the many people whose shoulders I stand on, especially
Terry, Terrin, and Jace**

ACKNOWLEDGEMENTS.

I take this opportunity to thank the people who helped, supported, encouraged, and guided me through this phase in my career. I am fortunate to have had people who cheered me on, even when I wanted to give up. I am honored and lucky to share this journey with them. It truly takes a village. They are my village!

First, I thank my advisors, Dr. Elizabeth (Libby) King, and Dr. Troy Zars, for allowing me into their labs and seeing the potential in me. I am grateful for their outstanding mentorship and friendships. I also thank you for your unwavering support of me, for pushing me, and for believing me. I truly appreciate all that you both have done, and I am truly honored to have trained under you. I would have it no other way. To Dr. Edward Catapane and Dr. Margaret Carroll, thanks for my first research experience, thanks for setting a high standard for excellent mentoring. I am grateful. To Dr. Chris Lorson, thanks for allowing me to do great science and for confirming my passion for research. My experience in your lab was phenomenal. To my committee members, Dr. David Schulz and Dr. Elizabeth Bryda, thanks for your guidance and support. To Dr. NaTashua Davis, an exceptional mentor. Thanks for opening doors for me. I appreciate all that you have done for me. To Dr. John David and Dr. Mark Hannink, thank you, thank you, thank you! I am grateful to all!

To my husband Terry, my daughter Terrin, and my son Jace. You are the real heroes. None of this would be possible without you. Thanks for taking this journey with me. Terry, thanks for being my best friend and biggest fan, thanks for listening to ALL my rants, thanks for understanding I had to work late nights and weekends, thanks for your unconditional love and support. Terrin, you have no idea how much you have pushed me for the past 13 years. I did this for you baby girl, and I hope mommy made you proud. I love you, my Pooka! To my mom and best friend Patricia (Alice), and my dad Patrick and Desmond, my biggest cheerleaders. Your daughter has made it! I thank you for ALWAYS being there for me. All the daily phone calls, listening to my rants, encouragement, and for your immense support, I am eternally grateful. I love you! To my sisters, Kolene, Kolette, Shereen, and Kelli, and my brother Jaime, thanks for not talking science with me. Thanks for the great laughs and support. To my grandma, affectionately known as Aunt Stella, this is for you! From the plantations to PhD, your granddaughter is a Dr. I hope I have made you proud, Aunt Stella, and it's the least I can do. To my aunt, Marvis, who is no longer with us, thank you for helping to develop my love for learning and my curiosity. This PhD is equally yours. To all other family members, I say thank you!

To Nancy, thanks for your support and for being a great friend who turned into a family. I love you, mi amiga. To Candice and Emi. Thank you! There are no words to explain how much your friendship and unconditional support means to me. You both were ALWAYS there for my family and me. This journey would not

have been as enjoyable without you. Thanks for the laughs, the memories, the photos. I am eternally grateful!!! To Deise and Aditi, thanks for all your support, lunch dates, laughs, and for listening to my rants. I would have it no other way. Thanks for allowing me to share your PhD journey with you.

To my lab mates, especially Enoch, Ronel, and Chris, thanks for all your support through this journey. To my dear friends Kathy, Patricia, Lindell, Mendy, Mellissa, Chris, Carrell, Rachel, and Eddie, thanks for all your support! To my Mi-STEM, Gilliam, GES, and TAE colleagues, I say thank you. To Debbie Allen, thanks for ALL your support. The DBS support staff, especially Melody, Anantha, and Nick, thanks for all that you have done to make my stay at DBS easier.

Thanks to everyone who supported me on this journey. I stand on many shoulders.

TABLE OF CONTENTS.

ACKNOWLEDGEMENTS.	ii
List of Illustrations (Figures and Tables)	vi
ABSTRACT	
Chp 1 – INTRODUCTION	1
1.1 Drosophila melanogaster as an Animal Model	3
1.2 The Genetics of Learning, Memory, and Thermal Tolerance	6
1.3 Quantitative Genetics – Phenotypes to Genotypes	15
1.4 Combining Approaches – Whole Genome Scans, Fine-Mapping, and Single Gene Mutagenesis	19
1.5 Conclusion	21
1.6 Figures/ Schematics:	22
Chp 2 – THE GENETIC BASIS OF LEARNING AND MEMORY IN A MULTI-PARENTAL POPULATION OF FRUIT FLIES	
2.1 Abstract:	25
2.2 Introduction:	26
2.3 Methods:	31
2.4 Results:	40
2.5 Discussion:	49
2.6 Acknowledgements:	55
2.7 Figure:	
CHP 3 – DISSECTING THE GENETIC BASIS OF THERMAL TOLERANCE IN FRUIT FLIES	62
3.1 Abstract:	63
3.2 Introduction:	64
3.2 Methods:	68
3.4 Results:	73
2. 5 Discussion:	78
2. 6 Figures:	85
CHP 4 – FUNCTIONALLY VALIDATING CANDIDATE GENES INFLUENCING IN THERMAL TOLERANCE D. MELANOGASTER	
4.1 Abstract:	104
4.2 Introduction	105

4.3 Methods:	109
4.4 Results:	114
4.5 Discussion:	116
4.6 Figures:	
4.7 Tables:	125
CHP 5 – CONCLUSION.	127
BIBLIOGRAPHY	132
VITA	187

List of Illustrations (Figures and Tables)

Chapter 1: Introduction.

Figure 1. Schematic explaining interval QTL mapping	25
Figure 2. Schematic showing different integrative approaches	26

Chapter 2: The Genetic Basis of Learning and Memory in a Multi-parental Population of Fruit Flies.

Figure 1. Schematic of phenotyping process	58
Figure 2. Phenotypic patterns in learning and memory	59
Figure 3. Genome scan of learning and memory phenotypes	60
Figure 4. Haplotype means at each shared QTL	61
Figure 5. Expression Data	62
Figure 6. Individual QTL BCI regions showing possible candidate <i>genes</i>	63

Chapter 3: Dissecting the Genetic Basis of Thermal Tolerance In Fruit Flies.

Figure 1. Phenotypic variation in thermal tolerance	
a. RILs	87
b. Founders	88
c. Correlation between thermal tolerance and Learning	89
Figure 2. Genomic regions that influence thermal tolerance	90
Figure 3. Genome-wide expression data	
a. MA plots.....	91
b. Volcano Plots (pool and condition)	92
c. Volcano Plots (pair-wise)	93
d. PCA	94
Figure 4. QTL 6 interval	95
Figure 5. Haplotype effects	96
a. Q6 interval	97
b. QTLs intervals	98
Table 1. Significant genes within QTLs	99
Table 2. Significant genes within Q6	100
Table 3. Q6 genes information	101

Chapter 4: Functionally Validating Candidate Genes Influencing in thermal tolerance *Drosophila melanogaster*.

Figure 1 Q6 BCI interval	122
Figure 2 Thermal tolerance measurements for <i>atpalpha</i>	123
Figure 3. Thermal tolerance measurements for <i>dmrt938</i>	124
Figure 4 Thermal tolerance measurements for <i>RhoGAP93B</i>	125

Table 1 List of the 20 UAS-RNAi lines	126
Table 2. List of structural variance	127

ABSTRACT

Learning, memory, and thermal tolerance are complex traits that are fundamental survival skills in many species. For example, small poikilotherms – like *Drosophila melanogaster*, have evolved physiological mechanisms to learn from and adapt to rapidly warming temperatures because their body temperature reflects that of their surroundings. Along with physiological mechanisms, these animals have also adapted cognitive practices, such as learning and memory, which help them escape extreme temperatures. Although scientific evidence shows that learning, memory, and thermal tolerance have been subjected to natural selection as a means of adapting to warmer temperatures, studies examining both traits (i.e., learning, and thermal tolerance) and how they relate genetically have not been done. Investigating the natural variation of the genes that control these traits is essential if we want to understand better how these traits arise within species. With recent advances in genomic technology, it is now feasible to tease apart - down to a single gene - regions of the genome that influence complex phenotypes. For my Ph.D., I took a quantitative genetics and genomics approach to dissect the genetic basis of learning, memory, and thermal tolerance in *Drosophila melanogaster*. I hypothesized that there will be a shared and independent genetic locus controlling learning and memory and that there will be a trade-off between learning and thermal tolerance.

To genetically dissect learning and memory performance, I used the *Drosophila* Synthetic Population Resource (DSPR), a multiparent mapping

resource, consisting of a large set of recombinant inbred lines (RILs), which allows for high-resolution genome-wide scans, and the identification of loci contributing to naturally occurring genetic variation. Using a highly sensitive apparatus known as the “heat box”, I trained flies with temperatures (24°C and 41°C, which is aversive) to learn to remain on one side of a chamber (place learning) and to remember this (place memory). An individual that spent most of the time on the side of the chamber associated with 24°C will be considered to have high learning and memory.

Immediately after, I assayed the flies for thermal tolerance, which is measured as the time to incapacitation at a constant high 41°C temperature for 9 ½ mins. I identified 16 different loci across the genome that significantly affect place learning and or memory performance, with 5 of these loci affecting both traits and eight loci that influence thermal tolerance. To identify transcriptomic differences associated with learning, memory, and thermal tolerance, I performed RNA-Sequencing (RNA-seq) on pooled samples of seven high-performing and seven low-performing RILs for all three traits and identified hundreds of genes with differences in expression. Integrating our transcriptomic results with the mapping results allowed us to identify nine promising candidate genes that control learning and twenty-one genes that fall between the most significant QTL that is possibly influencing thermal tolerance.

I then used a few different methods to identify and validate genetic variants from the list of candidate genes within the most significant QTL (Q6) that influences thermal tolerance. First, I performed a structural variance analysis and found a few

structural variants (SV; eg: inserts, deletions, or P-elements) located within candidate genes. Genes that contain SVs are known to influence phenotypic variation, and thus identifying any SV within candidate genes will provide a greater understanding of the genetic basis of a specific complex trait. Second, I functionally validated the candidate genes to identify causal genes using the UAS-GAL4/RNAi system. In this classical genetic system, a specific gene is disrupted through mutagenesis, which allows for the investigation of that gene on a specific phenotype. I then phenotyped lines to determine the phenotypic effect of each gene on thermal tolerance. We screened 20 UAS-RNAi lines (~1900 individuals) that were crossed with both pan neuronal and ubiquitously expressed Gal4 promoters and found 3 lines that were significant.

Collectively, these results suggest that learning, memory, and thermal tolerance are highly polygenic and that multiple genes with both large and small effects influence these traits. Future work will aim to validate the function of candidate genes that influence learning and memory and identify the molecular pathways that the causal genes that influence thermal tolerance are associated with.

Chp 1 – INTRODUCTION

The ability to adapt to shifts in temperature influences the distribution and survival of many species. Physiological mechanisms (thermal tolerance; (Tosini and Menaker 1995; Hamada *et al.* 2008) and behavioral plasticity (learning and memory; (Mery *et al.* 2007c; Kawecki 2010) are adaptive traits that allow organisms to compensate for warmer temperatures, however, there are very few studies that elucidate the underpinning of the genetic basis of those traits, and therefore the genetic basis of learning, memory, and thermal tolerance remain unclear. Poikilotherms, like *Drosophila*, are especially susceptible to rapidly warming temperatures (Stevenson 1985), since their body temperatures reflect that of their surroundings. As climate change continues to be a problem, it is increasingly imperative that we identify the genes contributing to the natural variation of learning, memory, and thermal tolerance. This knowledge would be fundamental in helping us comprehend how these traits have evolved and continue to evolve in a rapidly changing environment (Mery and Kawecki 2003; Dukas 2004; Calosi *et al.* 2008; Baudier *et al.* 2018). Interestingly, even if a specific species has not encountered a specific environment in the evolutionary past, behaviors allow individuals to develop responses (Hutchison and Maness 1979) and these behaviors are predicted to play a major role in thermal adaptation (Hutchison and Maness 1979; Charmantier *et al.* 2008; Chevin *et al.* 2013b, 2013a). Because of

the critical role learning and memory, and thermal tolerance continue to play in the survival of many species, it is essential that we better understand the variation of the genetic architecture that influences those traits.

Despite decades of work on learning and memory (Tolman 1924; McGuire and Hirsch 1977c; Lofdahl *et al.* 1992); (Aceves-Piña *et al.* 1983; Levin *et al.* 1992; Bourtchuladze *et al.* 1994; Silva *et al.* 1997); (Zars *et al.* 2000b; Sitaraman *et al.* 2008; Kahsai and Zars 2011; LaFerriere *et al.* 2011; Ostrowski and Zars 2014)(LaDage *et al.* 2012; Ko *et al.* 2020)(Zars *et al.* 2000b; Sitaraman *et al.* 2008; Kahsai and Zars 2011; LaFerriere *et al.* 2011; Ostrowski and Zars 2014) and thermal tolerance (Velazquez *et al.* 1983; Hoffmann *et al.* 2003; Dhaka *et al.* 2006); (Stevenson 1985; Tosini and Menaker 1995; Hamada *et al.* 2008); (Huey *et al.* 2009; Harley and Paine 2009; Allen *et al.* 2010; McKechnie and Wolf 2010; Seuront *et al.* 2019); (Overgaard *et al.* 2011; MacLean *et al.* 2019b), there are very few studies that have identified suggestive candidate variants affecting learning and or memory (Chandra *et al.* 2001; Steinberger *et al.* 2003; Papassotiropoulos *et al.* 2011; Mery 2013a; Parker *et al.* 2014; Knoll *et al.* 2016), and thermal tolerance (Norry *et al.* 2004, 2007a, 2007b, 2008; Morgan and Mackay 2006; Mery *et al.* 2007a; Rand *et al.* 2010; Papassotiropoulos *et al.* 2011; Rolandi *et al.* 2018). This is due to most studies being focused on single-gene approaches (Dudai and Zvi 1985; Zars *et al.* 2000b; LaFerriere *et al.* 2011; Ostrowski *et al.* 2015) where a specific gene is mutated and researchers investigate what happens to a particular phenotype (Zars *et al.* 2000b; Hong and Vierling 2000). This approach, however, could result in dubious findings, as **1**) learning, memory, and thermal tolerance are

complex traits, which by definition have multiple genes contributing to those phenotypes, and **2)** the genes that were found through mutagenesis may not be a true reflection of what occurs within natural populations. For this Dissertation, I took an unbiased approach and utilized quantitative and behavioral approaches to dissect the genes that are influencing learning, memory, and thermal tolerance in a wild-type population of *Drosophila melanogaster*.

1.1 *Drosophila melanogaster* as an Animal Model

D. melanogaster, commonly known as the fruit fly, has been used as an animal model for more than a century (Allen 1978). From the proposal to use *Drosophila melanogaster* (fruit flies) as a model organism by Charles W. Woodworth in the 1900s to a now fully annotated genome (Adams *et al.* 2000), the field of *D. melanogaster* research has made significant advances in a relatively short time. This is primarily due to the fact that fruit flies are phenomenal models for biological research. Fruit flies have a short life span and huge population, are relatively cheap to maintain and are easily genetically manipulated, reviewed in (Jennings 2011; Yamaguchi and Yoshida 2018). Fruit flies have been used in a broad range of research fields, including molecular mechanisms underlying genetic traits, behavior, and development (Jennings 2011; Kuhnlein 2012; Tolwinski 2017; Yamaguchi and Yoshida 2018; Douglas 2018). Interestingly, most of these findings have been essential for understanding human health and diseases, because approximately 75% of human disease-causing genes are believed to have a functional homolog in fruit flies (Reiter *et al.* 2001; Chien *et al.*

2002). In addition, because of decades of research with fruit flies, there are a plethora of genetic resources that are available, including a fully annotated genome (Adams *et al.* 2000). One of the successes that stem from annotating the fruit fly genome is the ability for scientists to continue to push the envelope on genetic research and has allowed scientists to create cutting-edge approaches to answer questions that would have otherwise been impossible (Reiter *et al.* 2001; Oleksiak *et al.* 2002; Rifkin *et al.* 2003; Cancer Genome Atlas Network 2015; McFaline-Figueroa *et al.* 2020). This breakthrough led to the ‘omics’ era. Omics technologies provided researchers with the capabilities to measure roles and relationships of different cellular molecules at any given snapshot (e.g. (Reiter *et al.* 2001; Oleksiak *et al.* 2002; Rifkin *et al.* 2003; Cancer Genome Atlas Network 2015; McFaline-Figueroa *et al.* 2020)). These technologies include genomics, the study of genes and their function; transcriptomics, the study of mRNA; proteomics, the study of proteins; metabolomics, which focuses on the cellular metabolism of molecules (Schneider and Orchard 2011). Broadly, these technologies allow for investigation and/or comparison from a single cell to specific tissue processes and between species and among species. Omics is high-throughput and usually produces a vast amount of specified data on the snapshot of a cell or tissue. This application is applied in numerous fields in biology, such as human aging (Valdes *et al.* 2013), cancer biology (Cancer Genome Atlas Network 2015), and food safety (Davies 2010), and has facilitated significant understanding of the genetic basis of many complex traits. For example, in this dissertation, I used transcriptomics to identify genes that influence learning, memory (Williams-Simon *et al.* 2019), and

thermal tolerance between different pools of cohorts. Additionally, Highfill et al. (Highfill *et al.* 2017) used transcriptomics to dissect the naturally segregating variants that contribute to nicotine resistance, and Stanley et al. (Stanley *et al.* 2017), used genomics to elucidate the genetic basis of the insulin-signaling pathway based on diets. Indeed, all three of these studies were conducted in the *D. melanogaster*.

One of the powerful resources found in *D. melanogaster* and other species that are used in dissecting complex traits is Multi-Parental Populations (MPPs; reviewed in (de Koning and McIntyre 2017a). The general idea behind MPPs is that the ending population is an integration of the genome of many founding parents. This is advantageous to dissecting complex traits because there is variation in genotypes (haplotypes), which ultimately leads to variation in phenotypes. Perhaps, MPPs sit in the intersection between classical genetics and natural ecological populations and leverage the advantages of both approaches. I say this because the genome of MPPs is undisturbed (natural population), however, MPPs are reared in the lab in a controlled environment (classical genetics populations). In *D. melanogaster*, there are two major MPPs, The Drosophila Genetic Panel Reference (DGRP; (Mackay *et al.* 2012) and The Drosophila Synthetic Population Resource (DSPR; (King *et al.* 2012a, 2012b; Long *et al.* 2014); (<http://FlyRILs.org>). In my research, I use the DSPR. Briefly, the DSPR consists of about 1,600 Recombinant Inbred Lines (RILs) from two separate populations, which were derived by crossing the 16 founding parents (8 each population) for several generations to create a mosaic of their genome (King *et al.*

2012a). In addition to the huge fly population, the DSPR also comes with statistical packages to conduct genome scans, permutations, and codes for data visualization (King *et al.* 2012b). The DSPR makes answering the type of scientific questions that I am interested in more feasible.

As a field, fruit fly research has led to a plethora of groundbreaking findings which has tremendously expanded our scientific knowledge on a broad spectrum of topics (Jennings 2011), such as development, neuroscience, evolution, and medicine. Today, *D. melanogaster* and its sister species continue to be one of the leading model organisms in the development of cutting-edge techniques that continue to push the scientific envelope and will continue to do so for many years to come.

1.2 The Genetics of Learning, Memory, and Thermal Tolerance

Learning and Memory – I broadly define learning as the change in an individual's behavior based on experience, and memory as a persistent change in behavior in the absence of the experience, which is based on (Pearce 2008b). This definition encompasses many different types of learning and memory, including associative (Pascual and Prat 2001; Pitman *et al.* 2009; Busto *et al.* 2010), classical (Pitman *et al.* 2009; Kahsai and Zars 2011; Oswald *et al.* 2015), operant (Wustmann *et al.* 1996; Zars *et al.* 2000b; Kahsai and Zars 2011), and place (LaFerriere *et al.* 2008, 2011; Sitaraman *et al.* 2008, 2012), learning and memory. In this project, I paired learning and memory although I recognize that learning and

memory are separate phenotypes. These phenotypes often paired together in research designs **1)** because individuals are not able to remember (memory) without first learning (Krashes *et al.* 2007; Aso and Rubin 2016; Takemura *et al.* 2017), and **2)** because of the high genetic correlation between the two (Takemura *et al.* 2017; Williams-Simon *et al.* 2019).

Learning and memory have been extensively studied over the decades. There have been multiple learning and memory studies in non-model organisms, for example, birds (Healy and Hurly 2004; Tello-Ramos *et al.* 2015, 2018; Croston *et al.* 2016; McQuillan *et al.* 2018), lizards (LaDage *et al.* 2012; Batabyal and Thaker 2019; Ko *et al.* 2020), butterflies (Langham 2004; Toure *et al.* 2020), and bees (Gould 1987; Tison *et al.* 2017; Siviter *et al.* 2018; Baracchi *et al.* 2020). These studies have demonstrated that learning and memory behaviors are conserved between different species and are an imperative survival skill in a wide range of taxa. Early on, many experiments established that learning and memory have a genetic basis and are heritable traits. These include studies in humans (Galton 1875; Kesner *et al.* 1992), single-gene mutagenesis approaches in both rodents and insects (Aceves-Piña *et al.* 1983; Levin *et al.* 1992; Bourtchuladze *et al.* 1994; Silva *et al.* 1997), and selection experiments (Tolman 1924; McGuire and Hirsch 1977c; Lofdahl *et al.* 1992). Collectively, these findings have uncovered that learning and memory have a wide phenotypic range between different species, suggesting the genetic complexity of those traits.

Initial progress toward identifying the genes involved in learning and memory focused on mutants. Specifically in *D. melanogaster*, there were several

early experiments that developed learning and memory mutants through mutagenesis including (Dudai *et al.* 1976; Livingstone *et al.* 1984), which ultimately led to the discovery of *dunce* (*dnc*; (Dudai *et al.* 1976), and *rutabaga adenyl cyclase* (*rut*; (Livingstone *et al.* 1984), two or more commonly known genes that have been shown to influence learning and memory in fruit flies. The development of new techniques allowed for more targeted single-gene mutagenesis, where researchers perturb a particular gene and then study what effect/s that has on the phenotype (Tang *et al.* 1999; Zars *et al.* 2000c, 2000a; Tonegawa *et al.* 2003; Godenschwege *et al.* 2004; Krashes and Waddell 2008; LaFerriere *et al.* 2008; Ostrowski *et al.* 2015). In addition, molecular pathways and neuronal connections were identified, including the parts of the fly brain that are responsible for short-term memory (Zars *et al.* 2000a). Lastly, binary genetic tools such as the GAL4/UAS (Brand and Perrimon 1993; Duffy 2002), and *lexA/LexAop* (Little and Mount 1982; Butala *et al.* 2009), which are systems that target specific genes ubiquitous or in specific tissues to modulate their function (Pfeiffer *et al.* 2010). These systems, in particular, the GAL4/UAS system have led to significant progress towards better understanding some of the molecular processes that are involved in learning and memory and other complex traits, both in fruit flies and other models (Kitamoto 2001; Marella *et al.* 2006; Garrity *et al.* 2010; von Philipsborn *et al.* 2011; Sitaraman *et al.* 2015; Hige *et al.* 2015). For instance, the GAL4/UAS system has been extensively used to localize gene function to specific tissue types, including neuronal tissues that influence learning and memory (Akalal *et al.* 2006; Kasuya *et al.* 2009; Kramer *et al.* 2011; Sitaraman *et al.* 2012; Hige *et*

al. 2015). Moreover, the RNA interference (RNAi) system targets the expression of single genes within, by degrading the messenger RNA (mRNA), therefore reducing expression significantly reviewed in (Kuwabara and Coulson 2000; Perrimon *et al.* 2010). The RNAi system has been functionally validated in many learning and memory paradigms and within different model organisms (Threlkeld *et al.* 2007; Takahashi *et al.* 2009; Pan *et al.* 2009; Awata *et al.* 2016). Finally, recently the CRISPR/Cas9 system has been developed for targeted gene insertion or deletion (Jinek *et al.* 2012; Ran *et al.* 2013; Bassett and Liu 2014; Gratz *et al.* 2015; Meltzer *et al.* 2019). Broadly, CRISPR/Cas9 system has been used in many different applications in fruit flies, including identifying variants that contribute to nicotine resistance (Highfill *et al.* 2017), courtship song (Ding *et al.* 2016), and mate preference (Seeholzer *et al.* 2018). Although there are only a few studies (Kan *et al.* 2020) that have used CRISPR in fruit flies to elucidate the genetic basis of learning and memory, the CRISPR/Cas9 system, in combination with other approaches has the potential to revolutionize our knowledge on the genetic basis of learning and memory.

There are a few studies that took genome-wide approaches to elucidate the genetic basis of learning and or memory. These studies were conducted in a variety of different species, including human (Papassotiropoulos *et al.* 2011), mouse (Steinberger *et al.* 2003; Parker *et al.* 2012, 2014; Knoll *et al.* 2016), fruit flies (Mery *et al.* 2007a), and honey bees (Chandra *et al.* 2001). Of these, even fewer studies have successfully mapped QTLs for learning and memory (Chandra *et al.* 2001; Steinberger *et al.* 2003; Parker *et al.* 2012, 2014), however, none of

these studies examined place learning and memory. In (Williams-Simon *et al.* 2019), and as part of this dissertation, Williams-Simon *et al.* identified 16 novel QTLs that are associated with place learning and or memory and have hypnotized 9 possible candidate genes in the DSPR. This work is explained in more detail in chapter 2 of this dissertation.

Together, the field has done remarkable work in identifying genetic variants that influence learning and memory, through various approaches, however, there is still much work to be done in the elucidation of causal genes and pathways that are involved in those traits. As the approaches to science continue to evolve rapidly, there is hope that in the near future we will have a functional and mechanistic understanding of the genetic basis of learning and memory.

Thermal Tolerance – Thermal tolerance affects every living organism. Temperature regulates and plays a critical role in a wide range of processes from cellular function, like action potentials (Yu *et al.* 2012a), to more systematic evolutionary processes, like migration (Sunday *et al.* 2012). Indeed, any change in optimum temperature can cause multiple negative effects downstream both on cellular and population levels (Chen *et al.* 2011a; Huey *et al.* 2012; Sunday *et al.* 2012; Urban 2015). Thermal tolerance can be placed in two major categories – cold and hot thermal tolerance (Dhaka *et al.* 2006; Sinclair *et al.* 2012; Hoffmann *et al.* 2013; Khaliq *et al.* 2014; Lancaster and Humphreys 2020), which can be further broken down into behavioral (Hertz *et al.* 1993; Hoffmann *et al.* 2003; Dillon *et al.* 2009) and physiological (Stevenson 1985; Tosini and Menaker 1995; Hamada *et al.* 2008) thermoregulation (thermal tolerance). How well organisms

are able to adapt to extreme temperatures is vitally important for survival in a rapidly changing environment (Huey *et al.* 2012; Khaliq *et al.* 2014; Lancaster and Humphreys 2020). This is especially true as climate change continues to threaten our planet (McCain and Colwell 2011; Letcher 2015), and force the redistribution of many organisms (Walther *et al.* 2002; Parmesan and Yohe 2003; Thomas 2010; Sunday *et al.* 2012) and reviewed in (Bozinovic *et al.* 2011). Recently, Bennett *et al.* (Bennett *et al.* 2018), created a database that comprises over 2,000 species' thermal tolerance from across the world, and thus, will help improve our understanding of the effects that climate change has on thermal tolerance of these species. As climate change continues to be a major ecological issue and evolutionary problem (Huey *et al.* 2009; Harley and Paine 2009; Allen *et al.* 2010; McKechnie and Wolf 2010; Seuront *et al.* 2019)(Kjærsgaard *et al.* 2010; Denny and Dowd 2012; Kingsolver and Buckley 2015; Buckley and Huey 2016)(Huey *et al.* 2009; Harley and Paine 2009; Allen *et al.* 2010; McKechnie and Wolf 2010; Seuront *et al.* 2019), it is imperative that we have a clearer understanding about the genetic variants that influence thermal tolerance. In this dissertation, I focus on hot thermal tolerance, and the phenotype I measure is incapacitation, which I define as the time it takes before an individual becomes incapacitated when presented with high temperatures. It is important to note that there are other measures of thermal tolerance including, cold and heat hardening (Rinehart *et al.* 2000; Kellett *et al.* 2005; Huang *et al.* 2007; Norry *et al.* 2008), and cold thermal tolerance (Bozinovic *et al.* 2016; Königer and Grath 2018; Sunday *et al.* 2019). As climate change continues to be one of the leading causes of species extinction

(Thomas *et al.* 2004; Warren *et al.* 2013; Román-Palacios and Wiens 2020) reviewed in (Pimm 2009; Cahill *et al.* 2013), my work will provide fundamental insights into the individuals within a population that will be the ‘winners and losers’ (Somero 2010, 2011), and, therefore survive the extremely high temperatures that are associated with global climate change.

It is well-established that thermal tolerance has a genetic basis. Research both in non-model and model organisms has shown that multiple genetic variants influence thermal tolerance with studies ranging from coral to ectotherms, to livestock, and even plants (Krishnan *et al.* 1989; Mori 1999; Sørensen *et al.* 2001; Fanguie *et al.* 2006; Kobayashi *et al.* 2010; Li *et al.* 2011; Collier *et al.* 2012; Yu *et al.* 2012b; Ladner *et al.* 2012; Dixon *et al.* 2015; Ørsted *et al.* 2019; Chakravarti *et al.* 2020). Like other complex traits, much of what we know about the genetic basis of thermal tolerance stems from single gene mutagenesis (Kobayashi *et al.* 2010; Chen *et al.* 2011b; Finka *et al.* 2012; Díaz *et al.* 2015; Charoensuk *et al.* 2017; Cleves *et al.* 2020). Through these approaches, two of the more common gene families that are associated with thermal tolerance have been identified (i.e. Transient Receptor Potential [TRP]; (Ramsey *et al.* 2006; Venkatachalam and Montell 2007), and Heat Shock Proteins [HSP]; (Burdon 1986). In fact, HSP and TRP are evolutionarily conserved in a variety of species, including plants, coral, insects, fish, and mammals (Iwama *et al.* 1998; Gurley 2000; Zhang *et al.* 2002; Clapham 2003; Vazquez *et al.* 2004; Frank *et al.* 2009; Gao *et al.* 2014; Lockwood *et al.* 2017; Cleves *et al.* 2020), suggesting that these families of genes are essential for thermal tolerance. Of the two, the HSP family has been extensively

studied. A recent study by Cleves et al. (Cleves *et al.* 2020), induced mutations in the gene for heat-shock transcription factor and demonstrated that there is a decrease in thermal tolerance in coral (*Acropora millepora*). Another example in wheat showed that induction of HSP gene expression has a role in thermal tolerance acquisition to high temperatures (Rampino *et al.* 2009). Lastly, Lockwood et al. (Lockwood *et al.* 2017), showed that maternal loading of small hsp23 increases thermal tolerance in *D. melanogaster* embryos, suggesting epigenetics play a role in thermal tolerance as well. Collectively, HSP has been shown to play a critical role in how well a variety of species tolerate extreme temperatures.

Drosophila species have been used extensively to study the genetic basis of thermal tolerance. This is primarily because *Drosophila* is widely distributed throughout the world and is found in multiple different latitudinal temperatures (Overgaard *et al.* 2011; MacLean *et al.* 2019b). Because of this large range of habitat temperatures, the thermal tolerance variation that exists in *Drosophila* (MacLean *et al.* 2019b) makes the species particularly attractive for studying the genetic basis of thermal tolerance. This variation has been well documented in many different *Drosophila* species (Sørensen *et al.* 2001, 2019; Hoffmann *et al.* 2002; Overgaard and Sørensen 2008; Overgaard *et al.* 2011; MacLean *et al.* 2019b, 2019a). *Drosophila mojavensis* has evolved to be able to withstand 40°C (Stratman and Markow 1998), while *Drosophila subobscura* are able to fly in air temperatures that are under 10°C (Gilchrist and Huey 2004). This phenotypic variation among species suggests that these differences have evolved and that there is a genetic basis to this trait. In *D. melanogaster*, in particular, the major

candidate genes identified to influence thermal tolerance are HSP (Tissières *et al.* 1974; Burdon 1986), TRP (Ramsey *et al.* 2006), and to a lesser extent the foraging (*for*) gene (Chen *et al.* 2011b). All three of these genes were discovered through single-cell mutagenesis or gene expression analysis of specific cells.

Another approach in *D. melanogaster* used to elucidate the genetic basis of thermal tolerance is genome scans (i.e. interval QTL mapping and GWAS). First, Norry *et al.* (Norry *et al.* 2004) and Vermeulen *et al.* (Vermeulen *et al.* 2008b, 2008a), from the same research group, selected for high vs. low resistance to thermal tolerance and knockdown, and identified QTLs that influence thermal tolerance between two geographically distinct populations of *D. melanogaster*. Second, Morgan and Mackay (Morgan and Mackay 2006) used two inbred populations divergent for thermal tolerance and identified QTLs that influence low and high-temperature resistance. Again in (Norry *et al.* 2007a, 2007b, 2008), (Norry *et al.* 2007a, 2007b, 2008), and (Norry *et al.* 2007a, 2007b, 2008), Norry, *et al.* identified QTLs for different parameters of thermal tolerance, including heat and chill-coma recovery, and identified chromosomal regions influencing those traits. Furthermore, Rand *et al.* (Rand *et al.* 2010) identified QTLs through microsatellite mapping in laboratory selected and natural populations of *D. melanogaster*. Finally, Rolandi *et al.* (Rolandi *et al.* 2018) used the DGRP, to identify QTLs that influence high thermal tolerance. All of these genome scans made strides towards establishing loci that influence thermal tolerance, however, there was very little overlap in the identified QTLs, highlighting the complexity of this trait.

Collectively, both the single gene and genome scan approaches identified candidate genetic variants that influence thermal tolerance, however, the natural variants that influence thermal tolerance are still unclear. In this dissertation, I take an integrative approach to identify causal genes that influence thermal tolerance in *D. melanogaster*

1.3 Quantitative Genetics – Phenotypes to Genotypes

Forward genetics is a powerful tool that has been used to elucidate the genetic basis of complex traits (including learning, memory, and thermal tolerance). This technique uses top-down approaches, i.e. from phenotype to genotype. This particular field of study is known as quantitative genetics (Falconer and Mackay 1983; Falconer 1996) and has been extensively used in animal breeding (Turner and Young 1969; Schrooten *et al.* 2000; Pampouille *et al.* 2018), crop production (VanToai *et al.* 2001; Hallauer *et al.* 2010; Kang 2020), our understanding of evolutionary processes (Kruuk *et al.* 2008) reviewed in (Kruuk *et al.* 2008), and human diseases (Jostins *et al.* 2012; Locke *et al.* 2015; Wray *et al.* 2018). Broadly, the field of quantitative genetics aims to measure the contribution and how the combination of genotype and environment influences a particular phenotype (**Fig. 1.1**). This process usually starts with the measurement of the phenotype of interest then works back to the genes and or environment that are contributing. To identify specific genetic variants using this approach, ideally, one would phenotype and genotype a huge population to identify genetic variants associated with that particular phenotype (Lynch *et al.* 1998). There are a few

guiding principles of quantitative genetics, which include heritability (H^2) and the genetic covariance between traits (Roff 1996), and genetic mapping (linkage and association; (Lander and Botstein 1989; Haley and Knott 1992). Heritability is the proportion of genetic variation of a trait that is attributed to the next generation and is usually additive, whereas genetic covariance determines how traits will change in correlation to each other (Roff and Mousseau 1987; Mousseau and Roff 1987; Roff 1996). I elaborate on two types of genetic mapping in the paragraph below. In this dissertation, I focus on those guiding principles as they relate to learning, memory, and thermal tolerance in the DSPR.

To identify the genetic variants that are associated with complex traits from a top-down approach, researchers typically use genomic scans, which includes Quantitative Trait Loci (QTL) mapping (Falconer 1996; Lynch *et al.* 1998; Kearsey 1998) or genome-wide association study (GWAS; (Hirschhorn and Daly 2005), which was pioneered in human genetics. These approaches allow for genome-wide inference or association of the relationship between genotype at various genomic locations and phenotype of interest, using phenotypic data and genomic markers, like microsatellites (Hamada *et al.* 1982; Senan *et al.* 2014) and Single Nucleotide Polymorphisms (SNP; (Morin *et al.* 2004). The main goal of genome-wide scans is to localize chromosomal regions that significantly affect the variation of a particular phenotype within a population (Fox and Wolf 2006), and therefore the ability to identify genes that contribute to the variation in the phenotype. Genomic scans typically cannot independently identify single candidate gene variants and are often paired with fine-mapping approaches, like RNA-seq (Highfill

et al. 2017; Yang *et al.* 2019; Binenbaum *et al.* 2020) and candidate gene approaches (Byrne and McMullen 1996; Hu *et al.* 2008). However, genomic scans have several other benefits, which includes, high power to detect large effect QTLs and high mapping resolution (Carlson *et al.* 2004; Winkler *et al.* 2010; Shriner *et al.* 2011; King *et al.* 2012a; Crawford and Nielsen 2013), but perhaps the most significant benefit is the ability to systematically search the entire genome in an unbiased manner to identify chromosomal regions that influence a particular trait.

Despite the success of quantitative genetics, there are a few major drawbacks. **1)** QTLs mapping can only detect allelic diversity that segregates in the parents of RILs or F₂ generation – in the case of two-way QTL mapping – (Mackay 2001), **2)** the resolution to detect QTLs is depend on the amount of recombination, although that problem can be resolved easily by allowing RILs to recombine more (King *et al.* 2012a), **3)** Linking phenotypic variation to small effect QTLs (Maher 2008). **4)** Validation of candidate genes has been challenging because single gene deletion and knockouts were not a simple task in natural populations, because natural variation is often distributed continuously (Falconer 1996), making it difficult to validate single variants. Validation methods, like quantitative complementation (Mackay 2001; Turner 2014) and the reciprocal hemizyosity test (Stern 2014) – which takes advantage of the CRISPR/Cas9 system – make this challenge less daunting. Several empirical studies have shown the feasibility of identifying specific genes following mapping and validation approaches, both in *D. melanogaster* and other species (Marriage *et al.* 2014; Ding *et al.* 2016; Zhang *et al.* 2017; Highfill *et al.* 2017; Ng'oma *et al.* 2020). Other

limitations of quantitative genetics include small sample sizes (Beavis and Paterson 1998), the 'missing heritability' (Manolio *et al.* 2009), and the Beavis effect phenomenon (Beavis *et al.* 1991; Beavis 1994; Beavis and Paterson 1998; Xu 2003; King and Long 2017). The missing heritability effect is directly related to GWAS approaches where several variants of large effect only explain a small proportion of heritability. In other words, the portion of phenotypic variance that is seen in a particular trait within a population is attributed to additive genetic factors (Hindorff *et al.* 2009). A classic example of the missing heritability is human height, which is a complex trait with 80% estimated heritability, however, one of the first studies only explains about 5% of the phenotypic variance (Visscher 2008). A later study by the same group drastically increased the sample size and was able to map hundreds of loci, which increased the estimates to explain upwards of 20% of the phenotypic variance (Wood *et al.* 2014), however still 'missing' information. One proposed cause of this phenomenon is low power and mapping resolution for variants of modest effect (Risch and Merikangas 1996; Risch 2000; Altmüller *et al.* 2001). A potential fix to this problem would simply be to increase the sample size (Manolio *et al.* 2009). The Beavis effect states that average identified QTLs that are associated with a phenotype are typically overestimated. The Beavis experiments simulated independent QTLs, where each QTL explained a range of 0.75 to 9.5% of the proportion of the phenotypic variance with sample sizes ranging from 100 to 1000. If the sample size was 100, then the QTL was overestimated, slightly overestimated if the sample size was 500, and close to actual when the sample size was 1,000 (Beavis and Paterson 1998). King and Long (King and Long

2017) calculated the Beavis effect both in the DSPR and in the DGRP, which are two mapping populations in *D. melanogaster*, and found that overestimation of QTLs is determined most strongly by sample size and is only minimally impacted by the mapping design. Again, suggesting the importance of larger sample sizes.

Quantitative genetics have been fundamental in our understanding of many complex traits, despite the many drawbacks. As genome-wide technologies continue to rapidly improve, they become cheaper and more accessible, therefore I am hopeful that the future of field quantitative genetics will be bright.

1.4 Combing Approaches – Whole Genome Scans, Fine-Mapping, and Single Gene Mutagenesis

Identifying the genetic basis of complex traits has been a longstanding grand challenge in Biology. This problem has been approached from many different perspectives, but the potential synergy that could result from an integrative approach (**Fig. 1.2**) is rarely realized. Much of our knowledge of the genes that influence complex traits have been uncovered using classical genetics approaches (Dudai *et al.* 1976; Zars *et al.* 2000b; Cleves *et al.* 2020). The sequencing of the *D. melanogaster* genome (Adams *et al.* 2000) and the introduction of RNA interference (RNAi) reviewed in Kuwabara and Coulson (Kuwabara and Coulson 2000), and CRISPR/Cas9 genome modification (Jinek *et al.* 2012), have made functional validation of candidate genes highly feasible. Single gene approaches are valuable and have made an impactful contribution to what we know about complex traits, however, this approach is limited in throughput

(Zhu and Zhao 2007). Conversely, mapping approaches identify genome-wide associations between genotypes and phenotypes but typically lack sufficient resolution to identify single genes (Mackay 2001; King *et al.* 2012a), and is therefore often paired with identification and validation approaches like RNA-seq to identify gene variants (Highfill *et al.* 2017; Williams-Simon *et al.* 2019; Ng'oma *et al.* 2020; Binenbaum *et al.* 2020). Other approaches include identification of structural variants (Frazer *et al.* 2009; Eichler *et al.* 2010) and evolve and resequence (Miles *et al.* 2011; Turner and Miller 2012; Burke *et al.* 2014b; Jha *et al.* 2015) approaches. SVs, which are broadly defined as all variants that are not single nucleotide variants, include insertion-deletions, block substitutions, inversions of DNA sequences, and copy number differences (Frazer *et al.* 2009). SVs are thought to be major contributors to human genetic variation (Sebat *et al.* 2004; Redon *et al.* 2006; Human Genome Structural Variation Working Group *et al.* 2007) and other in complex traits (Frazer *et al.* 2009; Eichler *et al.* 2010; Chakraborty *et al.* 2018), and, therefore are important factors to consider when dissecting the genetic basis of complex traits. In (Chakraborty *et al.* 2018, 2019), Chakraborty *et al.* characterized the SVs in *D. melanogaster* and found that there are more than 20,000 SVs present in the fruit fly genome of which ~40% are invisible to high specificity short-read genotyping approaches, such as next-generation sequencing. Furthermore, they functionally validated their findings in the DSPR and found that SVs are overrepresented in candidate genes associated with QTLs in the DSPR, indicating that uncovering SVs that are associated with

candidate genes within QTLs might be critically important to identify causal variants that are associated with complex traits.

1.5 Conclusion

In this dissertation, I take a multipronged approach in an integrated research design, leveraging the strengths and compensating for the weaknesses of each approach, allowing for the comprehensive discovery of the genes that influence learning, memory, and thermal tolerance using fruit flies. Overall, I **hypothesized** that there are many large effect genomic loci that regulate these traits, there is a shared genetic basis between learning and memory and there is a trade-off between learning and thermal tolerance. The **overarching goal** of this dissertation is to take a multifaceted approach to dissect the natural genetic architecture of learning, memory, and thermal tolerance in *Drosophila melanogaster*. This research is broken up into three specific **aims**: **1)** Identify the major genetic loci and candidate genes that are influencing learning and memory through genome scans, QTL mapping, and RNA-seq, **2)** Identify the major genetic loci and candidate genes that are influencing thermal tolerance through genome scans, QTL mapping, and RNA-seq, **3)** functionally validate candidate genes that are important for learning, memory, and thermal tolerance leveraging the RNAi system and other computational pipelines. To date, I have identified the QTLs and candidate genes that influence all three traits and validated the genes that influence thermal tolerance. Collectively, these novel findings have bridged the gap in our understanding of the genetic basis of learning, memory, and thermal

tolerance and could serve as the building blocks to better understand the wide range of learning, memory, and thermal regulation diseases that affect humans and other mammals, including climate change

1.6 Figures/ Schematics:

Fig 1.1:

Schematic explaining interval QTL mapping. Bars represent the genome and memory scores are plotted in box plots on the right. This schematic is modified from King lab and Broman and Sen (Broman and Sen 2009).

Fig 1.2:

Schematic showing different integrative approaches that I will be using to genetically dissect learning, memory, and thermal tolerance.

Chp 2 – THE GENETIC BASIS OF LEARNING AND MEMORY IN A MULTI-PARENTAL POPULATION OF FRUIT FLIES

Author Affiliation: Williams-Simon P.A.¹, Posey C.¹, Mitchell S.¹, Ng'oma E.¹, Mrkvicka, J.A.^{1,2}, Zars T.¹, and King E.G.¹

¹Division of Biological Sciences, University of Missouri, 401 Tucker Hall, Columbia, Missouri 65211, USA

²Current address: Interdisciplinary Graduate Program in Genetics, Graduate College, University of Iowa, Iowa City, IA 52242, USA

This chapter and the content thereof have been published by Genes, Brains, and Behaviour. Permission was requested and was granted to publish this article for dissertation/ thesis type. License number 5100321136978.

2.1 Abstract:

Learning and memory are critical functions for all animals, giving individuals the ability to respond effectively to changes in their environment. Within populations, individuals vary in performance, however, the mechanisms underlying this variation in performance are largely unknown. To uncover the genetic basis for variation in learning and memory performance, I used the DSPR, a multiparent mapping resource in the model system *Drosophila melanogaster*, consisting of a large set of recombinant inbred lines (RILs) that naturally vary in these and other traits. Fruit flies can be trained in a “heat box” to learn to remain on one side of a chamber (place learning) and can remember this (place memory) over short timescales. Using this paradigm, I measured place learning and memory for ~49,000 individual flies from over 700 DSPR RILs. I identified 16 different loci across the genome that significantly affect place learning and/or memory performance, with 5 of these loci affecting both traits. To identify transcriptomic differences associated with performance, I performed RNA-Seq on pooled samples of 7 high performing and 7 low performing RILs for both learning and memory and identified hundreds of genes with differences in expression in the two sets. Integrating our transcriptomic results with the mapping results allowed us to identify several promising candidate genes, advancing our understanding of the genetic basis underlying natural variation in learning and memory performance.

These results are a critical first step towards characterizing the genetic differences that determine an individual's learning and memory abilities.

2.2 Introduction:

The ability to learn and remember is critically important, allowing individuals to adjust their behavior in response to stimuli to cope with changing environments. Learning and memory performance varies widely in different contexts across the animal kingdom, with some species having evolved higher learning and/or memory abilities depending on the selective pressures they have been exposed to (Dukas 2008a; Pearce 2008a; Snell-Rood *et al.* 2011a; Mery 2013a; Morand-Ferron and Quinn 2015a). Behavioral experiments in butterflies (Goulson and Cory 1993; Snell-Rood and Papaj 2009), chickadees (Pravosudov *et al.* 2003; Tello-Ramos *et al.* 2018), moths (Cunningham *et al.* 2004; Ito *et al.* 2008), honey bees (Brandes 1988; Hammer and Menzel 1995; Sandoz and Menzel 2001), fruit flies (Quinn *et al.* 1974; Yurkovic *et al.* 2006), rodents (Tang *et al.* 1999; Peleg *et al.* 2010; Shumake *et al.* 2014), and humans (Blokland *et al.* 2011a) have shown collectively that there is genetically based variation in these traits and have established the different classes of learning and memory within species. This variation provides the raw material for natural selection to act on when higher or lower learning and memory performance is selected for in a population. Experimental evolution studies have empirically demonstrated that populations can evolve higher or lower learning performance (McGuire and Hirsch 1977a; Mery and Kawecki 2002a; Mery

et al. 2007a). These results show that there are natural genetic variants that give rise to differences in performance in learning and memory. The identity of the vast majority of these variants in nearly all populations, however, remains unknown (for exceptions see: (Mery *et al.* 2007a; Papassotiropoulos *et al.* 2011).

Many empirical studies aim to understand the mechanistic basis of different classes of learning and memory within the broad context of animal behavior, using both classical and operant learning and memory paradigms. Specifically, many classical molecular genetic studies have identified individual genes involved in learning, and memory in many different animal models, such as *Caenorhabditis elegans* (*C. elegans*; (Gomez *et al.* 2001; Amano and Maruyama 2011; Lau *et al.* 2013), *Mus musculus* (Silva *et al.* 1992a, 1992b; Tonegawa *et al.* 2003), and *Drosophila melanogaster* (*D. melanogaster*; (Dudai *et al.* 1976; Levin *et al.* 1992; Zars *et al.* 2000b; LaFerriere *et al.* 2011). For example, in *C. elegans*, mutants with a defective *nmr-1* encoding an N-methyl-D-aspartate (NMDA) receptor subunit fail to form both short-term and long-term memories (Amano and Maruyama 2011). In fruit flies, mutations in the *rutabaga* (*rut*) gene (Levin *et al.* 1992; Zars *et al.* 2000b) and in the *dunce* (*dnc*) gene (Dudai *et al.* 1976; Widmann *et al.* 2016) have both been shown to influence multiple classes of learning and memory. These studies, along with follow-up studies showing the mechanism of action of these genes (Qiu and Davis 1993; Ueda and Wu 2012), have provided crucial information about which genes are required to function for proper learning and memory.

Despite this success, it is not clear whether segregating variants within these same genes are causing the variation observed in performance among

individuals in natural populations. It is possible that the genes identified via these approaches are so central to the processes of learning and memory that their function is highly conserved, and most of the individual-level variation in learning and memory is due to genetic variants in other genes (Mery 2013a). Another reason to expect that the variants identified via mutant studies might not correspond to those seen in natural populations is that many of these mutants show deleterious pleiotropic effects. For example, severe *dnc* mutant flies have female sterility and partial lethality (Byers *et al.* 1981; Ueda and Wu 2012), phenotypes that would be strongly selected against in the wild. Thus, identifying the natural genetic variants underlying why some individuals in a population perform better or worse than others is critical to our understanding of the mechanistic basis of learning and memory.

Given the complex processes that govern learning and memory, which certainly involve a large number of potentially interacting genes, single-gene approaches do not necessarily capture what happens on a systematic or organismal level (Mery 2013b). Mapping studies in natural populations, such as quantitative trait loci (QTL) studies and genome-wide association studies (GWAS), have the potential to allow for the identification of multiple genetic variants influencing a complex trait in the real context in which the variants occur. However, genetic mapping studies have been challenging to perform for learning and memory phenotypes. Assaying learning and memory is often labor-intensive, requiring repeated behavioral trials (Pearce 2008b), making it difficult to assay the large numbers of individuals typically required for high power to detect all but the

highest effect QTL (Broman and Sen 2009; King *et al.* 2012a; Keele *et al.* 2018). In human populations, while learning, memory, and general cognitive function are known to be heritable (Blokland *et al.* 2011b; Mery 2013b), they are also influenced strongly by a suite of environmental factors, which obviously cannot be systematically controlled in human studies. High levels of variation in those environmental factors will negatively affect the power to detect a genetic association (Broman and Sen 2009; King *et al.* 2012a; Keele *et al.* 2018). There have been a few successes where a causative natural variant for learning or memory has been identified. These include a GWAS in humans of short-term working memory that identified a polymorphism in the *SCN1A* gene, a voltage-gated sodium channel (Papassotiropoulos *et al.* 2009), a study in fruit flies linking the *foraging* gene, a cGMP-dependent protein kinase, to associative olfactory learning and memory (Mery *et al.* 2007b), and a QTL mapping study in mice implicating the *Hcn1* gene, a hyperpolarization-activated cyclic nucleotide-gated channel, in fear conditioning (Knoll *et al.* 2016). However, by and large, very few QTL for learning and memory have been discovered (but see (Chandra *et al.* 2001; Steinberger *et al.* 2003; Parker *et al.* 2012, 2014), and even fewer specific causative variants (aside from the few exceptions noted above) have been identified in any system thus far.

Multiparental populations, consisting of a large number of recombinant inbred lines (RILs) generated from multiple inbred founder lines crossed for multiple generations, have the potential to allow for high-resolution, highly powered genetic mapping studies. The *Drosophila* Synthetic Population Resource (DSPR)

is one such multi-parental population, consisting of two sets of ~800 RILs, each generated from an 8-way, 50-generation cross (King *et al.* 2012a, 2012b; Long *et al.* 2014). This system, with a large number of RILs and a long period of crossing allows for highly powered QTL mapping to an interval that typically includes tens of genes rather than the wide intervals including hundreds of genes that are typical of more traditional 2-line QTL studies. In addition, in the *D. melanogaster* system, a specialized piece of equipment, the “heat box”, originally developed by Wustmann *et al.* (1996), allows for a high-throughput assay of place learning and memory (Zars *et al.* 2000c; Sitaraman *et al.* 2008). Place learning is a type of operant learning that occurs when a fly learns to associate a specific place with a consequence. In the heat box, up to 16 individuals can be assayed concurrently for place learning and memory in less than 10 minutes, making it feasible to assay the large number of individuals necessary for QTL mapping for these key phenotypes.

In this study, I use the DSPR to genetically dissect place learning and memory. After assaying 48,940 individual flies from 741 RILs from the DSPR, I mapped 16 QTL, with 5 shared QTL between learning and memory. In addition, I performed RNA-Seq to measure differential gene expression of high versus low performing cohorts of RILs and used these data to narrow the set of candidate genes within our QTL intervals. The loci I identify have not been previously associated with place learning or memory, representing a step forward in understanding the genetic basis of these traits.

2.3 Methods:

Mapping population – I used a multi-parental population of *D. melanogaster*, the DSPR (King *et al.* 2012a, 2012b; Long *et al.* 2014) (<http://FlyRILs.org>). The DSPR consists of two sets of approximately 800 RILs (genotype), each created via an 8-way advanced intercross design. To create these RILs, eight inbred founder lines were crossed and allowed to randomly mate for 50 generations followed by 25 generations of inbreeding. This creates a set of recombinant inbred lines whose genomes are mosaics of the eight original founder lines. The founder lines have been fully re-sequenced, and the RILs have been genotyped at over 10,000 SNPs. King *et al.* (2012a) developed a hidden Markov model that infers the likely founder ancestry at all genomic locations for all RILs, providing full genome information for all RILs. King *et al.* (2012a, 2012b) give the complete details of the formation of the DSPR and associated resources. In this experiment, I used nearly all of the lines in one of the two sets of RILs, the population A RILs, which consisted of 741 RILs.

Fly husbandry – All stock flies were raised on the standard cornmeal-yeast diet, and kept at 18° C, with 60% relative humidity, in wide vials (Fisher Scientific, Cat. No.: AS-513). Two weeks before the behavioral assays, about 6 males and 10 female flies were flipped onto a new flask (Fisher Scientific, Cat. No.:AS-355) of food and allowed to mate at 25°C and lay eggs. Seven days post-oviposition, all the adults were removed from the flask to ensure that only the F1 generation remained in the flask. Fourteen days post-oviposition, adult flies were anesthetized on ice, and I collected 60-80 female flies for the behavioral assay. These female

flies were placed onto two separate new food vials in groups of 30-40 individuals per vial. The flies were then allowed to recover for at least 24 hrs before the assay. Flies included in the assay ranged in age from 1d. For each RIL I phenotyped a minimum of 40 individuals, with an average of 66, and a range of 40-117. A total of 10-15 RILs were set up and phenotyped in any given week.

Phenotyping – Through a behavioral assay known as place learning, we are able to train flies with a highly sensitive apparatus, the “heat box” (Wustmann *et al.* 1996; Putz and Heisenberg 2002; Kahsai and Zars 2011). Single flies are placed in individual chambers that are lined top and bottom with Peltier elements, which allow for fast temperature changes within a chamber. The cooling (24°C) and warming (41°C) of the chambers – which takes about 4 seconds to change from one temperature to the other – are entirely controlled by the position of the fly. If the fly is positioned on the cool-associated side (24°C), then the entire chamber reflects that temperature, the same is true for the hot-associated side (41°C). There is an infrared light within each chamber that tracks the position of the flies during the assay. Because 41°C is a highly aversive stimulus, a typical fly will learn to remain on the cool-associated side of the chamber. In addition, most flies will remember to remain on the cool-associated side of the chamber once the punishment is removed (i.e., the chamber no longer heats when the fly enters the hot-associated side).

Using the above paradigm, I assayed learning and memory in a high throughput way. When all chambers are functioning, a total of 16 flies can be tested within the heat box concurrently, with both learning and memory assays performed

consecutively on a single individual fly. I define learning as the increasing avoidance of the hot-associated side, and memory as the persistent avoidance of the hot-associated side, even if the aversive stimulus no longer exists (Dukas 2008b; Pearce 2008b; Snell-Rood *et al.* 2011b; Mery 2013b; Morand-Ferron and Quinn 2015b). To quantify learning and memory, flies are given a total of 9.5 mins within the chamber: 30 secs for pre-testing, 6 mins for learning, and 3 mins memory testing (Ostrowski *et al.* 2015). During these testing periods, performance indices (PIs) were calculated with specialized software (Heat Calc 2.14) (Sitaraman *et al.* 2008; LaFerriere *et al.* 2011), which reflect the relative position preference of an individual fly. I calculated PIs for both phenotypes, using the amount of time an individual spends on one side of the chamber divided by the total amount of time of the assay (**Figure 1**): The maximum PI is 1, which corresponds to a fly that has perfect avoidance of the hot-associated side. A PI of zero indicates a preference for neither side, and a PI of -1 indicates a fly that remains on the hot-associated side for the entire assay.

Several quality control filters allowed us to remove any individual assays with problems and ensure the validity of all measurements. These analyses and all following were performed in (R Core Team 2018). First, I eliminated any individuals that did not cross the midline of the chamber at least once, to ensure all individuals experienced the aversive stimulus. Second, the temperature of each chamber is tracked over the course of the assay, and I used these data to identify problematic chambers (i.e., any chamber that was not warming or cooling to the correct temperature). Most often, these were chambers that were not reaching the

target temperature of 41°C. I flagged any chamber as problematic if the average difference between the actual temperature within the chamber and the target temperature exceeded 5°C during the assay. Finally, I eliminated any individual that showed a long period of inactivity at the end of the assay to remove any individuals that were potentially immobilized for any reason, which would skew their learning or memory PI.

Heritabilities and QTL mapping – I estimated the broad-sense heritability of learning and memory and the genetic correlation between these phenotypes by estimating the genetic and phenotypic variance and covariance components from a linear mixed model using the `lme` and `VarCorr` functions in the `nlme` package (Pinheiro *et al.* 2013), followed by a jackknife, which allowed us to obtain standard errors of our estimate (Roff and Preziosi 1994; Roff 2008). Prior to fitting this model, both learning and memory were quantile normalized to ensure normality. This transformation was particularly important for the individual-level data because, with the PIs bounded by -1 and 1, the data are highly non-normal. The jackknife removes each observation once, the model is then fit and the quantitative genetic parameter (e.g., heritability or the genetic correlation) is estimated, and a pseudo value is calculated. Thus, for our dataset of 741 RILs, each RIL is deleted, one at a time, to produce 741 pseudo values. Because the `lme` function uses restricted maximum likelihood, it is possible for the model to not converge. This occurred for 28 cases for the heritability of learning, 50 cases for the heritability of memory, and 77 cases for the genetic correlation, thus, these pseudo values were not included in the calculation of our estimates. Our estimate did not differ

substantially from the estimate obtained from fitting the mixed model alone without using the jackknife. I also performed a model comparison using a likelihood ratio test to determine the significance of RIL (i.e., genotype). We compared the model fitting only the intercept to the model including RIL, fitting both via maximum likelihood such that model comparisons would be valid.

For each of the 741 RILs we measured, we took the average PI for learning and memory from each RIL and applied Haley-Knott regression (Haley and Knott 1992). This statistical approach regresses each phenotype on the eight founder haplotype probabilities (Broman and Sen 2009; King *et al.* 2012a) by fitting the following model:

$$y_i = \sum_j^7 p_{ij}b_{ij} + e_i,$$

where y_i is the phenotype (i.e., learning or memory) of the i th RIL, p_{ij} is the probability that the i th RIL has the j th haplotype at the locus, b_{ij} is the vector of effects for the j th haplotype, and e_i is the vector of residuals. I fit this model at ~10,000 regularly spaced positions across the genome. Prior to performing this genome scan, I transformed the learning phenotypic data, using a power transformation, raising our average learning values to the third power, which improved the normality of the data (**Figure X**). To identify statistically significant QTL, I performed 1000 permutations (Churchill and Doerge 1994) of the dataset and performed a genome scan on each permuted dataset as above. The same individuals were assayed for learning, memory, and thermal tolerance. I do not report the results of thermal tolerance here; however, because all three

phenotypes are part of the same experiment, I permuted these three phenotypes together and performed genome scans at once. I then determined the number of false-positive QTL at different significance thresholds. I identified individual, distinct QTL by removing any QTL that was within 2 cM of a more significant QTL. I calculated the empirical false discovery rate (FDR, expected false positives/total positives) for each threshold and determined the threshold corresponding to a 5% FDR. I also calculated the threshold corresponding to a 5% family-wise error rate (FWER, the probability of one or more false positives experiment wide) by determining the lowest p-value for each set of genome scans from each permutation and calculated the 95% quantile of the resulting set of p-values (Churchill and Doerge 1994). For each peak, I calculated the estimated effects of each haplotype at each QTL, the percent variance explained by the QTL (PVE), and the Bayesian Credible interval (BCI), following the methods described previously for the DSPR (King *et al.* 2012a). I considered QTL as separate QTL if their BCIs did not overlap. One memory QTL, Q15 is near a wider learning QTL, Q14, and thus it is possible Q15 is a shared, pleiotropic QTL. However, within learning, the peak at Q14 and a smaller peak near Q15 cannot be distinguished as their BCIs overlap. Additionally, while Q14 and Q15 are near one another, their BCIs do not overlap. Thus, I decided to be conservative with respect to concluding Q15 was a shared QTL and consider it a QTL for memory only.

Rna-sequencing and Analysis – I selected 14 RILs from the high ($n = 7$) and low ($n = 7$) 5% for learning and memory of the first 140 RILs I phenotyped. Since the maximal and minimal PIs are bounded by 1 and -1, I felt confident the

RILs in the top and bottom cohorts would still have higher or lower than average PIs after assaying all RILs. Obviously, the ranking of these RILs changed once I assayed an additional 601 RILs, however, they do still fall within the high and low cohorts. I collected 25 female heads from each of the 7 RILs in the high and low cohorts. Five heads from each RIL within a cohort were pooled into 5 biological replicates, totaling 35 heads per tube. To optimally identify consistent differences in expression between high and low cohorts, I chose to pool RILs together to average over biological differences and reduce the number of samples for RNA-Seq. I also included 5 heads per RIL to account for individual-level variability.

To prepare the samples, first, the 7 selected RILs underwent the same set-up protocol as done in the phenotypic assay (see phenotyping section above for details). Briefly, female flies (fourteen days post-oviposition) were selected by anesthetizing them on ice between 0800 and 1000 and allowed to recover for 24 hrs at 25°C. Immediately after, the flies were flash-frozen in liquid nitrogen. Second, the samples were vortexed for no more than 10 seconds to remove heads from bodies and stored at -80°C. Third, the samples were freeze-dried overnight to prevent degradation using a lyophilizer (Labconco, Cat No.: 77550-00). Both the high and low cohorts were processed in the same batch for each phenotype.

The samples were then shipped to RaPiD Genomics, LLC (fee for service) where the total RNA was extracted with Dynabeads mRNA direct kit from life technologies, mRNA was then fragmented and converted into double-stranded cDNA, followed by the standard proprietary library prep for one lane of Illumina

HiSeq 3000 instrument to generate paired-end (PE) reads. The first 16 samples were PE 150-bp, while the remaining samples were PE 100-bp.

Before aligning the reads, I trimmed the samples using the software *cutadapt* (Martin 2011), then ran a quality control test using *fastqc* (Andrews 2010). The *fastqc* analysis did not reveal any major problems with the reads. I specifically assessed the summaries for per base sequence qualities and the per sequence quality scores. I aligned reads to the *D. melanogaster* reference transcriptome (Ensembl Release Version: 84) using HISAT2 (Kim *et al.* 2015) and assembled and quantified transcripts using StringTie (Pertea *et al.* 2015, 2016). I did not allow the assembly of novel transcripts not present in the *D. melanogaster* reference. I used the "prepDE" python script available from the StringTie manual to calculate the read counts for each gene. To assess differentially expressed genes between our high and low cohorts, I used the DESeq2 package (Love *et al.* 2014). First, I filtered all genes that had zero counts in all samples. I then performed surrogate variables analysis (SVA) on normalized counts using the *sva* package (Leek *et al.* 2012, 2019) to estimate for unknown batch effects while accounting for treatment. The SVA identified one surrogate variable, which I included as a covariate when testing for differentially expressed genes. To identify significantly differentially expressed genes I used the DESeq2 package (Love *et al.* 2014) to test for overall treatment effects. I then performed contrasts between the high and low cohorts of learning and the high and low cohorts of memory to identify significantly differentially expressed genes. To visualize and rank the genes I used the function *lfcShrink*, which performs shrinkage on $\log_2(\text{Fold Changes})$, which have been

shown to produce better estimates. All $\log_2$ (Fold Changes) reported here are the shrinkage estimated values using the “normal” estimator.

To identify possible candidate genes within our QTL intervals, we first determined which of the genes within each interval of interest were significantly differentially expressed between high and low-performing cohorts. For QTL mapped to a single phenotype, this interval was the BCI. For QTL mapped to both phenotypes, we considered only the overlap region between the BCI's for both learning and memory. In addition, we integrated a previous genome-wide eQTL dataset using the DSPR. King et al. (2014) performed eQTL mapping on ~600 RIL crosses, crossing the population A and population B RILs and measuring the expression of nearly all genes in the *D. melanogaster* genome using microarrays. We used this dataset to identify which genes show evidence for a *cis* eQTL in the DSPR. Finally, I identified differentially expressed genes with previous evidence implicating a potential role in learning and memory by manually examining the annotation for these genes in FlyBase (Thurmond *et al.* 2019) and noting annotations such as neurological process, behavior, or neurotransmitter, which might implicate the gene could be involved in the process of learning or memory.

Data Availability – Raw learning and memory phenotypic data are available from zenodo (<https://zenodo.org/>): <http://doi.org/10.5281/zenodo.2595557>. RNA-Seq data is available from the NCBI Short Read Archive (Leinonen *et al.* 2011) under BioProject accession number: PRJNA527143 and SRA accession number: SRP188472 <https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?study=SRP188472>. Raw re-

sequencing data of the founder lines is deposited in the NCBI SRA under accession number SRA051316, and the RIL RAD genotyping data are available under accession number SRA051306. Founder genotype assignments from the hidden Markov model are available as two data packages in R (<http://FlyRILs.org/>) and are available from the Dryad Digital Repository (<http://dx.doi.org/10.5061/dryad.r5v40>). All data associated with the DSPR eQTL dataset is also available publicly (see (King *et al.* 2014) for experimental details) in NCBI's Gene Expression Omnibus (Barrett *et al.* 2013) and is accessible through GEO Series accession number GSE52076. All of the above DSPR data are also available centrally at <http://FlyRILs.org>. The complete code used to perform all analyses is available at GitHub (https://github.com/EGKingLab/LearnMem_DSPR).

2.4 Results:

Phenotypic Patterns – I measured place learning and memory in 741 RILs from the DSPR (**Figure 1**) and showed these lines vary widely in both these traits (**Figure 2**). For each line, I measured a minimum of 30 females, with an average of 53 per line. The lines range in average learning performance index (PI) from 0.215 to 0.933 from the lowest-performing RIL to the highest, while average memory PIs range from -0.016 to 0.874 (**Figure 2**). The estimated broad-sense heritability are moderate (Learning: $H^2 = 0.20$, 95% CI = 0.18–0.22 ; Memory: $H^2 = 0.10$, 95% CI = 0.09–0.12), and the effect of RIL (i.e., genotype) is highly significant (Learning = 6747, $p < 0.0001$; Memory = 2585, $p < 0.0001$),

demonstrating a genetic basis for these traits. Learning and memory are genetically correlated ($r_g = 0.53$, 95% CI = 0.47–0.59; Figure 2), indicating that some of the same loci influence both learning and memory. This relationship would be expected given memory formation cannot occur without learning, though high learning performance does not necessarily guarantee high memory performance. Indeed, while most RILs with high learning PIs also have high memory PIs, some RILs show high learning but poor memory, indicating some independence between these processes as well.

Genomic Scans (QTL) – To identify loci affecting place learning and memory, I performed a genome scan following established methods for mapping QTL in the DSPR (King *et al.* 2012a, 2012b). To establish both a 5% family-wise significance threshold (FWER) and a 5% false-discovery rate threshold (FDR), I performed permutations (see Methods for details). I identified 16 QTL (Q1 - Q16) at a 5% FDR: 9 are unique to the learning phenotype, 2 are unique to the memory phenotype, and 5 were mapped for both phenotypes (**Figure 3; Table 1**). Of these, 10 QTL are also significant at the FWER threshold (**Table 1**). For each QTL, I obtained a 95% Bayesian Credible Interval (BCI; (Manichaikul *et al.* 2006; Broman and Sen 2009) to define the region in which the causative variants are expected. With the exception of Q12 for learning, which spans the centromere, these intervals are narrow, averaging 643 kb in physical distance and 2 cM in genetic distance. All of the QTL I identified are of moderate effect, with the percentage of phenotypic variance explained by a locus averaging 4% and ranging from 3% to 7% (**Table 1**). In the DSPR, we are able to assign the likely haplotype identity to

every genomic segment in every RIL (King *et al.* 2012a). Therefore, for our mapped QTL, I can estimate the effect of harboring each of the 8 founder haplotypes at the QTL location on our phenotypes of interest. I show these estimates, along with the haplotype assignments for each individual RIL measurement for all shared QTL in **Figure 4**. As has been found previously in the DSPR (King *et al.* 2014; Najarro *et al.* 2015) and in other mapping panels (Cubillos *et al.* 2013; Giraud *et al.* 2014), these effect estimates show evidence for multiple causative variants resulting in an allelic series, not a single causative biallelic SNP that would show haplotype effects in defined “high” and “low” groups. The similar patterns in our effect estimates for learning and memory for a single shared QTL also support the hypothesis that these are pleiotropic loci, rather than separate linked QTL, which would be expected to show contrasting patterns in the two phenotypes.

Gene Expression – We performed RNA-Seq on pooled sets of female heads to identify differentially expressed genes in cohorts of high or low-performing RILs for place learning and memory. For each, I selected 7 high performing and 7 low performing RILs and pooled a total of 35 female heads in a single sample, with 5 heads from each RIL contributing to each sample (**Figure 2**). I then performed RNA-Seq on 5 biological replicates of these pooled samples (see Methods for details). This design allowed us to account for an individual- and RIL-level variability while keeping the number of RNA-Seq samples reasonable. The resulting dataset provides a genome-wide expression profile of high and low performers for place learning and memory (**Figure 5**). For place learning, 2076

genes were differentially expressed between the high-performing and low-performing cohorts. Of these, 947 genes were upregulated and 1129 were downregulated in the high learning cohort relative to the low learning cohort. For place memory, 590 genes were differentially expressed between the high and low performing cohorts, with 347 upregulated genes and 243 downregulated genes in the high memory cohort relative to the low memory cohort. There are 220 genes that are significantly differentially expressed in both the learning dataset and the memory dataset, the majority of which ($n = 169$) trend in the same direction for both phenotypes. There are several genes, including *dunce*, *amnesiac*, *white*, *radish*, *rutabaga*, *arouser*, and *tribbles*, that have been previously implicated as playing a role specifically in place learning and memory (Ostrowski and Zars 2014), which are labeled in **Figure 5**. Of these, only *radish* is significantly differentially expressed between high and low learning cohorts, with higher expression in the low learning cohort ($\log_2(\text{Fold Change})$: -0.4, $p_{\text{adj}} = 2.4 \times 10^{-5}$). Comparing memory cohorts, only *dunce* and *white* were significantly differentially expressed, and both were more highly expressed in the high memory cohort (*dunce*: $\log_2(\text{Fold Change})$: 0.38, $p_{\text{adj}} = 0.005$; *white*: $\log_2(\text{Fold Change})$: 0.6, $p_{\text{adj}} = 1.6 \times 10^{-5}$).

Identification of candidate genes – While our QTL credible intervals are fairly narrow, they do not provide the single-gene resolution that would be necessary to immediately identify a potential causative gene. I used our RNA-Seq data to identify potential regulatory candidate genes within our QTL BCI. For QTL shared between learning and memory, the causal gene should lie within the overlapping region of the learning and memory BCIs, assuming these QTL truly

represent a shared causal gene. This assumption allowed us to further narrow the search intervals for these shared QTL. I then determined which significantly differentially expressed genes fall within the interval of interest for the 16 identified QTL. I used two additional sources to inform the identification of potential candidate genes. First, the DSPR is a widely used resource, providing datasets that have the potential to integrate with this study. A previous study by King et al., (2014) performed genome-wide eQTL mapping of female heads for ~600 RIL crosses. These RILs were crossed between the population A RILs and the population B RILs, while our assays only used the population A RILs. Nevertheless, this dataset allows us to identify which genes within an interval have been previously identified as having a significant *cis* eQTL in the DSPR. Second, I examined the FlyBase (FB2018_06) controlled vocabulary terms associated with the genes in our QTL intervals to identify any genes with a potential role in learning or memory (Thurmond *et al.* 2019). Each QTL BCI is shown in Figure 6, with the significantly differentially expressed genes within the interval labeled, except for Q12, which spans the centromere and thus has an interval too wide to display effectively. Some QTLs are fairly wide, such as Q1, Q2, Q6, Q7, and Q13, and thus include a moderate number of differentially expressed genes with several possible candidates. In one case, Q12, the interval of interest includes no significantly differentially expressed genes. For all of my other QTL, I was able to identify potential candidate genes.

Among the QTL mapped only for learning, Q3, Q8, Q10, Q14, and Q16 are narrow enough to suggest a small number of possible candidate genes. Only six

genes that fall within the Q3 interval are differentially expressed, and just one has a *cis* eQTL in the King et al., (2014) dataset: *CG7058*, which is expressed at higher levels in the low learning cohort relative to the high learning cohort ($\log_2(\text{Fold Change}) = -0.7$, $p_{\text{adj}} = 9.3 \times 10^{-5}$). Little is known about the function of this gene beyond it is known to be expressed in the nervous system (Brown *et al.* 2014). Of the eight genes that fall within the interval for Q8, four have a *cis* eQTL in the King et al., (2014) dataset. Two of these genes have been shown previously to be involved in the nervous system. *Liprin-γ* is expressed at higher levels in the low learning cohort relative to the high learning cohort ($\log_2(\text{Fold Change}) = -0.27$, $p_{\text{adj}} = 0.016$) and is one of three *Liprin* genes in the *D. melanogaster* genome that are involved in synapse formation. Astigarraga et al., (2010) used null mutants of all three of the *Liprin* genes to show that *Liprin-α* and *Liprin-β* both promote synapse formation, and *Liprin-γ* acts antagonistically to both of these. In addition, *windpipe* (*wdp*) was more highly expressed in the high learning cohort relative to the low learning cohort ($\log_2(\text{Fold Change}) = 0.37$, $p_{\text{adj}} = 0.001$). This gene has been implicated in synaptic target recognition in the neuromuscular junction via a screen in which *wdp* was overexpressed in the target tissue in larvae and the number of mistargeting events was quantified (Kurusu *et al.* 2008). In the Q10 interval, of the six differentially expressed genes, three have previous evidence for a *cis* eQTL (King *et al.* 2014). *Multiplexin* (*Mp*) is more highly expressed in the low learning cohort relative to the high learning cohort ($\log_2(\text{Fold Change}) = -0.38$, $p_{\text{adj}} = 0.029$), encodes a Collagen XV/XVIII type protein, and has been shown to be critical for presynaptic homeostasis in the neuromuscular junction (Wang *et al.* 2014). Within

the Q14 interval, three significantly expressed genes have previous evidence for a *cis* eQTL (King *et al.* 2014). One of these genes, *Octopamine β 1 receptor* (*Oct β 1R*), is one of the octopamine receptors in flies and is more highly expressed in the low learning cohort relative to the high learning cohort ($\log_2(\text{Fold Change}) = -0.41$, $p_{\text{adj}} = 0.008$). In insects, octopamine is a major neurotransmitter that has been shown to play a critical role in learning, particularly reward-based learning (Balfanz *et al.* 2005; Burke *et al.* 2012). A previous study by Sitaraman *et al.*, (2010) found no difference in place learning in flies with greatly reduced octopaminergic signaling, suggesting octopamine signaling is not necessary for place learning. However, Koon and Budnik (Koon and Budnik 2012) also showed that the *Octopamine β 1 receptor* (*Oct β 1R*) inhibits synaptic growth, acting in opposition to *Octopamine β 2 receptor* (*Oct β 2R*), which promotes synaptic growth. The *Oct β 1R* inhibitory effect occurs via inhibition of the cAMP pathway, which is the same pathway affected by the well-known learning and memory mutants *dnc* and *rut* (Koon and Budnik 2012). Only four genes are significantly differentially expressed within the Q16 interval, and two of these have a known *cis* eQTL (King *et al.* 2014). One of these, *CG34353*, has been classified with terms such as axon guidance and synapse organization via the PANTHER classification system (Mi *et al.* 2017), though these functions have not been confirmed in the *D. melanogaster* system. This gene is more highly expressed in the low learning cohort relative to the high learning cohort ($\log_2(\text{Fold Change}) = -0.46$, $p_{\text{adj}} = 0.01$).

I mapped just two QTL specifically for our memory phenotype, and only Q15 was sufficiently narrow to identify possible candidate genes. There were just four

significantly differentially expressed genes within this interval, none of which have a *cis* eQTL in the King et al., (2014) dataset. All four of these genes are more highly expressed in the high memory cohort relative to the low memory cohort. The gene with the largest difference is the well-known *hedgehog* gene (*hh*; $\log_2(\text{Fold Change}) = 0.90$, $p_{\text{adj}} = 6.4 \times 10^{-5}$), which has multiple, critical roles in development (Thurmond *et al.* 2019).

For the QTL mapped for both learning and memory, our ability to narrow the interval of interest to the region of BCI overlap between both traits leads to a substantially smaller list of possible candidates. Within the interval of interest for Q4, both *Mothers against dpp* (*Mad*) and α/β hydrolase 2 (*Hydr2*) have a *cis* eQTL in the King et al., (2014) dataset. In addition, both of these genes are significantly differentially expressed between the high and low learning cohorts, with higher expression in the low learning cohort relative to the high learning cohort (*Mad*: $\log_2(\text{Fold Change})$: -0.37, $p_{\text{adj}} = 0.015$; *Hydr2*: $\log_2(\text{Fold Change}) = -0.38$, $p_{\text{adj}} = 0.015$). While these genes are not significantly differentially expressed between the memory cohorts, the trend is in the same direction in the learning cohorts for both genes (*Mad*: $\log_2(\text{Fold Change})$: -0.13, $p_{\text{adj}} = 0.67$; *Hydr2*: $\log_2(\text{Fold Change}) = -0.19$, $p_{\text{adj}} = 0.45$). *Mad* is a transcription factor that regulates the expression of BMP response target genes (Vuilleumier *et al.* 2019) and has also been shown to play a role in inhibiting synapse formation, as one part of two antagonistic signaling pathways that act together to produce the correct number of synapses (Jordán-Álvarez *et al.* 2017). In the Q5 interval of interest, just two genes are significantly differentially expressed, *Odorant-binding protein 28a* (*Obp28a*) and *Cyp4d21*,

neither of which have evidence for a *cis* eQTL in the King et al., (2014) dataset. Both of these are significant only between the learning cohorts, with higher expression in the high learning cohort (*Obp28a*: $\log_2(\text{Fold Change})$: 0.96, $p_{\text{adj}} = 1.9 \times 10^{-4}$; *Cyp4d21*: $\log_2(\text{Fold Change}) = 0.52$, $p_{\text{adj}} = 0.01$). Neither of these genes has any previous evidence implicating their involvement in place learning or memory. There are six significantly differentially expressed genes within the Q9 interval of interest, all of which have a *cis* eQTL in the King et al., (2014) dataset. Of these, *mwh* is the only gene that is significantly differentially expressed between both the learning cohorts and the memory cohorts (Learning: $\log_2(\text{Fold Change})$: 0.37, $p_{\text{adj}} = 0.004$; Memory: $\log_2(\text{Fold Change}) = 0.36$, $p_{\text{adj}} = 0.01$). In both cases, *mwh* is expressed more highly in the high-performing cohorts. This gene is primarily involved in the formation of hairs (Thurmond et al. 2019). Within the Q11 interval, there are two significantly differentially expressed genes, *Glutathione S transferase O1 (GstO1)* and *ergic53*, both of which have a *cis* eQTL in the King et al., (2014) dataset. *GstO1* is more highly expressed in the high learning cohort but is more lowly expressed in the high memory cohort, though the difference is only significant for learning (Learning: $\log_2(\text{Fold Change})$: 0.47, $p_{\text{adj}} = 0.03$; Memory: $\log_2(\text{Fold Change}) = -0.21$, $p_{\text{adj}} = 0.56$). The gene *ergic53* is more highly expressed in the low learning and low memory cohorts relative to the high performing cohorts (Learning: $\log_2(\text{Fold Change})$: -0.41, $p_{\text{adj}} = 0.03$; Memory: $\log_2(\text{Fold Change}) = -0.37$, $p_{\text{adj}} = 0.11$). An RNAi screen specific to neurons identified *ergic53* as a candidate gene involved in the perception of very high temperatures (Neely et al. 2010). Given the punishment in our learning paradigm is high temperature, this is

a potentially exciting candidate gene; however, it is counterintuitive that the low learning cohort shows higher *ergic53* expression compared to the high learning cohort.

2.5 Discussion:

I successfully used a large multiparent population to identify multiple loci affecting place learning and memory in the fruit fly model system. I integrated an RNA-Seq dataset that provided a genome-wide characterization of differential expression between high and low-performing cohorts to identify potential candidate genes at mapped loci. All of the identified loci represent novel loci not previously associated with place learning or memory.

Our study joins a small set of previous studies that have identified naturally occurring genetic variants influencing learning and/or memory (Chandra *et al.* 2001; Steinberger *et al.* 2003; Mery *et al.* 2007b; Papassotiropoulos *et al.* 2009; Parker *et al.* 2012, 2014; Knoll *et al.* 2016), none of which examined place learning and memory. Selection and quantitative genetic approaches have long suggested that natural variants in the genome support variation in learning and memory. But specifically how genetic variation gives rise to better or worse performance has been challenging. Selection experiments in systems such as blow flies (McGuire and Hirsch 1977b) and fruit flies (Mery and Kawecki 2002b; Mery *et al.* 2007c) have shown that natural genetic variants can be combined to yield better performance in learning tasks. However, only rarely have QTL for learning and memory been successfully mapped (Chandra *et al.* 2001; Steinberger *et al.* 2003;

Parker *et al.* 2012, 2014), as I have done here, and even more rarely have the specific genetic variants within a gene influencing learning and/or memory been identified (but see: (Mery *et al.* 2007b; Papassotiropoulos *et al.* 2009; Knoll *et al.* 2016). As we build a larger collection of QTLs that influence learning and memory in different systems, ideally validating the specific variants involved, we will gain a much greater understanding of the genetic mechanisms governing these processes.

Some of the major advantages of using a stable multiparent mapping panel such as the DSPR include the ability to measure multiple phenotypes on the same set of lines and the availability of additional datasets from other studies that can be integrated to address major questions. With these strengths, we achieve a comprehensive picture of the genetic basis of both place learning and place memory. By measuring both phenotypes on the same set of lines, I showed that both traits are genetically correlated and identify specific loci that influence both traits. A constant challenge for QTL studies following mapping is determining which genes within the mapped interval are the most likely candidate genes. In traditional two-way QTL mapping, these intervals are typically very wide and usually include hundreds of genes (Mackay 2001). The multiple generations of crossing in the DSPR lead to smaller haplotype segment sizes and thus higher mapping resolution (King *et al.* 2012a) however, single-gene resolution is still not possible, requiring strategies for identification of likely candidate genes. Mapping the same QTL for multiple traits allowed us to focus on an even narrower interval of interest. This approach also allowed us to significantly narrow the interval of interest by

considering only the area of overlap between the two BCI's. By bringing in additional phenotyping, in the form of RNA-Seq of high performing and low performing cohorts, I could focus on genes that were significantly differentially expressed between these cohorts. Changes in gene expression in identified shared QTL in the same direction lent more support for a given gene. Finally, a particularly useful feature of the DSPR is the constantly increasing database of previous studies using this same set of lines (<http://FlyRILs.org>). We specifically used a large, genome-wide eQTL dataset (King *et al.* 2014) to identify which genes within our intervals of interest had previously been shown to harbor a *cis* eQTL. This is a powerful approach to identify genes that are affecting learning and memory via changes in gene expression. Other studies using multiparent populations have used a similar approach of data integration to more effectively identify candidate genes for traits such as resistance to toxins in fruit flies (Marriage *et al.* 2014; Najarro *et al.* 2015), grain yield and flowering time in maize (Dell'Acqua *et al.* 2015) and body weight in mice (Aylor *et al.* 2011).

While I successfully identified potential candidate genes for many of our QTL, our approach does rely on several important assumptions. First, I focused on identifying candidate genes that are affecting learning and memory via regulatory variants, with a change in gene expression leading to a phenotypic change. However, it is also possible some of the genes underlying our mapped QTL are coding variants that do not lead to differences in the expression of the candidate gene. Our method of identifying candidate genes would miss such genes. Some of the few previously identified natural variants that have been found to influence

learning and/or memory in other systems do show differences in gene expression (Mery *et al.* 2007b; Papassotiropoulos *et al.* 2009), and it has been hypothesized that many of the variants underlying QTL may be regulatory (Gibson and Weir 2005; Gilad *et al.* 2008). Second, I did not expose the lines we used for RNA-Seq to the heat box prior to sample collection. The training and memory test in the heat box happens over a short timescale, with the entire assay lasting just under 10 minutes, and thus we considered the differences in the baseline gene expression levels between cohorts to be a more likely predictor of the RILs' performance on this test. Third, we performed RNA-Seq on pooled samples of female heads, which is an efficient way to average over a substantial amount of inter-individual and inter-RIL variability to pull out the set of genes consistently differentially expressed between high and low performers. However, it is also possible that this approach misses more subtle, but potentially important variation. For example, if differences in learning and memory are dependent on differences in gene expression in a specific set of cells, I could miss this effect using our approach. Future studies employing higher resolution expression analysis, such as the single-cell approaches that are currently being used to characterize expression patterns across the fly brain (Croset *et al.* 2018), have the potential to provide this fine-scale data. Finally, I note that I do not yet have additional functional information confirming that the genes I identify as candidate genes are causative, and it is certainly possible that the true causative gene for any QTL is one I have not identified here. Future follow up studies using approaches such as quantitative complementation (Turner 2014; Stern 2014) or targeted studies of gene

expression of potential candidate genes in the DSPR RILs along with learning and memory measurements will allow us to begin to determine exactly which of the genes in our intervals are causative.

Despite these caveats, our study represents a critical first step towards a more holistic characterization of the genetics of learning and memory in an important model system. The majority of previous investigations of the genetic basis of learning and memory have focused on single gene approaches. Several mutants have been identified that specifically affect place learning and/or memory (Ostrowski and Zars 2014). However, despite the success of these studies, a disconnect remains between the genes identified via single-gene approaches and the identity of the genetic variants leading to individual-level variation in learning and memory in natural populations. In this study, none of the previously identified genes shown to specifically influence place learning and memory via mutant studies (e.g., *dunce*, *amnesiac*, *white*, *radish*, *rutabaga*, *arouser*, and *tribbles*) are within our QTL intervals. In addition, very few of these previously identified genes were significantly differentially expressed between high and low-performing cohorts (see Results). This result is not wholly unexpected, given that learning and memory are expected to be highly complex, polygenic traits, and the mutations identified in single-gene approaches are of large effect, are often quite deleterious, and would presumably be selected against in any natural population (Waddell and Quinn 2001; Mery 2013b). A similar disconnect between candidate gene approaches and the identity of natural causative variants has been found for other traits such as lifespan (Remolina *et al.* 2012; Burke *et al.* 2014a; Stanley *et al.*

2017). Our study has instead taken a genome-wide approach, identifying loci involved in learning and memory without making major perturbations to the system, such as null mutations, blocking neurotransmitters, or eliminating major cell types. With this approach, I have localized multiple QTL that influence whether a genotype learns and remembers well or poorly. Several of these QTL are likely pleiotropic loci, influencing both learning and memory. Others only map to a single phenotype and are potential loci contributing to some independence in these two processes. I also characterize genome-wide expression differences between high and low-performing cohorts for these two phenotypes, providing a valuable profile of how high vs. low learning and high vs. low memory genotypes differ at the transcriptome level. Several other studies investigating the role of gene expression in learning have taken a before versus after approach, studying expression changes following training to identify which genes alter expression during and following learning (Cavallaro *et al.* 2002; Dubnau *et al.* 2003; Rapanelli *et al.* 2009). Our dataset instead provides a high performing versus low performing approach, identifying a set of hundreds of genes, whose regulatory differences potentially influence whether an individual will learn and remember well or poorly.

There is a growing appreciation for the complexity of learning and memory and the need to study these traits in realistic settings (Mery 2013b; Smid and Vet 2016). Recently there has been a shift from investigating single genes to identifying the sets of neurons and circuits involved in learning and memory (Sitaraman *et al.* 2008, 2017; Burke *et al.* 2012; Ostrowski *et al.* 2015). In addition, mapping studies in humans are beginning to uncover the natural genetic variants influencing

learning and memory (Papassotiropoulos *et al.* 2009; Heck *et al.* 2014). Although it is still early, the challenges associated with studying human learning and memory suggest that a complete picture of how a genome can support fundamental learning and memory mechanisms remains in the future. Examination of these processes in animal models like fruit flies, with systems genetic resources such as the DSPR, provide a chance to identify mechanisms comprehensively through careful behavioral measures, the integration of intermediate phenotypic data, and advanced quantitative genetics approaches.

2.6 Acknowledgements:

I thank Dr. Stuart Macdonald for supplying us with the DSPR RILs, Aditi Mishra for numerous helpful discussions about the project, and Research Computing Support Services at the University of Missouri, especially Christopher Bottoms and Jacob Gotberg, for their bioinformatics computing support. This material is based upon work supported by the National Science Foundation under Grant No. IOS-1654866 (T.Z. and E.G.K.), National Institutes of Health R01GM117135 (E.G.K.), MU Research Council Grant URC-16-007 (T.Z.), National Institute of Health IMSD R25GM056901 (P.A.W-S.), and an HHMI Gilliam Fellowship (P.A.W-S.).

2.7 Figure:

Fig 1:

Schematic of the phenotyping process of 741 RILs (48,940 individuals). RILs are stored at 18°C, until 1 day of “set up”, then 6 males and 10 females are placed into a flask and allowed to mate. Fourteen days post-oviposition adult F1 flies were collected into two separate vials. These flies were then phenotyped within the heat box. The rectangles depict a chamber within the heat box, blue (cool) and red (hot) represent 24°C and 41°C respectively, which are the temperatures that are associated with one-half of the chamber. If an individual is on the cool associated side of the chamber, then the whole chamber reflects that temperature and the same is true for the hot associated side. After 30 s pre-testing, learning is phenotyped for 6 mins., and immediately after the same individual is phenotyped for memory for 3 mins. Traces are the activity within a chamber of an individual fly. This individual has a higher learning score, because it spent less time on the hot associated side, whereas the memory score is medium because it started exploring the chamber toward the end of the trial.

Fig 2:

Phenotypic patterns in place learning and place memory among the DSPR RILs. a) Performance index (PI) scores for learning for each RIL in the DSPR. The RILs are sorted from lowest average learning to highest average learning. The PIs for each individual measured for a given RIL are displayed (black points) and the mean value ± 1 SE for each RIL is plotted in red. RILs selected for RNA-Seq in the low-performing cohort are labeled with cyan triangles, and the high-performing cohort is labeled with purple triangles. b) PI scores for memory for each RIL in the DSPR plotted in the same way as learning in (a). c) The relationship between learning and memory. The mean values for learning and memory for each RIL are plotted with black points. Horizontal red bars show the learning mean ± 1 SE and the vertical blue bar shows the memory mean ± 1 SE.

Fig 3:

Genome scan of learning and memory phenotypes. Different colors denote the different phenotypes of learning and memory. The dotted horizontal line denotes the 5% family-wise error rate threshold, and the solid horizontal line denotes the 5% false discovery rate. QTL peaks reaching either threshold is labeled with the QTL id. Shaded blue boxes denote chromosome arms.

Fig 4:

Haplotype means at each shared QTL. The average learning (power transformed) and memory PI are plotted for each RIL. Each RIL is colored by its most likely haplotype identity at the location of the QTL. The large black outlined points are the estimated average effect of each haplotype, and the horizontal and vertical bars show ± 1 SE of this estimated effect for learning and memory respectively. The individual RILs are plotted with semi-transparent points.

Fig 5:

RNA-Seq expression levels between high and low cohorts of learning and memory in the DSPR. a, b) MA plot of learning cohorts (a) and memory cohorts (b) showing the $\log_2(\text{Fold Change})$ versus the normalized average expression for each gene. Colors indicate the adjusted p-value for a given gene. c, d) Volcano plots of the comparison between high and low performing cohorts for learning (c) and memory (d). $\log_2(\text{Fold Change})$ are shrinkage adjusted and represent the high-performing cohort relative to the low-performing cohort. Colors indicate the overall average expression level. The points corresponding to a set of genes that have been previously associated with place learning and memory are labeled.

CHP 3 – DISSECTING THE GENETIC BASIS OF THERMAL TOLERANCE IN FRUIT FLIES

Authors: **Patricka Williams-Simon**, Ronel Ghidey, Enoch Ng'oma, Troy Zars,
and Elizabeth King

3.1 Abstract:

Thermal tolerance is a complex trait that is a fundamental survival skill in many species. For example, everyday tasks such as foraging, finding a mate, and escaping predation, are highly dependent on how well an organism can tolerate extreme temperatures. Understanding the natural variants of the genes that control this trait is of high importance if we want to better comprehend how this trait evolves in natural populations. Here, I take a quantitative genetics approach to dissect the genetic basis of thermal tolerance using *Drosophila* Synthetic Population Resource (DSPR). This multi-parental population consists of approximately 1,800 Recombinant Inbred Lines (RILs), which allows for high-resolution genome-wide scans, and the identification of loci contributing to naturally occurring genetic variation. Using a highly sensitive apparatus known as the “heat box” I presented the flies with a constant 41°C for a total of 9 ½ mins. and recorded the time it takes for an individual to become incapacitated as our measure of thermal tolerance. I found that thermal tolerance ranges from 24.8 sec – 361 sec, which is a ~14-fold difference between the lowest to the highest tolerant RILs. I then performed a genome scan and identified several loci (QTL) affecting thermal tolerance, including one large effect locus. I performed RNA-seq to identify differentially expressed candidate genes within these QTL that affect thermal tolerance. Future work will aim to validate the function of these genes.

3.2 Introduction:

Thermal tolerance is a complex physiological trait that determines the survivability of many species. Some examples of systems where thermal tolerance has been shown to be a critically important trait include coral, (Dixon *et al.* 2015), fish (Schulte *et al.* 2011; Anttila *et al.* 2013; Grindler *et al.* 2020), fruit flies (Overgaard *et al.* 2008; Bozinovic *et al.* 2016), birds (Nord and Nilsson 2011; DuRant *et al.* 2012, 2013), frogs (Pintanel *et al.* 2019; Díaz-Ricaurte *et al.* 2021), and lizards (Angilletta *et al.* 2002a). Thermal tolerance is particularly important for ectotherms (Huey and Berrigan 2001; Angilletta *et al.* 2002b; Frazier *et al.* 2006; Hoffmann 2010; Sunday *et al.* 2011; Gunderson and Stillman 2015), whose body temperature reflects that of the ambient temperatures. These organisms have evolved physiological mechanisms, like thermal tolerance, (Tosini and Menaker 1995; Hamada *et al.* 2008), which have allowed individuals to compensate for the change in temperatures (Partridge *et al.* 1994; Hoffmann and Sgrò 2011). Thermal tolerance has been shown to directly affect many biological processes on multiple levels, ranging from the speed of cellular reactions to ecological processes such as population growth (Frazier *et al.* 2006; Angilletta and Angilletta 2009; Schulte *et al.* 2011; Huey *et al.* 2012). Thermal tolerance has also been a critical factor as climate change becomes a bigger worldwide problem (Walther *et al.* 2002; Parmesan and Yohe 2003; Thomas 2010; McCain and Colwell 2011; Sunday *et*

al. 2012; Letcher 2015), and force the redistribution of many organisms. Despite the obvious importance of thermal tolerance across every life stage in all living organisms (Boher *et al.* 2010; Hammond and Hofmann 2010; Webster *et al.* 2011; Diederich and Pechenik 2013; Nguyen *et al.* 2014; MacLean *et al.* 2019b), much remains unknown about the genetic variants that contribute to thermal tolerance.

Several empirical studies demonstrated the phenotypic variation in thermal tolerance that exists between and among species including, (Cuculescu *et al.* 1998; Addo-Bediako *et al.* 2000; Sørensen *et al.* 2001; Pörtner 2002; Nærgaard *et al.* 2003; Fangué *et al.* 2006; Verble-Pearson *et al.* 2015; MacLean *et al.* 2019b). This variation has been well studied in the *Drosophila* species (Sørensen *et al.* 2001, 2019; Overgaard and Sørensen 2008; Overgaard *et al.* 2011; MacLean *et al.* 2019b, 2019a). This is due in part to the fact *Drosophila* is distributed throughout the world in a wide range of varying temperatures (Overgaard *et al.* 2011), making the *Drosophila* genus an ideal model organism to elucidate the genetic basis of thermal tolerance. *Drosophila melanogaster*, in particular, has been extensively investigated in multiple different paradigms, ranging from field-based to laboratory approaches (Wertheim *et al.* 2006; MacLean *et al.* 2019b; Grainger *et al.* 2021), from theoretical to empirical (Andersen *et al.* 2018; Chen 2019), and from cold to hot tolerance (Sunday *et al.* 2019). Traditionally, most genotypes to phenotype approaches in *D. melanogaster* have taken the single gene or candidate genes approach (Clapham 2003; Ramsey *et al.* 2006; Venkatachalam and Montell 2007; Rampino *et al.* 2009; Fowler and Montell 2013; Lockwood *et al.* 2017; Cleves *et al.* 2020). Therefore, much of our knowledge of the genes that influence complex

traits, like thermal tolerance, has been uncovered using these classical genetics approaches. From the early 2000s, as genome-wide systematic techniques became more feasible, there were also several studies that used genome association studies to identify QTLs and genomic regions that regulate thermal tolerance in the fruit fly (Norry *et al.* 2004, 2007a, 2007b, 2008; Morgan and Mackay 2006; Vermeulen *et al.* 2008b, 2008a; Rand *et al.* 2010; Rolandi *et al.* 2018).

Quantitative trait loci (QTL) mapping is a powerful tool that has been used to elucidate the genetic architecture of thermal tolerance in *D. melanogaster*. Despite the capability of these approaches, there are only a few studies that involve identifying QTLs that influence thermal tolerance. First, Norry *et al.* (Norry *et al.* 2004) and Vermeulen *et al.* (Vermeulen *et al.* 2008b, 2008a), from the same research group, selected for high vs. low resistance to thermal knockdown, and identified QTLs that influence thermal tolerance between two geographically distinct populations of *D. melanogaster*. Second, Morgan and Mackay (Morgan and Mackay 2006) used two inbred populations divergent for thermal tolerance and identified QTLs that influence low and high-temperature resistance. Furthermore, Rand *et al.* (Rand *et al.* 2010) identified QTLs through microsatellite mapping in laboratory selected and natural populations of *D. melanogaster*. Finally, Rolandi *et al.* (Rolandi *et al.* 2018) used a multi-parental population of fruit flies, the *Drosophila* Genetic Reference Panel (DGRP) to identify QTLs that influence high thermal tolerance. The results from the high-temperature studies (Norry *et al.* 2004; Morgan and Mackay 2006; Rand *et al.* 2010) show overlap in QTL on

chromosome 3R, which had some overlapping regions, however (Rolandi *et al.* 2018), did not identify any significant QTLs on chromosome 3R. Both GWAS and interval QTL mapping lack the resolution to identify single genes (Mackay 2001), but more recently researchers have been pairing those approaches with transcriptomics to identify candidate genes that are influencing complex traits (Cubillos *et al.* 2017; Highfill *et al.* 2017; Williams-Simon *et al.* 2019; Wen *et al.* 2019; Ng'oma *et al.* 2020; Lecheta *et al.* 2020). The integration of these approaches has made identifying candidate genes that influence complex traits more feasible.

In this chapter, we combine quantitative genetics with genomics to dissect the genetic basis of thermal tolerance in a Multi-Parental Population (MPP) of *D. melanogaster*. MPPs are relatively new systems that are used to tease apart complex traits in different model organisms (reviewed in (de Koning and McIntyre 2017a)(de Koning and McIntyre 2017b). Here we used the *Drosophila* Synthetic Population Resources (DSPR) developed by King *et al.* (King *et al.* 2012a, 2012b); <http://flyrils.org>). We measured over 700 Recombinant Inbred Lines (RIL) for thermal tolerance and identified large effect QTLs. We integrated this approach with RNA-seq between high and low cohorts and identified several candidate genes that are located within our QTLs. We, therefore, recommend that follow-up studies are needed to functionally validate those findings.

3.2 Methods:

DSPR and Fly Husbandry – We used the DSPR, which is a multi-parental population of *D. melanogaster*, (King *et al.* 2012a, 2012b; Long *et al.* 2014); <http://FlyRILs.org>). The DSPR consists of approximately 1,600 Recombinant Inbred Lines (RILs), with two subpopulations (A and B), which consist of approximately 800 RILs each. Both subpopulations have unique founders except for one that is shared, AB8. Each RIL is a mosaic of the 15 founders (parental) lines, which have been fully re-sequenced, and the RILs have been genotyped at over 10,000 SNPs. This structure makes genomic analysis of this population more feasible, because of the ability to be able to track entire haplotypes (King *et al.* 2012a). To generate the RILs, the founders were mixed for 50 generations, with two major goals - first to create genetic diversity, then for high mapping resolution - and then the population was inbred for another 25 generations to create stable lines (King *et al.* 2012b). We phenotyped 741 RILs in subpopulation A, scoring thermal tolerance.

These RILs were raised at 18°C, with 60% relative humidity on a cornmeal diet, and added to live yeast to the top of the food after it cooled. Two weeks before phenotyping for thermal tolerance, 10 female and 6 male flies were flipped onto a new flask (Fisher Scientific, Cat. No.:AS-355) of food and stored at 25°C for one week to allow for mating and egg-laying. After the first week (7 days, post-oviposition) all the adults (parental generation) were removed from the vials to ensure that there was no mixing of generations. After Two weeks, (14 days post-oviposition) adult flies were anesthetized on ice for no more than 10 mins., and 60-

80 female flies were selected for the thermal tolerance assay. Flies were then divided equally and placed into two separate vials of food and were allowed to recover for at least 24 hrs period before phenotyping.

Phenotyping – Flies were assayed within a well-controlled and highly sensitive apparatus known as the “heat box” (Wustmann *et al.* 1996). Within the heat box, there are 16 individual chambers, which allows for concurrent testing of flies. Temperatures are set at 24°C (~75°F) for the first 30 secs (acclimation), then immediately shifts to 41°C (~106°F) for 9 ½ mins for a total of 10 mins. We define thermal tolerance as the amount of time it takes before a fly becomes incapacitated for at least 60 seconds, after which an incapacitation time is assigned. Two groups of flies of the same genotype – up to 27 individuals total – were phenotyped within the heat box on any given day. We planned on assaying 6 groups per RIL, however, because of some technical difficulties with the heat box, where some chambers here do not accurately reflect assigned temperatures. Because of such a setback, we assayed more groups to ensure we had a minimum of 30 individuals for each RIL. We also measured the founding parents for thermal tolerance. We followed the same phenotypic setup as we did for the RILs.

QTL mapping and Heritability – We performed a genome scan on 741 RILs using a custom R script as described in Williams-Simon *et al.* (Williams-Simon *et al.* 2019). We took the average incapacitation times (phenotype) from each RIL and applied the Haley-Knott regression (Haley and Knott 1992). This statistical approach regresses each phenotypic average on the eight founder

haplotype probabilities (Broman and Sen 2009; King *et al.* 2012a) by fitting the following model:

$$Y_i = \sum_{j=1}^77 p_{ij} b_{ij} + e_i,$$

where **y_i** is the phenotype (i.e. thermal tolerance) of the *i*th RIL, **p_{ij}** is the probability the *i*th RIL has the **j**th haplotype at the locus, **b_{ij}** is the vector of effects for the **j**th haplotype, and **e_i** is the vector of residuals (Broman and Sen 2009). The phenotype with the lowest p-value was considered most associated at that locus. To identify statistically significant QTLs, we tested for False discovery rate (FDR), which identifies the expected false positives/ total positives, and the Family-wise error rate (FWER) tests for the probability of one or more false positives (QTLs) in the entire experiment. The FWER is stringent of the two, and therefore, has a higher threshold and confidence. I permuted thermal tolerance along with our place learning and memory phenotypes (Williams-Simon *et al.* 2019) because the individuals that were assayed for thermal tolerance were first assayed in a learning and memory assay and were exposed to 41 prior. Initially, we hypothesized that there would be a trade-off between individuals with high learning to have low thermal tolerance because they would have a higher avoidance for the high-temperature associated side of the heat box, however, that was not the case (**Fig 1c**).

RNA-sequencing – I performed RNA-seq to determine differentially expressed genes between the top and bottom cohorts of thermal tolerance. First, I selected 14 RILs from the high (n = 7) and low (n = 7) 5% cohorts of the first 140 RILs I phenotyped and followed the same set-up protocol (see phenotyping section above). After oviposition day 14 adult female flies were flash-frozen in liquid nitrogen and stored at -80, until further processing. In addition, I also took the same RILs and cycled them through an ‘incubation assay’ for 1 min intervals at 41°C and 24°C (room temp) for a total of 6 mins., and left control samples on the benchtop at a constant 24°C also for 6 mins. Then, all the vials – from both experimental and control groups – were left on the benchtop at room temp for 3 hrs, to ensure that the gene expression changes had time to occur but also that there was no time to allow for gene expression turn over. Immediately after the 3 hrs, flies were flash-frozen in liquid nitrogen and stored at -80°C. For both the heat box and incubator assays, only female flies ranging from 15-22 days old post-oviposition were assayed. For the incubator assay, all RILs were tested between 8 am and 10 am, to help reduce possible confounding effects that would affect gene expression. The samples were then freeze-dried overnight to prevent degradation of RNA, using a lyophilizer (*Labconco, Cat No.: 77550-00*). Second, the samples were shipped to Rapid Genomics (fee for service) for transcriptome processing. The total RNA was extracted with Dynabeads mRNA direct kit from life technologies, mRNA was then fragmented and converted into double-stranded cDNA, followed by the standard proprietary library prep for one lane of Illumina *HiSeq 3000* instrument to generate paired-end (PE) reads. The first 16 samples were PE 150-bp, while the remaining

samples were PE 100-bp. Third, we analyzed the RNA-seq data using a bioinformatic pipeline, as described before in Williams-Simon et al. 2019. Briefly, we trimmed the samples using the software *cutadapt* (Martin 2011), then ran a quality control test using *fastqc* (Andrews and Others 2010), which did not show any significant issues with the reads. We then used parts of the *Tuxedo pipeline* (Pertea et al. 2016) to align the reads against the *D. melanogaster* reference transcriptome (Ensembl Release Version: 84) using *HISAT2* (Kim et al. 2015) and *StringTie* (Pertea et al. 2015, 2016) assembled and quantified transcripts. To calculate the read count for each gene we used a *prepDE* python script from *StringTie*. The data was then transferred to the *DESeq2 package* (Love et al. 2014), – using the *prepDE* python scripts in *StringTie* – which uses the count data to identify differentially expressed genes. We filtered the zero counts samples, performed surrogate variables analysis (SVA) on normalized counts (Leek et al. 2012, 2019) to estimate any unknown batch effects. We identified 2 significant surrogate variables (SV), which we included in all following bioinformatics analyses. We then used the *DESeq2 package* (Love et al. 2014) to test for overall treatment effects and performed contrasts between pool (high and low cohorts) and condition (temperature: 41°C or room temp). Within the pool and condition treatments, we separated our groups into thermal tolerance high (TH) and thermal tolerance low (TL), and thermal tolerance high incubated (TH_Incu), thermal tolerance low incubated (TL_Incu). We included batch in our model because some samples were processed on different days. The model is as follows:

$$DESeq(test="LRT", reduced=\sim SV1 + SV2 + batch + condition)$$

DESeq(test="LRT", reduced=~ SV1 + SV2 + batch + pool)

To visualize and rank the genes we used the function *lfcShrink*, which performs shrinkage on log2(Fold Changes). All log2(Fold Changes) reported here are the shrinkage estimated values using the “*normal*” estimator. We then used that information to identify candidate genes that are differentially expressed between the different groups and aligned them with the BCI of the mapped QTLs (Q1-Q8). In addition, we compared our expression data set to a genome-wide eQTL study that also used the DSPR (King *et al.* 2014), which identifies cis-eQTLs. Lastly, we took advantage of the FlyBase resource to manually investigate the keywords (response to stimuli, thermal regulation, ion channels, and sensory and motor neurons) that could indicate that the genes are involved in thermal tolerance.

3.4 Results:

Phenotypic distributions – We quantified thermal tolerance in 741 RILs (39,392 individuals) in the DSPR using the heat box. We considered the time it took for an individual to become incapacitated, assigned an incapacitation score, then took the average over a RIL. We measured a minimum of 30 females, an average of 53 females per RIL (**Fig 1a**). We found that there is a wide range of phenotypic variation with approximately a 15-fold range from 24sec – 361sec between individuals with low and high incapacitation times (**Fig 1a**). We expected a wide range in thermal tolerance measurements because the original founding parents of the RILs were collected from different continental locations with varying temperatures, therefore the thermal tolerance is expected to vary. We established

that the estimated broad-sense heritability (H^2) for thermal tolerance is $H^2 = 0.24$, 95% CI = 0.22–0.27 and the effect of RIL is highly significant measuring at 125241.2, $p < 0.0001$, indicating that there is a genetic basis of thermal tolerance. Together this data confirms that thermal tolerance has a genetic basis because there is a significant amount of variation that is due to variation in genetic factors.

We also phenotyped the DSPR founders (A1, A2, ..., A7) for thermal tolerance (**Fig 1b**). We measured a total of 510 individuals, with an average of 73 individuals per founder (range: 56 – 81). Founder A7 had the highest average phenotype of 242 secs (range: 2 – 450 secs), while A6 had the lowest average of 105 secs. We did not phenotype A8, because the line did not thrive in our hands. There could be many factors that contribute to the line not thriving, including differences in food, or other rearing conditions.

Quantitative Trait Loci: I performed genomic scans following the methods of (King *et al.* 2012a, 2012b), to identify QTLs and permuted the 5% Family-wise error rate (FWER) and False Discovery Rate (FDR; see methods for details), to determine the significance of the QTLs and identified 8 QTLs (Q1-Q8; **Fig 2**) at the 5% FDR and higher. Of those 8 QTLs, 5 of them reached or passed the FWER threshold, which is the threshold with more confidence. Any significant QTLs are suggestive regions that influence thermal tolerance, with Q6 having the strongest effect. For each significant QTL, we determined the region of causative variants by obtaining a Bayesian Credible Interval (BCI; (Manichaikul *et al.* 2006; Broman and Sen 2009). For Q6 (**Fig 4**), the BCI is 20.93 – 21.09 on chromosome 3R (**Table 2**). We did not calculate the distances of Q1 because it spans the centromere and

Q2 and Q8 both had negative numbers; thus we did not include Q1 and Q8 in my analyses moving forward.

We assigned haplotype identity for each of our significant QTLs (except Q1). One advantage of the DSPR is that we are able to assign haplotype identity to any given location in every RIL (King *et al.* 2012b). Therefore for each of our significant QTL, we can estimate the founder (A1, A2, A3, A4, A5, A6, A7, and A8) haplotypes' effect at the specific QTL location for thermal tolerance (**Fig 5b**). At Q6 (**Fig 5a**), A5 founder has the average highest incapacitation score of 173 secs, while founder A4 has the lowest score of 125 secs (range), and A3, A6, and A7 have average scores of 130, 155, and 162 seconds respectively.

RNA-seq and genome-wide expression patterns – First, we performed RNA-seq on a pooled set of female heads to identify differentially expressed genes between low and high cohorts of thermal tolerance. In total, we had 6 pooled samples with 5 biological replicates for each. For each sample, we selected 7 RILs from the low-performing cohort and 7 from the high-performing cohort. The same 14 RILs were also processed through a control assay (see methods for more details): 2 sets (low and high) as experimental controls and 2 sets were temperature controls. For each of the 6 samples, we pooled a total of 35 female heads in a single sample, with 5 heads from each RIL contributing to each sample. This design was the most cost-effective and more efficient to account for both individual and RIL level variability. We then performed RNA-Seq and obtained genome-wide expression profiles (**Fig 3**) for 2 separate cohorts which we refer to as thermal tolerance high (TH) and thermal tolerance low (TL), and thermal

tolerance high incubated (TH_Incu), thermal tolerance low incubated (TL_Incu). These were then placed into three batches, pool (high and low), condition (incubated and room temp. cohort), and interaction between pool and condition. Room temperature is our control group.

First, we determined genome-wide gene expression patterns (**Fig 3a,b**). There were 2642 genome-wide genes that were differentially expressed between high and low cohorts for thermal tolerance. In this comparison, 1351 genes were upregulated and 1291 were downregulated from the baseline (**Fig 3a**). In the incubated vs room temp., there were 2862 genome-wide genes that were differentially expressed between high and low cohorts for thermal tolerance. In this comparison, 1253 Genes were upregulated and 1609 was downregulated from the baseline (**Fig 3b**).

To determine the level of genome-wide expression patterns over high and low cohorts and temperature, we performed principal components analysis (PCA) on all samples using an expression matrix from which batch effects had been removed (**Fig. 3d**). Before running the PCA, we first removed the batch effects (described in the methods section). PC1 explains 29% of the variance, while PC2 explains 22% of the variance. Both high and low pools at high temperature (Incu) clustered together, whereas there is substantial variance between high and low at room temperature (RT).

Candidate Genes – We identified several novel candidate genes using a combination of approaches. We narrowed our search for a possible candidate gene(s) from genome-wide expression to concentrate on genes that were

differentially expressed and also located with the BCI of our QTLs (**Table 1**; Q2–Q8). There were 29 genes within Q2 that were significantly expressed within the condition batch and 18 with pool. Of the 29 genes in the condition batch, 5 of them are in the HSP family of genes (*Hsp22*, *Hsp23*, *Hsp26*, *Hsp27*, *Hsp67Ba*) and 1 gene from the TRP (*TrpA1*) family. Only 1 HSP (*Hsp23*) and no TRP genes were significant within the pool batch. We found no genes that were significant when we computed an interaction between pool and condition. All other genes are considered novel findings within this paradigm. We continued our statistical analysis for the remaining QTLs (Q3–Q8) throughout the rest of the dataset to identify similar patterns. From here on, we are only reporting expression data for the pool (low and high) cohorts, however, all the data for the condition could be found in **Table 3**. There are 49 significant genes within Q3, 35 within Q4, and 5 within Q5, while Q6, Q7, and Q8 have 8, 15, and 66 genes respectively.

Of the 8 significant QTLs, Q6 was most significant (**Fig 2**), indicating that there is a large effect variant located within the BCI of the QTL, and therefore we consider Q6 a QTL of high interest. Q6 is located on chromosome 3R there are 21 genes located within that region (**Fig 5a**). Eight of the 21 genes were significantly expressed, 2 of these genes are *CG31207* and *CG7079* and unfortunately, we do not know much about them to conduct further investigation. Of the 6 remaining genes: *annexin B9 (anxB9)*, *cortactin (cortactin)*, *dynein heavy chain (dhc93AB)*, *malvolio (mvl)*, *Na pump alpha subunit (atpalpha)*, *rho GTPase activating protein (RhoGAP93B)*, *atpalpha* was the only gene that was upregulated, with log2 fold change of 0.08 (padj = 0.0000041). All the other genes were downregulated, with

Dhc93AB being the most significantly downregulated with a log2 fold change of -3.426297291 ($p_{adj} = 0.01$). We then manually looked for annotations in Flybase (Thurmond *et al.* 2019) for the genes listed above. We searched: response to stimuli, thermal regulation, ion channels, and sensory and motor neuron involvement, however we were not able to build a strong correlation between our list of 8 genes and those terms. For instance, *mvl*, *dhc93AB*, *calx*, *atpalpha* are all 'involved' with ion channels, and both *mvl* and *dhc93A* are significantly expressed, however, we were not able to make any other correlations, suggesting the genetic complexity of thermal tolerance. Finally, we compared our dataset to an eQTL study done by King *et al.* (Long *et al.* 2014). This comparison allows for the identification of cis-regulatory genes that might be affecting our candidate genes under Q6. Unfortunately, we did not identify any cis-regulatory genes that influence the genes within our dataset.

3. 5 Discussion:

We identified 8 genomic regions that influence thermal tolerance in the DSPR, including one large effect locus on chromosome 3R. We also identified a large set of genes that are differentially expressed between cohorts that differ in high vs low thermal tolerance. We integrated our RNA-seq findings with our QTLs to identify 21 suggestive candidate genes. These findings, of novel loci that influence thermal tolerance, add to our understanding of the genetic basis of this key trait.

Our results expand on previous studies that took genome-wide approaches to elucidate the genetic basis of thermal tolerance in *D. melanogaster* (Norry *et al.*

2004, 2007a, 2007b, 2008; Morgan and Mackay 2006; Rand et al. 2010; Rolandi et al. 2018). It is important to note that the measurements for thermal tolerance within these studies are different, so direct comparison might not be possible. Nevertheless, it is essential to observe if there might be any overlap or similar patterns within genomic regions, especially because pleiotropic interactions are a major contribution to complex traits, like thermal tolerance. We first compared our QTL dataset to Norry et al. (Norry et al. 2008), because they characterized thermal tolerance in two divergent populations that were measured for thermal tolerance in both males and females. We think this is a fair comparison because this QTL was identified using heat-hardening measures, therefore the exposure temperature is similar to our measurements. Note that we only compared the female QTL dataset. We found that our Q2 had an overlap with the data that was presented by Norry et al., however, when we compared to the list of candidate genes only *hsp23* was common among the two groups. This is not surprising, because in Frydenberg et al., (.2003), performed a DNA sequencing variation of *hsp23*, *hsp26*, and *hsp27*, and found that *hsp23* on different latitude clines in *D. melanogaster* and found that *hsp23* changes in latitudinal frequencies. Both the founding parents from DSPR – which we used in this study – and the RILs used in Norry et al., (Norry et al. 2008) were collected in different climatic locations. We also compared our study to the findings from (Rolandi et al. 2018). This group elucidated the genetic basis of thermal tolerance by measuring the upper thermal limits of 34 lines from the DGPR and performed GWAS to identify QTLs. The DGRP – like the DSPR – is one of the multi-parental populations in fruit flies.

Briefly, they reported 8 major SNPs that are associated with high temperatures in both males and females. We did not find any overlaps between the SNPs they identified as candidates and our QTLs nor did they correspond with genes that were significantly expressed within our QTLs. We also cross-referenced the genes found in these studies to the genes that are significantly differentially expressed between our high and low cohorts and we found no overlaps, except for *hsp23*. Based on our comparisons to these genomic-scan datasets, we conclude our study has identified novel QTLs and candidate genes that influence thermal tolerance in the *D. melanogaster*.

Our findings emphasize the natural phenotypic and genetic variation in thermal tolerance (Sørensen *et al.* 2001, 2019; Norry *et al.* 2004, 2008; Overgaard and Sørensen 2008; Overgaard *et al.* 2011; Rolandi *et al.* 2018; MacLean *et al.* 2019b, 2019a). These findings present several lessons to be learned. The first is that different genomic regions influence different types of thermal tolerance, although there might be some pleiotropy. There are very many 'levels' (cold and hot) of thermal tolerance which is controlled by different genomic regions, and even within a specific 'level' there are different ways to measure these processes, including heat-hardening, heat ramping, constant hot temperatures (Dhaka *et al.* 2006; Overgaard and Sørensen 2008; Sørensen *et al.* 2019; Grinder *et al.* 2020). But, despite the mode of measurement, ultimately the question is, are these organisms able to withstand extreme temperatures, and what biological processes are contributing to these temperature effects? For example, we did not identify any *hsp70* genes in our integrated approach comparing high and low thermal tolerance

groups. There could be many factors that may have contributed to this pattern, below we expand on a few. Most of the studies that identified *hsp70* using *D. melanogaster* had a longer exposure or repeated exposure (Norry *et al.* 2008; Rampino *et al.* 2009; Cleves *et al.* 2020), whereas – in the case of the condition treatment – our measurement of thermal tolerance was a total of 10 mins. Second, it is also important to recognize that most of the studies in *D. melanogaster* that examined the role of HSPs have taken a mutagenesis approach, and while this approach is powerful, it is not focused on natural genetic variation that is segregating. Our focus on natural genetic variation is one of the advantages of the DSPR. In the DSPR system, the fly genome is unmodified, however, we are able to rear them in the lab within a controlled environment, therefore reducing possible confounding effects. The DSPR in combination with quantitative genetics approaches allows for unbiased population genomic scans to uncover natural segregating variants that would otherwise be ‘impossible’ using single gene approaches. Because of this systematic approach we are able to identify many loci that influence complex traits, like thermal tolerance. Lastly, we were also surprised to find a study by Jensen *et al.* (Jensen *et al.* 2010) that suggests that *hsp70* is not related to adult heat tolerance variation in adult *D. melanogaster*, which supports our findings.

Genomic scans are often paired with RNA-seq to further dissect the genetic basis of complex traits. In this chapter, we integrated RNA-seq to identify differentially expressed genes between high and low thermal tolerance cohorts within our 8 QTL regions. This integration continues to be a powerful approach

used to narrow down potential candidate genes within QTL that affect complex traits. We compared both low and high (baseline thermal tolerance) groups and incubated vs. room temp for genome-wide expression. We found that different genes are significantly expressed between the two different groups (pool and condition) although the genotype is the same. The difference in gene expression suggests that having baseline levels of those genes are necessary but not sufficient to withstand high temperatures, and therefore being exposed to higher temperatures increases thermal tolerance (Lansing *et al.* 2000; Kellett *et al.* 2005; Norry *et al.* 2008; Sgro *et al.* 2010). Second, without this integration, we would only have loci and not candidate genes. This is because the resolution of QTL mapping could be very low and sometimes there are tens to hundreds of genes that are located within a region. For example, there are 472 genes located within the BCI of Q8 for the high and low cohorts, however, through RNA-seq, we were able to narrow down this list to 66 candidate genes. There are other follow-up studies that could be done to further narrow down this list, including eQTL mapping (King *et al.* 2014). However, in this study, we focused on narrowing the list of genes that were in our most significant QTL. Notably, we might be excluding some valuable candidates, however, we were mindful of the feasibility of the study, and these genes that we present here (Table 2) are worth doing follow-up studies on.

While we successfully identified novel genomic regions and candidate genes that influence thermal tolerance, are some important limitations to this study to consider. First, most of the naturally separating variants that influence complex traits have small effects (Boyle *et al.* 2017), and given research feasibility, it is

difficult to identify small-effect QTLs. Second, thermal tolerance is very polygenic, where many genes influence the trait, which makes it difficult to identify candidate genes even within large effect QTLs (Boyle *et al.* 2017). Third, the Beavis effect, where the confidence of QTLs gets stronger with more individuals measured, but the term 'more' is very arbitrary. In the present study, we measured more than 48,000 individuals and set very strict thresholds FWER and FDR, therefore we are confident QTLs are reliable. We also acknowledge the phenotypic variation we observed within a RIL. The flies were measured in the heat box (see methods for explanation), which is a highly controlled environment. However, even within a highly controlled environment, there might be some environmental and technical influences that we are not able to control. For example, which position within a vial the pupas are located and how much food an individual ingests. Those factors could be influencing the variation we see within a RIL. We accounted for those possible errors by assaying a huge number of individuals over a few days and calculating the means of a particular RIL. We also recognize that a block design would have ideally been a better approach for measuring thermal tolerance, where we would measure 3 different replicates of the same genotype (RIL). However, we were not able to design a blocked type experiment due to the number of individuals that were measured (741 RILs, and 39,392 individuals). It is also important to note, that we initially hypothesized that there will be a tradeoff between learning (see Chp 2) a thermal tolerance, ie RILs that had a low learning score would have a higher incapacitation time, however, the results suggest that there is no correlation between the two traits (**data not shown**). Because of our initial hypothesis, we

measured thermal tolerance on individuals directly after measuring for learning and memory, where it is possible that some individuals were exposed to 41°C, at different time durations. We selected the RILs for RNA-seq after measuring 140 RILs (high and low = n7 RILs), because we saw there was a clear pattern of RILs remaining in the high or low cohort. Unfortunately, after measuring the additional 601 RILs, the positions of the RILs that we selected for thermal tolerance changed. Most of them remained within the high and low measurements, except RIL 11042 had a lower incapacitation time than 10024 (**Fig 1a**). Ideally, we would have dropped those datasets from the RNA-seq analysis, however, we pooled the samples so that is not possible. It is the case, however, that the high and low pools contain a majority of high and low RILs, respectively. We, therefore, caution that our results could change if we had performed RNA-seq after measuring all 741 RILs. Therefore, we present that the 21 genes within Q6, should be considered as candidate genes until further studies suggest otherwise.

In this Chapter, my aim was to identify the genetic basis of thermal tolerance using the DSPR. I have successfully **1)** measured and quantified thermal tolerance in 741 RILs, **2)** identified 8 genomic loci within the fruit fly genome that influence thermal tolerance through internal QTL mapping, and **3)** provided a list of novel candidate genes that are controlling thermal tolerance. Collectively, these results suggest that the genetic underpinnings of thermal tolerance are very complex rather than just a few large-effect genes. Therefore, I propose that functional validation of the candidate genes, along with the identification of key pathways

involved in thermal tolerance is necessary for our further understanding of the genetic basis of thermal tolerance.

3. 6 Figures:

A.

Fig 1a:

Scatter plot showing phenotypic variation for thermal tolerance in the DSPR. RILs (arranged min to max by mean) are on the x-axis and performance index scores are on the y axis. Every black dot denotes an individual within a RIL and red bars are the mean ± 1 SE. The 7 high (purple) and low (aqua) are labeled on the x-axis.

B.

Fig 1b:

Scatter plot showing DSPR founders phenotypic variation for thermal tolerance. Colors represent a founder. Each dot denotes a RIL. Boxes and black lines represent the mean and ± 1 SE of the mean.

Fig 1c:

The relationship between learning and thermal tolerance. The mean values for learning and thermal tolerance for each RIL are plotted with black points. Salmon lines show the learning mean $\pm$ 1 SE and the teal lines show the thermal tolerance mean $\pm$ 1 SE.

Fig 2:

QTL map showing genomic regions that influence thermal tolerance in the DSPR. The $-\log_{10}(\text{P-value})$ is shown on the y axis, while the genomic position on the x. Perforated (5% family-wise error rate) and solid (%false discovery rate) gray horizontal lines are significance thresholds. Any of the orange lines that meet or pass the FWER and the FDR are significant QTLs, which are labeled Q1-Q8.

**Fig
3a:**
RNA-
Seq

expression levels between high and low cohorts and temperature. A) MA plot between high and low pools, B) MA plot of incubated vs. Room Temp. showing the $\log_2\text{FC}$ versus the normalized average expression for each gene. Colors indicate the adjusted p-value for a given gene.

Fig 3b:
Volcano plots showing genome-wide fold change expression between high

Fig 3c:

Volcano plots showing the pairwise comparisons between different groups.

Fig 3d:

Principal component analysis (PCA) showing the global effect of temperature on high and low cohorts. Colored circles represent high or low cohorts in different temperature environments (incubated at 41°C = Incu and room temperature at 24°C = RT). Two PC dimensions are shown here: PC1 and PC2.

Fig 4:

Q6 BCI interval. Red dots represent a gene. Brightly colored dots mean that the gene is significantly differentially expressed, which is also labeled. The horizontal line represents the baseline level of expression while the vertical dashed lines represent the lower and upper BCI for Q6.

Fig 5a:

Scatter plot showing founder mean effect at Q6 BCI. Legend represents the founder means (A1, ..., A8). If the founder is not listed in the legend it means that the specific founder does not have an effect on thermal tolerance at that loci. Dots represent RILs. Black lines are the ± 1 SE of the mean.

B

Fig 5b:

Scatter plot showing founder haplotype mean effect at QTLs except for Q1 and Q6. Legend represents the founder means (A1, ..., A8). If the founder is not listed in the legend it means that the specific founder does not have an effect on thermal tolerance at that loci. Dots represent RILs. Black lines are ± 1 SE of the me

Tables:

QTL ID	Position (bp)	Number of Genes in BCI	Name Significant Genes
Q2	3L:8816900-9416900	82	<i>hsp23; rdl; bol; Fdx1; teq; ValRS-m; CG5653; CG5194; Use1; PGRP-LA; PGRP-LCPGRP-LF; CG4080; Cpr67B; pix; Galk; dally; GluRIB</i>
Q3	3R:5534278-8284278	445	<i>Dfd; dsx; Gld; Rm62; Ccp84Ag; RpA-70; Xe7; noi; Mlp84B; CG2846; alpha-Est1; alpha-Est9; Gasp; CG2082; MTA1-like; Pi4Kllalpha; CG2926; CG2911; CG1213; Osi24; Osi2; Osi18; CG17919; CG15186; CG15177; CG10068; thw; CG1142; pyd3; CG10919; CG17816; CG2767; CG7900; CG7878; Ada2b; CG31549; CG33722; snoRNA:Psi28S-3385b; snoRNA:Psi28S-3091b; Alh; plx; CheA84a; lncRNA:CR42651; CG42675; gpp; asRNA:CR44093; m; lncRNA:CR45902; lncRNA:CR45906</i>
Q4	3R:15264278-17374278	307	<i>ea; ss; Tm1; Sap47; ATPsynO; ND-23; Rh6; Scp2 SerRS-m; IKKbeta; CSN5; Manf; CG10345; Sra-1; Ccm3; CG6125; Trissin; AOX3; CG5399; CG5404; Hmt-1; EndoU; CG8927; Zip89B; CG14880; CG14894; CG14892; CG14907; beat-IIb; CG6006; CG42342; nsI1:CG45218; asRNA:CR46029; asRNA:CR46058</i>
Q5		32	<i>DI; ATPsynD; CG6040; Cyp12a5; CG42359</i>

	3R:1909427 8-19344278		
Q6	3R:2093427 8-21094278	21	<i>AnxB9; Atpalpha; Mvl, Dhc93AB; Cortactin; CG7079; RhoGAP93B; CG31207</i>
Q7	3R:2600427 8-26634278	78	<i>E(spl)m2-BFM; E(spl)m3-HLH; Vps2; dysf; CG6142; CG12290; CG6154; CG6036; ppk15; CG5432; CG31324; scrib; asRNA:CR44321; lncRNA:CR46060; lncRNA:CR46084</i>
Q8	3R:2697427 8-30404278	472	<i>CecB; CecC; janA; l(3)mbt; Mlc1; Pkc98E; Jon99Ciii; spz; Klp98A; Pglym78; ATPsyngamma; Men-b; TwdIT; CG14257; CG17189; caix; CG6051; CG5909; CG14262; CG14260; CG5590; CG12428; BCAS2; mfrn; beat-VI; CG1646; Nepl18; wat; CG14518; CG11841; CG11842; CG11843; CG11873; Pdhb; Ctl2; CG14516; CG14511; rdog; CG11951; CG2310; CG2010; Obp99d; Obp99b; CG7834; CG7824; CG15514; Zip99C; CG15515; Capa; Nph; CG9747; CG9733; CG31038; snoRNA:Psi18S-1377d; ALiX; Tpi; CG42557; larp; SP1029; lnx3; Doa; sima; tau; asRNA:CR45209; mil; CG46335</i>

Table 3.1:

Significant QTLs and all the genes that are differentially expressed within each QTL for the **pool** group.

QTL ID	Position (bp)	Number of Genes in BCI	Name Significant Genes
Q2	3L:8816900-9416900	82	<i>Hsp22; Hsp23; Hsp26; Hsp27; Hsp67Ba; CG4456; Hsp67Bc; Rdl; Fdx1; smg; TrpA1; CG5021; CG5026; CG5644; Doc2; Dhpr; PGRP-LC; CG4461; CG4080; Cpr67B; CG3982; GNB3; CG32032; CG32039; Galk; nwk; asRNA:CR43481; CR43482; mfr</i>
Q3	3R:5534278-8284278	445	<i>alphaTub84B; Ccp84Ag; Ccp84Ab; Ccp84Aa; TfIIIFalpha; Xe7; unc-45; Sas-4; Mcm2; CG2846; alpha-Est1; alpha-Est10; alpha-Est9; CRMP; Sodh-1; Taf7; agt; CG1979; Madm; Rpn5; CG1236; NPFR; CG17917; CG17919; CRAT; godzilla; Dmtn; Neurochondrin; CG15186; CG1965; MAGE; CG14609; CG2616; CG3014; CG2993; CG17816; CG10445; CD98hc; CG2781; CG2698; CG10903; CG11035; CG7878; Arl8; CG18249; CG9636; CR18228; Rheb; Obp83g; Nlg1; CG31248; sgll; CG31548; Pcmt; puc; wtrw; CG42675; CG42724; SmD2; Sec23; twr; lncRNA:CR43685; lncRNA:CR44936; lncRNA:CR45551; rn; lncRNA:CR46225</i>
Q4	3R:15264278-17374278	307	<i>ea; Sb; Dhfr; Pxd; Akt1; mtSSB; Fkbp39; cher; AhcyL2; CCT3; Rh6; Dad; CG10340; Hel89B" Manf; CG4338; CG14866; CG6236; mRpL9; CG6136; h-cup; CG6125; CG6118; AOX2; CG9590; blp; GATAe; CG8927; CG10311; Der-2; Gr89a; CG14892; CG17560; CG14907; CG8907; AdSL; CG5225; Keap1; CG5246; TwdlW; Mur89F; tara; Sulf1; Pak3; modSP; CG31287; CG31296; CG31301; sra; CG42232; CG42342; Rbp; Det; CG44013; lncRNA:CR44133; lncRNA:CR45635; lncRNA:CR45639; lncRNA:CR45641; AOX1</i>
Q5	3R:19094278-19344278	32	<i>Cyp12a4; unc79; CG31475</i>
Q6		21	<i>r-l; Mvl; Dhc93AB; Calx; dmr93B; HHEX; G31189; CG31207</i>

	3R:20934278-21094278		
Q7	3R:26004278-26634278	78	<i>Ald; RpL34a; ppk15; CG5432; CG14238; CG5455; CG42498; lncRNA:CR42765; CG43117; tx; scrib; lncRNA:CR43434</i>
Q8	3R:26974278-30404278	472	<i>janA; kay; Mlc2; ncd; Jon99Cii; Jon99Ciii; spz; stg; Klp98A; Dhc98D; side; Pcd; Noa36; Nepl21; mino; Usp1; Rpn2; TwdlC; CG6295; CG5909; CG5984; CG14260; mrt; CG3368; bigmax; CG3339; CG13972; CG5639; dsd; unc80; CG12428; mRpS22; mfrn; Nep7; Gfat2; snu; Nepl18; Nepl19; CG11841; Atg14; CG14511; rdog; CG11898; CG11951; CG14506; CG2321; Trc8; CG7568; ppk20 ppk30; Obp99a" Obp99c; fig; CG7834; CG7789; neo; CG15517; CG15515; Vps13B; Sas-6; CG15525; Nph; CG7912; CG7943; CG7946; CG9733; CG31029; CG31036; Cisd2; ALiX; alph; CG42558; Bet5; yata; CG42813; sro; spg; asRNA:CR44046; lncRNA:CR44953; tau; mil; CG4633</i>

Table 3.2:

Significant QTLs and all the genes that are differentially expressed within each QTL for the **condition** group.

Gene Name	Location on Chrm 3R (bp)	Log ₂ (FC)	Adjusted P-value	Differentially Expressed
AnxB9	21065241-21070917	-0.17	0.17 (con) 0.01 (pool)	Yes (pool)
Alp5	21005160-21006990	0.31	0.27 (con) 0.48 (pool)	
Atpalpha	20948714-20976239	0.09	0.18 (con) 0.01 (pool)	Yes (pool)
Calx	20978183-21015613	0.32	0.01 (con) 0.06 (pool)	Yes (con)
CG12278	21044068-21044905	-0.33	0.65 (con) 0.91 (pool)	
CG17278	21021044-21024163	-0.01	0.98 (con) 0.87 (pool)	
CG17279	21050014-21050975	-0.11	0.86 (con) 1.00 (pool)	
CG31189	21044898-21045842	0.35	0.03 (con) 0.31 (pool)	Yes (con)
CG31191	20909213-20945530	0.17	0.47 (con) 0.17 (pool)	
CG31207	21046032-21047027	0.36	0.00 (con) 0.00 (pool)	Yes (both)
CG43446	20986504-20987936	0.04	0.97 (con) 0.87 (pool)	
CG5697	21015949-21017120	-0.01	1.0 (con) 0.37 (pool)	
dmrt93B	21074007-21076798	-3.43	0.01 (con) 0.78 (pool)	Yes (con)
CG7079	21047462-21048829	0.06	0.84 (con) 0.03(pool)	Yes (pool)
Cortactin	21061680-21064471	-0.20	0.11 (con) 0.00 (pool)	Yes (pool)
Dhc93AB	21024580-21042499	-0.55	0.027 (con) 0.01 (pool)	Yes (both)
HHEX	21076862-21080926	-2.99	0.01 (con) 0.09 (pool)	Yes (con)
lncRNA:CR44178	21088792-21095744	-0.01	0.99 (con) 0.29 (pool)	
Mvl	21051383-21060800	-0.31	0.00 (con)	Yes (both)

			0.00 (pool)	
r-1	21070557-21073319	-0.32	0.02 (con) 0.24 (pool)	Yes (con)
RhoGAP93B	21083043-21097827	-0.11	0.30 (con) 0.03 (pool)	Yes (pool)
Rlip	21017684-21020753	0.11	0.46 (con) 0.91 (pool)	

Table 3.3:
Q6 genes information. See results for more details

CHP 4 – FUNCTIONALLY VALIDATING CANDIDATE GENES INFLUENCING IN THERMAL TOLERANCE *D. MELANOGASTER*

Authors: **Patricka Williams-Simon**, Camille Oster, Kevin Middleton, Troy Zars,
and Elizabeth King

4.1 Abstract:

Genome-wide studies have great resolution to identify candidate genes, however, in most cases, most of these remain speculative, because of the lack of functional validation. In this chapter, we take a two-pronged approach to validate candidate genes that influence thermal tolerance in the *Drosophila* Synthetic Population Resource (DSPR). We performed Structural Variant (SV) analysis to identify any structural variants that were located within our genomic region of interest. We found 3 SVs that are located within our QTL. We also screened 20 UAS-Gal4-RNAi mutants that were expressed in different tissue types to functionally validate candidate genes. We identified 3 large effect causal variants. Our novel findings expand on previous literature, and we recommend further confirmatory validation studies.

4.2 Introduction

Functionally validating candidate genes that influence complex traits is of great importance if we want to identify causal genes. However, validating genes that influence complex traits has been a daunting task, and therefore many genetic studies surrounding complex traits have identified relatively few causal variants (Ding *et al.* 2016; Highfill *et al.* 2017; Gaspar *et al.* 2020). Identifying genes that influence complex traits will ultimately lead to connecting the dots between cellular processes all the way up to population dynamics. For example, identifying genetic variants answers fundamental questions like whether a gene is cis- or trans-regulated (Rockman and Kruglyak 2006), whether the trait is dormant (Lemos *et al.* 2008), what molecular pathways are the genes involved in (Arnold *et al.* 2012; Wheeler *et al.* 2019), and what other phenotypes are the genes controlling (pleiotropy) (Solovieff *et al.* 2013; Visscher and Yang 2016). One of the ways to characterize these processes is through first identifying the genes that are involved, and then making the connection to molecules, systems, and eventually understanding how standing genetic variation affects population levels.

Complex traits are typically very polygenic (Pritchard *et al.* 2010; Chakravarti and Turner 2016; Shi *et al.* 2016; Boyle *et al.* 2017), where multiple genes (sometimes thousands) influence a particular trait. Height is the typical

model example when discussing polygenic traits. For example, Wood *et al.* (Wood *et al.* 2014), performed a genome-wide association study and identified 697 loci that influence height, however that number only explained about 20% of the phenotypic variation. Again suggesting that complex traits, like height, are very polygenic with mostly small effect variants (Goldstein 2009). One explanation/theory is that as the number of genes that influence a trait increases, then the genetic contribution to each gene becomes correspondingly smaller, which is an extension to the “infinitesimal model” (Barton *et al.* 2016, 2017). Additionally, complex traits are tightly correlated, a phenomenon known as linkage disequilibrium (LD), where favorable genes that are of close proximity on the chromosome are inherited together (Smith and Haigh 1974). Another explanation could be that many complex traits are regulated by non-coding variants that are presumed to affect gene regulation (Welter *et al.* 2014; Pickrell 2014; Li *et al.* 2016). There are reports from Fraser *et al.* (Fraser and Xie 2009) and Li *et al.* (Li *et al.* 2016), that gene regulation contributes significantly to the variation that is observed in complex traits. Taken together, these results might seem daunting, but over the past few years, researchers have pushed the envelope to develop tools to put together some of these missing pieces (Yang *et al.* 2011; Turner 2014; Stern 2014; Ding *et al.* 2016; Schaid *et al.* 2018).

Scientists have used many different methods to try to identify & validate candidate genes. These methods include, RNA-seq (Williams-Simon *et al.* 2019; Ng’oma *et al.* 2020), gene ontology (GO) (Ashburner *et al.* 2000; Ng’oma *et al.* 2020), structural variants (SV) analysis (Chakraborty *et al.* 2019), evolve and

resequence (Turner *et al.* 2011). The major goal of these approaches is to further narrow down associated regions in order to identify specific genetic variants hypothesized to contribute to complex traits (Najarro *et al.* 2015; Highfill *et al.* 2017; Williams-Simon *et al.* 2019; Ng'oma *et al.* 2020) and reviewed in (Schaid *et al.* 2018). Often, researchers already have evidence that there is an association of a genotype and a phenotype at a specific genomic region (loci) and assume at least one candidate variant is located within that region (Schaid *et al.* 2018). Below we briefly elaborate on some of these approaches, **1)** RNA-seq (reviewed in (Stark *et al.* 2019), which identifies differentially expressed genes between cohorts, **2)** Gene Ontology (Ashburner *et al.* 2000), which uses so-called GO terms to associate between gene products, **3)** structural variant analysis (Chakraborty *et al.* 2019) and reviewed in (Frazer *et al.* 2009; Eichler *et al.* 2010), which identifies variants (e.g indels, TEs, and inversion), which are known to contribute to variation in complex traits (Sebat *et al.* 2004; Redon *et al.* 2006; Human Genome Structural Variation Working Group *et al.* 2007); (Frazer *et al.* 2009; Eichler *et al.* 2010; Chakraborty *et al.* 2018), and **4)** Evolve and resequence (Turner *et al.* 2011; Turner and Miller 2012; Long *et al.* 2015; Barghi *et al.* 2017) is a type of artificial selection where the researcher will select for a divergent population over several generations and perform genomic sequencing to identify changes in either gene expression or allelic frequencies. It is important to recognize these methods only allow one to identify hypothesized causal variants, but it is still critical to perform validation to conclusively show the variant is causal.

Only through functional validation can we truly consider a gene to have a causal variant for a particular trait. Genome engineering tools like the CRISPR-Cas9 (Jinek *et al.* 2012) and RNAi (Kuwabara and Coulson 2000) systems allow for gene knockdown within endogenous backgrounds, therefore one could measure the effects of a specific gene on a particular phenotype within any genomic background. The Gene knockdown approach is used extensively in classical genetics, in particular in *D. melanogaster*. In fact, the vast majority of what we know about the genetic basis of many complex traits stems from the classic gene mutagenesis approaches (Zars *et al.* 2000c; Nishi and Nanjo 2011; Baggett *et al.* 2018; Cleves *et al.* 2020). The field of quantitative genetics has leveraged the strengths of these approaches to functionally validate candidate genes. For example, Turner (Turner 2014) proposed using quantitative complementation (QC), which is an older technique to identify and validate causal genes that influence complex traits. Around the same time, Stern (Stern 2014) also proposed using a technique known as Reciprocal Hemizyosity Test (RHT) that has shown successful applications in empirical studies (Ding *et al.* 2016; Gaspar *et al.* 2020). Both QC and RHT have integrated the relatively new CRISPR technologies into classical approaches to functionally validate candidate variants and therefore identify causal variants. There is also the GAL-UAS/RNAi system that allows specific or ubiquitous disruption in the expression of a particular gene. In (Perkins *et al.* 2015), Perkins *et al.* generated a genome-wide collection of RNAi stocks of over 11,400, which covers more than 71% of the fruit fly genome. Essentially, creating a resource of the fly population where a single gene is disrupted within a

stock, giving rise to thousands of stocks, where one gene is disrupted. This allows for the investigation of what a particular gene would have on a phenotype whenever the gene is disrupted. The RNAi system has been functionally validated in paradigms (Neely *et al.* 2010; Awata *et al.* 2016; Highfill *et al.* 2017; Smith and Macdonald 2020).

In this chapter, we integrated SV analysis with the TRIP UAS-Gal4-RNAi to functionally validate the candidate genes we identified from Quantitative Trait Loci (QTL; see chapter 3 for more details). Briefly, in Chapter 3 we identified 8 QTLs (Q1-Q8), of those Q6 being the most significant with 21 genes within the QTL. We obtained the UAS-RNAi lines, each with a disruption in 1 gene, crossed with specific Gal4 lines, and measured thermal tolerance using the hot plate. From the screen we found 3 genes that were significant. We also investigated whether there were SVs present within the 21 genes and identified a few SVs. Taken together, our dataset provides a novel set causal genes

4.3 Methods:

Fly husbandry – All fly stocks were raised on the standard cornmeal-yeast diet, and kept at 18° C, with 60% relative humidity, on 12 hours light-dark cycles. Stocks were kept in those conditions for the duration of the experiments. Two weeks before the behavioral assays, about 6 males and 10 female flies were flipped onto a new flask (Fisher Scientific, Cat. No.:AS-355) of food and allowed to mate at 25°C and lay eggs. After seven days, all the adults were removed from the

flask to ensure that only the F1 generation remained in the flask. Five – Seven days post-oviposition, adult flies were anesthetized on ice, and we collected 60-80 female flies for the behavioral assay. These female flies were placed onto two separate new food vials in groups of 30-40 individuals per vial. The flies were then allowed to recover for at least 24 hrs before the assay. Flies included in the assay were 6–8 days post-oviposition. For each RIL we phenotyped a minimum of 40 individuals, with an average of 66, and a range of 40-117. A total of 10-15 RILs were set up and phenotyped in any given week.

RNAi Lines and Crossing Design – We sourced the UAS-RNAi transgenic fly lines from the TRiP BDSC ([Transgenic RNAi Project](#); Bloomington Drosophila Stock Center; (Perkins *et al.* 2015)). BDSC reported that the TRiP RNAi lines might have a low dose of *Wolbachia pipientis*, therefore we treated the flies with tetracycline (Sigma-Aldrich, Cat No.:T4062-5G, Batch: SLCF9393). We followed a tetracycline treatment protocol from Hoffman *et al.*, (Hoffmann *et al.* 1986). We incorporated 0.3 mg/mL final concentration of tetracycline into an agar base fly food recipe and allowed the flies to remain on that food for two generations flipping them onto new agar food after 2 weeks. One of the lines did not survive during the tetracycline treatment (BDSC stock No.: 63713). After the treatment, we switched the population back to our ‘normal’ fly food recipe and allowed the lines to re-acclimate to the food for several generations before beginning experiments. We did not confirm that the lines were *Wolbachia* free, but, because this is considered general practice in the field (Hoffmann *et al.* 1986; Poinot and Merçot 1997; Bordenstein and Werren 1998; Dobson *et al.* 2002; Fry *et al.* 2004), therefore we

are confident the lines are free from Wolbachia. The TRiP did not have RNAi lines for all 21 genes in the Q6 interval, either the RNAi lines were not created for that gene because much is not known about that particular gene, or it created a lethality. We, however, obtained RNAi lines for 14 genes from the BDSC, some of which have a couple or a few RNAi lines (**Table 4.1**). We checked each UAS line for off-target effects and none were reported to have off-target effects (<https://fgr.hms.harvard.edu/up-torr>). In total we screened **20** UAS-RNAi lines using to 3 different promoters: first ubiquitously expressed using an *actin* (*Act5C-Gal4*) driver (BDSC–3954: *y1 w**; *P{Act5C-GAL4}17bFO1/ TM6B, Tb1*), the second pan neuronally expressed using *embryonic lethal abnormal vision* (*elav*) driver (BDSC–25750: *P{w[+mW.hs]=GawB}elav[C155]w[1118]*; *P{w[+mC]=UAS-Dcr-2.D}2*), and third pan neuronally expressed using *neuronal synaptobrevin* (*nSyb-GAL4*) driver (BDSC–51941: *y[1]w[1118];P{y[+t7.7] w[+mC]=nSyb-GAL4.P}attP2*). All the UAS-RNAi lines were in either P2 (Chr 3) or P40 (Chr 2) background, which we used as controls for quantifying thermal tolerance.

We duplicated the lines from the parental stock and allowed them to mate. On day 3, we cleared the vials of all adults, then on days 9-12, we collected 3 males (*Gal4*) and 5 females (*UAS-RNAi*) to create *Gal4-UAS-RNAi* F₁ progeny. The *Act5C-Gal4* line has a tubby (*TM6B, tb1*) phenotypic marker, which we selected against at the pupae stage of development in the F₁ because tubby is a pupae marker. Therefore, to make a fair comparison we also transferred all lines on days 19-22 to control for transferring the tubby lines. On days 20-23 We anesthetized the flies using CO₂, collected 10 F₁ females (0-1 day post-

oviposition), placed them on a new vial with food, and allowed them to stay overnight before measuring thermal tolerance.

Measuring Thermal Tolerance – We measured the time it takes for individuals to become incapacitated. Flies were individually placed into small clear cylinder tubes that are 5 x 65 mm (diameter x length; monitor tubes, TriKinetics Inc USA) and plugged with pipe cleaners on both ends to prevent the flies from escaping. We placed the tubes onto a custom-built aluminum plate that heats and cools rapidly. We are able to measure up to 33 individuals simultaneously. The monitor tubes are placed on top of PELT-PLATE, which is controlled by PELT-5 (Sable Systems International, Inc), and the fly's activity is tracked by a camera using a python script in real-time. The duration of each measurement is 10 mins, and the PELT-PLATE is set to 41°C to reflect the temperature used in the box (see chapter 3 for more details). A thermocouple is also placed into a monitor tube for each run to monitor the accuracy of the temperature that is radiating from the PELT-PLATE and the temperature that is within the monitor tubes. We performed a block type experiment for the DSPR founders, RILs, and RNAi lines. For the founders and RILs, we had 2 biological replicates, assaying 33 individuals each time. For the RNAi lines, we had 3 biological replicates and assayed 10 individuals each replicate (**data not shown**). Before each trial, we confirmed that the aluminum plate was reflecting room temperature, before placing the monitor tubes into the wells, and that all the monitor tubes were placed in the correct positions. We would like to note that all lines were randomly chosen for the first replicate and kept that order for the second or third replicates. We also randomly choose the

wells into which the monitor tubes were placed. Ideally, we would have measured the validation lines with the same apparatus that we used to measure thermal tolerance in Chapter 3. But unfortunately, we were not able to assay these individuals using the heat box, which was used when we quantified thermal tolerance in the 741 DSPR RILs, because the heat box was no longer operational, however, we were able to keep our measuring parameters the same. For comparison, we measured a subpopulation of 50 RILs from ones that were assayed in the heat box, which was distributed across the range (**data not shown**) using the heat plate apparatus. Thermal tolerance is also more straightforward to measure than learning and memory, thus we were able to construct a higher throughput system.

To compute significance on the dataset we linear model (*lm*), followed by an *ANOVA* to compare all three UAS-Gal4-RNAi cross and the background control. The *lm* model is below, where incapacitation is the *incapacitation* time and *set_id* is the line cross identifier. We also calculated the False discovery rate (FDR) on our adjusted p-values, see model below, where *nn* is the p-value. The FDR measures the proportion of the data that is expected to be false (Benjamini and Hochberg 1995; Benjamini and Yekutieli 2001). We then did a manual search on FlyBase (<https://flybase.org/>), using gene ontology terms to investigate previous findings of these genes being implicated with thermal tolerance.

lm(incapacitation ~ set_id, data= Gene)

p.adjust(nn, method="BH")

Structural Variants – SV, for example, insertions, deletions, and transposable elements contribute to the phenotypic variations in complex traits (Frazer *et al.* 2009; Eichler *et al.* 2010; Chakraborty *et al.* 2018). Chakraborty *et al.* (Chakraborty *et al.* 2019) used the DSPR to identify genome-wide SVs in both populations A and B. To determine if there were any SVs located within the genes at our most significant QTL (Q6; Release 6 coordinates: 26004178-26634378) we followed the Chakraborty *et al.* (Chakraborty *et al.* 2019) protocol. To summarize, this approach identifies TEs, insertions, and deletions (indels) that would otherwise be invisible by genomic scans, like QTLs and GWAS, because they are not easily identified by short-read data and variant calling software. These types of variants can be major contributors to the variation that we see in complex traits (Frazer *et al.* 2009; Eichler *et al.* 2010). To ensure that we captured all the SVs located in the QTL region, we included any genes that had any portion within the confidence interval, even if the entire gene did not fall within the QTL region.

Data Analysis and statistics: All Data analyses and statistics were performed in R 3.6.2 and lower, python 3.9.0, and deeplabcut 2.1.10.4.

4.4 Results:

We quantified thermal tolerance in 20 UAS/Gal4-RNAi mutant lines (1969 individuals) and 2 control backgrounds. We measured a minimum of 21 individuals

and a maximum of 60 for a total of 10 mins. at 41°C. We found that there is a wide phenotypic variation in the measured lines with an average of 108.59 secs (range = 0.954–348.004). We ran a linear model on our data set to compare all groups (crosses) within a gene and found 3 significant UAS/Gal4-RNAi lines (**Table 1**). These three lines have a disruption in gene expression for *atpalpha*, *dmrt93B*, and *rhoGAP93B*. For *atpalpha* (**Fig 2a**), we found when the mutant RNAi line is crossed with Act-Gal4 (3954) has an average incapacitation score of 163 secs., whereas a lower average incapacitation score when crossed with both pan neuronal Gal4 lines, with *nSyb* (51941 = 122 secs.) having a slightly higher average score. Line crossed with *elav* (25750 = 90 secs.) has a very similar average score to the control (36304). Compared to the control, the data suggests that 51411-3954 is most significant. For *dmrt93B*, we were only able to complete two crosses (**Fig 2b**). In this comparison the control (36303) had a higher incapacitation score than the UAS-RNAi that crossed with the *elav*, pan-neuronal promoter, suggesting that *dmrt93B* is influencing thermal tolerance. For *rhoGAP93B*, both pan-neuronal crosses (25750 = 125 secs. and 51941 = 102 secs.) and control has higher incapacitation scores than the ubiquitously expressed line (3954 = 76 sec.; **Fig 2c**). These three genes are also significantly differentially expressed in our pool RNA-seq data set, *atpalpha* was upregulated (positive fold change) while *dmrt93B* and *rhoGAP93B* were both downregulated (see chapter 3 for more details).

We then performed a manual search on FlyBase to investigate previous findings of these genes being implicated with thermal tolerance. Of the three genes, *atpalpha* was characterized by two separate groups to be involved in the

response to temperature stimulus (Feng *et al.* 1997; Palladino *et al.* 2003), to have a jump response – i.e. rapid response to an external stimulus – (Ashmore *et al.* 2009), and predicted to be involved in tissue homeostasis (Palladino *et al.* 2003). The main biological functions that have been characterized for *dmrt93A* are regulation of transcription and sex differentiation (Zhu *et al.* 2000; Gaudet *et al.* 2011). *RhoGAP93B* is involved in regulating GTPase activity and mediates axon guidance (Lundström *et al.* 2004; Hu *et al.* 2005).

We also performed a structural variant analysis (Chakraborty *et al.* 2019) on our Q6 BCI to identify any structural variance located within the region. Structural variance is known to cause variation in complex traits (Fraser and Xie 2009; Eichler *et al.* 2010; Chakraborty *et al.* 2018), and therefore identifying them is a critical step in dissecting complex traits. We found 3 SVs in our QTL interval (**Table 4.2**), two insertions, and 1 deletion. These Svs were located within three genes, *annexinB9* (*anxB9*), *Na/Ca-exchange protein* (*calx*), and *doublesex-Mab related 93B* (*dmrt93B*). Both *anxB9* and *dmrt93B* were differentially expressed in our pool group (**Fig 4.1**) and *calx* was significantly expressed in our condition group (**Table 3**; see chapter 3 for more details).

4.5 Discussion:

In this chapter, we functionally validated candidate genes that influence thermal tolerance in fruit flies. We measured thermal tolerance in a total of 20 RNAi mutants, each with a specific gene disrupted, from this approach we identified 3

genes that have a causal effect on thermal tolerance. We also identified 3 structural variants that are present in our data set.

The TRIP (Perkins *et al.* 2015) is a powerful resource that provides the ability to screen more than 71% of the genes in the fly genome because there is a mutant line for almost all the genes. We took advantage of this resource and ordered lines with gene disruption for the genes that fall within our QTL6. In total, we screened 20 lines that we crossed to two pan-neuronal (*elavGAL4* and *nSyb*) fly lines and 1 ubiquitously expressed line (*act-Gal4*). We then measured thermal tolerance in those lines together with controls and computed significance using a linear model. We found three lines that were significant. Those lines had a knockdown expression of *atpalpha*, *dmrt93B*, and *rhoGAP9*. Comparing all 3 genes to the control we found that the mutant line for *dmrt93B* and *rhoGAP9* had the lower average incapacitation score, suggesting that in the absence of those genes thermal tolerance is lower, which indicates that they are necessary for high thermal tolerance.

More specifically, whenever the *atpalpha* fly mutant was crossed with a pan-neuronal Gal4, the average incapacitation score was lowest, compared to being crossed with a ubiquitous expression Gal4 line and control. This result indicates that *atpalpha* influences thermal tolerance in the nervous system. Our results are supported by Feng *et al.* (Feng *et al.* 1997), Palladino *et al.* (Palladino *et al.* 2003), and Ashmore *et al.* (Ashmore *et al.* 2009). They reported that the *atpalpha* response to temperature stimulus (Feng *et al.* 1997; Palladino *et al.* 2003) is part of the jump response to external stimulus (Ashmore *et al.* 2009). The *atpalpha*

gene is also predicted to be involved in tissue homeostasis (Palladino *et al.* 2003). Strikingly, we observed some of the phenotypes (hyperexcitability and seizure-like activity) reported by Palladino *et al.* (Palladino *et al.* 2002), in our *atpalpha* mutant. *Atpalpha* encodes Na⁺ /K⁺ ATPase subunit, which gradients determine the resting potential of cells and drive important secondary processes, like thermal tolerance (Palladino *et al.* 2002). Many of these sodium pumps are evolutionarily conserved reviewed in (Blanco and Mercer 1998; Mobasheri *et al.* 2000), indicating its importance in biological processes. In fruit flies, nematodes, and mice, mutations of ATP subunits result in paralytic phenotypes, pharyngeal function, paralysis of extremities respectively (Schubiger *et al.* 1994; Magyar *et al.* 1994; Davis *et al.* 1995; Molthagen *et al.* 1996; Shima *et al.* 1998). All three pieces of evidence suggest that the *atpalpha* gene plays a critical role in the neuromuscular junction processes (NMJ), which directly affects both the nervous and muscular systems. Dysfunction in the *atpalpha* gene is also reported to affect humans. An older study, Renkawek *et al.* (Renkawek *et al.* 1992), hypothesized that an impairment of pump activity caused neonatal seizures associated with spongiform encephalopathy and low Na⁺ /K⁺ ATPase, which resulted in the death of siblings. Temperature causes an over-excitability of *atpalpha* (Palladino *et al.* 2002), which suggests that the NMJ system becomes overwhelmed causing paralysis and therefore low thermal tolerance. However, contrary to these studies, our data suggest that the reduction of *atpalpha* only in the neurons increases thermal tolerance when compared to the control. This could indicate that the *atpalpha* on the muscle side of the NMJ system

is either compensating for the neurons and therefore causes an increase in thermal tolerance. This is an area for continued empirical investigation.

Although, *dmrt93B* and *rhoGAP93B* have not been directly implicated in thermal regulation before they are involved in the regulation of transcription (Gaudet *et al.* 2011) and axon signal transduction (Lundström *et al.* 2004; Hu *et al.* 2005) respectively. Our results also suggest that thermal tolerance is controlled by specific genes in specific tissue types. For example, in our *atpalpha* mutant pan-neuronal had the lowest average thermal tolerance, whereas in the *rhoGAP93B* mutant ubiquitous expression had the lowest phenotype. Recently, Jorgensen *et al.* (Jørgensen *et al.* 2020) measured thermal tolerance in 5 different species of *Drosophila* and found that neuronal dysfunction correlates with critical thermal maximum, which supports our findings. There is very little literature indicating that *dmrt93B* and *rhoGAP93B* are involved in thermoregulation, therefore these results present novel findings that need to be followed by more confirmatory findings.

Structural variants (e.g. insertions, deletions, and transpositions) within genomic regions are known to play a critical role in the genetic variation that is observed in complex traits (Fraser and Xie 2009; Eichler *et al.* 2010; Chakraborty *et al.* 2018), however, they are difficult to identify through standard methods (Alkan *et al.* 2011). More recently, there have been several proposed models to help lessen this problem including (Chakraborty *et al.* 2019), which we took advantage of to identify any structural variants within our most significant QTL. We identified 3 variants within Q6, two insertions, and 1 deletion, and all 3 are located within 3 different genes. We found an insertion in *calx* and *anxB9* and a deletion in

dmrt93B, which was also significant in the RNAi screen. All three genes are thought to play a role in some form of cell homeostasis. For example, *calx* is known to respond to endoplasmic reticulum stress (Chow *et al.* 2013) and *anxB9* helps maintain cell polarity (Tjota *et al.* 2011). All three genes are also differentially expressed in either pool or condition groups (see chapter 3).

In this chapter, we used an integrative approach to validate causal genes that influence thermal tolerance in fruit flies. Despite few experimental design limitations, we present these 5 (*atpalpha*, *anxB9*, *calx*, *dmrt93B*, and *rhoGAP93B*) genes to be causal variants that have a large effect on the thermal tolerance phenotype in fruit flies. We recommend follow-up confirmatory functional studies. However, collectively these results are novel findings that are important to elucidate the genetic basis of thermal tolerance.

4.6 Figures:

Fig 1:

Q6 BCI interval. Red dots represent a gene. Brightly colored dots mean that the gene is significantly differentially expressed, which is also labeled. The horizontal line represents the baseline level of expression while the vertical dashed lines represent the lower and upper BCI for Q6.

Fig 2a:

Thermal tolerance measurements for UAS-RANi mutant (*atp α*) lines crossed with Gal4 lines. Lines labeled 25750 and 51941 are expressed under pan neuronal promoters, *elav* and *nSyb* respectively, whereas 3954 is ubiquitously expressed using *Actin* as a promoter. The line 36304 is the UAS-RANi background control. Colored dots represent an individual within a cross. Colored boxes are group means and black lines denote ± 1 SE of the mean.

Fig 2b:

Thermal tolerance measurements for UAS-RANi mutant (*dmrt938*) lines crossed with Gal4 lines, labeled 25750 (*elav*), which drives pan neuronal expression. The line 36303 is a background control. Colored dots represent an individual within a cross. Colored boxes are group means and black lines denote ± 1 SE of the mean.

Fig 2c:

Thermal tolerance measurements for RANi mutant (*RhoGAP93B*) lines crossed with Gal4 lines. Lines labeled 25750 and 51941 are expressed under pan neuronal promoters, *elav* and *nSyb* respectively, whereas 3954 is ubiquitously expressed using *Actin* as a promoter. The 36303 line is the UAS-RANi background control. Colored dots represent an individual within a cross. Colored boxes are group means and black lines denote ± 1 SE of the mean.

4.7 Tables:

Gene Name	Gene BDSC ID	Adjusted P-value
Alp5	57526	0.19
AnxB9	38523	0.41
AnxB9	58111	0.86
Atpalpha	33646	0.40
Atpalpha	32913	0.86
Atpalpha	51411	0.00
Calx	28306	0.69
Cortactin	32871	0.08
Cortactin	44425	0.86
Dhc93AB	51511	0.86
dmrt93B	27657	0.79
CG7079	61174	0.40
CG17279	60071	0.86
dmrt93B	50656	0.00
HHEX	60119	0.40
CG31207	60074	0.86
Mvl	55316	0.69
r-1	55183	0.45
RhoGAP93B	31167	0.69
RhoGAP93B	35027	0.01

Table 1:

List of the 20 UAS-RNAi lines we measured for thermal tolerance. Bolded rows are significant. See results section for more details.

Location on Chrm (bp)	Gene Name	Type of SV	Founder	SV ID
3R:20997328- 20997362	<i>calx</i>	Insertion	A6	0000011920
3R:21067722- 21067722	<i>anxB9</i>	Insertion	A3	0000008970
3R:21074479- 21074712	<i>dmrt93B</i>	Deletion	A5	0000017347

Table 2:

List of structural variance that are located within Q6 BCI. See results for more details.

CHP 5 – CONCLUSION.

This dissertation provides novel insights into the genetic underpinnings of learning, memory, and thermal tolerance in the fruit fly model. I specifically used a multi-parental fruit fly population, known as the *Drosophila* Synthetic Population Resource (DSPR), which consists of hundreds of Recombinant Inbred Lines (RILs). I combined quantitative and classical genetics, with behavioral, molecular, and computational biology to dissect the genetic basis of learning, memory, and thermal tolerance. This integrative approach allowed me to expand on existing studies and present novel findings, which I expand on below.

In chapter two, I quantified place learning and memory in 741 (RILs) and identified 16 novel Quantitative Trait Loci (QTL) that influenced learning and memory. I then combined my genome scan data with RNA-seq to identify significantly expressed genes between pooled high and low RILs that were quantified for learning and memory. In addition, I compared the data collected from the RNA-seq to an eQTL study to identify genes that were cis-regulated. Taken together, we present novel genomic regions that elucidate the genetic architecture that influences place learning and memory in fruit flies. For this chapter, I propose a few future directions. First, I propose to functionally validate the list of 9 candidate genes in a subset of the 741 RILs that were measured, by disrupting gene expression of those genes within the DSPR background, using either RNAi or CRISPR, and then re-phenotype those RILs for learning and memory. If a RIL had a high score for the initial measurement, but the learning or memory score declines

considerably in the mutant, then that would suggest that the disrupted gene is influencing the phenotype. Second, I propose to use the gene expression data, which we already have, to **1)** generate gene ontology (GO) terms to then investigate possible pathways the 9 candidate genes are integrated into. **2)** identify genes that were significant genome-wide and also alternatively spliced. **3)** determine which genes are cis- or trans-regulated through expression QTL approaches. These suggested approaches will identify molecular pathways that influence learning and memory.

In chapter three, I measured thermal tolerance in 741 RILs and identified 8 (Q1-Q8) genomic regions that influence thermal tolerance in the DSPR, including one very large effect locus on chromosome 3R. I also identified a large set of genes that are differentially expressed between cohorts that differ in high vs low thermal tolerance. We integrated our finds from these RNA-seq findings with our QTLs to identify multiple candidate genes across the genome. Of these genes, we presented the 21 genes that fall within Q6 as possible candidate genes. Our findings from this chapter build upon previous work that aims to identify the genetic basis of thermal tolerance. Below, I propose a few follow-up studies for this chapter. I recommend further investigation into functionally validating the 21 genes in the Q6 interval. Validation can be done through a two-step process, first using classical genetics approaches, like TRIP-RNAi lines to screen the 21 genes in the region effect on the phenotype, then use quantitative complementation (QC) approach to further validate the narrower list of genes. The theory behind QC is to create a mutation in a gene of interest, then back cross that line for several

generations – with the DSPR founders in our case – while selecting a physical marker for the individuals with the mutation. Once the lines are introgressed, for about 6 generations the region of interest should be small enough, then re-measure thermal tolerance in those lines. One of the major differences between using the TRIP-RNAi and the QC approaches is that we would create the gene disruption in the DSPR background, allowing for a more equal comparison. These suggested follow-up studies will identify causal genetic variants that influence thermal tolerance.

In Chapter Four, we functionally validated candidate genes that influence thermal tolerance using two parallel but complementary approaches. First, I performed a structural variance analysis on the 21 candidate genes that are located within our most significant QTL, then measured thermal tolerance in RNAi lines that have a knock-down expression for most of the candidate genes. I identified 3 SVs within the QTL region through a structural variance approach. For this study, I present these 5 (*atpalpha*, *anxB9*, *calx*, *dmrt93B*, and *rhoGAP93B*) genes to be causal variants that have a large effect on the thermal tolerance phenotype in fruit flies. For this project I recommend two next steps, first a proteomics and metabolomics expression profiling, to investigate protein and molecular interactions, and therefore pathways that the genes are connected to. Second, whole-mount in-situ hybridization using specific probes and fluorescent markers to determine in which tissue types expression of those genes are localizing. This can be complemented with qRT-PCR for mRNA expression profiling in specific tissue types.

These findings have many broad implications and could be applicable to many fields. Here, I list a few of these broader implications. First, in the field of evolution, our results are fundamental to help close the knowledge gap on the genetic underpinnings that contribute to learning, memory, and thermal tolerance in a natural population and presents novel knowledge to help us comprehend how these traits have evolved and continue to evolve in rapidly changing environments. Second, in the field of ecology – specifically conservation – and as climate change continues to be one of the leading causes of species extinction our findings of the genes that influence high learning, memory, and thermal tolerance could be used as a way of predicting which species survive extreme temperatures. Third, because fruit flies and humans have 75% of the same genes that cause diseases, we can potentially use our findings as a building block to better understand the genetic basis of human-related diseases. For example, *ATPalpha* is linked to seizures in humans (Renkawek *et al.* 1992) and it is one of the genes that we identified and validated to play a role in thermal tolerance. Interestingly, before the flies within our thermal tolerance assay became incapacitated, they exhibited a seizure-like phenotype. In addition, and despite many decades of work to elucidate the genes that are involved in the broad spectrum of learning and memory disorders and ages that affect humans, there is still a major gap in our knowledge of the causative genes. Perhaps, our novel findings of genes that influence learning and memory in fruit fly present an opportunity to better understand the genetics that influences those disorders that might eventually lead to therapeutics. There is a lot more work to be done to make direct correlations between genes

that we find in fruit flies to human disorders, but our data presents a starting point for further investigation.

To summarize, in this dissertation I successfully identified genetic variants that influence learning, memory, thermal tolerance in a fruit fly population. I also functionally validated genes that influence thermal tolerance and presented several novel genetic variants as causal. This work expands on previous studies to help elucidate the genetic basis of learning, memory, and thermal tolerance and provides a foundation to further understand the functions and pathways that contribute to those phenotypes. These novel findings could have broad implications in the fields of evolution, ecology, and as a starting point for even human genetics and therapeutics.

BIBLIOGRAPHY

- Aceves-Piña, E. O., R. Booker, J. S. Duerr, M. S. Livingstone, W. G. Quinn *et al.*, 1983 Learning and memory in *Drosophila*, studied with mutants. Cold Spring Harb. Symp. Quant. Biol. 48 Pt 2: 831–840.
- Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne *et al.*, 2000 The genome sequence of *Drosophila melanogaster*. Science 287: 2185–2195.
- Addo-Bediako, A., S. L. Chown, and K. J. Gaston, 2000 Thermal tolerance, climatic variability and latitude. Proc. Biol. Sci. 267: 739–745.
- Akalal, D.-B. G., C. F. Wilson, L. Zong, N. K. Tanaka, K. Ito *et al.*, 2006 Roles for *Drosophila* mushroom body neurons in olfactory learning and memory. Learn. Mem. 13: 659–668.
- Alkan, C., B. P. Coe, and E. E. Eichler, 2011 Genome structural variation discovery and genotyping. Nat. Rev. Genet. 12: 363–376.
- Allen, G. E., 1978 *Thomas Hunt Morgan : the man and his science*. Princeton University Press.
- Allen, C. D., A. K. Macalady, H. Chenchouni, D. Bachelet, N. McDowell *et al.*, 2010 A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. For. Ecol. Manage. 259: 660–684.
- Altmüller, J., L. J. Palmer, G. Fischer, H. Scherb, and M. Wjst, 2001 Genomewide scans of complex human diseases: true linkage is hard to find.

- Am. J. Hum. Genet. 69: 936–950.
- Amano, H., and I. N. Maruyama, 2011 Aversive olfactory learning and associative long-term memory in *Caenorhabditis elegans*. *Learn. Mem.* 18: 654–665.
- Andersen, M. K., N. J. S. Jensen, R. M. Robertson, and J. Overgaard, 2018 Central nervous system shutdown underlies acute cold tolerance in tropical and temperate *Drosophila* species. *J. Exp. Biol.* 221.:
- Andrews, S., 2010 *FastQC: a quality control tool for high throughput sequence data*.
- Andrews, S., and Others, 2010 FastQC: a quality control tool for high throughput sequence data.
- Angilletta, M. J., Jr., and M. J. Angilletta, 2009 *Thermal Adaptation: A Theoretical and Empirical Synthesis*. OUP Oxford.
- Angilletta, M. J., T. Hill, and M. A. Robson, 2002a Is physiological performance optimized by thermoregulatory behavior?: a case study of the eastern fence lizard, *Sceloporus undulatus*. *J. Therm. Biol.* 27: 199–204.
- Angilletta, M. J., P. H. Niewiarowski, and C. A. Navas, 2002b The evolution of thermal physiology in ectotherms. *J. Therm. Biol.* 27: 249–268.
- Anttila, K., R. S. Dhillon, E. G. Boulding, A. P. Farrell, B. D. Glebe *et al.*, 2013 Variation in temperature tolerance among families of Atlantic salmon (*Salmo salar*) is associated with hypoxia tolerance, ventricle size and myoglobin level. *J. Exp. Biol.* 216: 1183–1190.
- Arnold, M., D. C. Ellwanger, M. L. Hartsperger, A. Pfeufer, and V. Stümpflen,

- 2012 Cis-acting polymorphisms affect complex traits through modifications of microRNA regulation pathways. *PLoS One* 7: e36694.
- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler *et al.*, 2000 Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25: 25–29.
- Ashmore, L. J., S. L. Hrizo, S. M. Paul, W. A. Van Voorhies, G. J. Beitel *et al.*, 2009 Novel mutations affecting the Na, K ATPase alpha model complex neurological diseases and implicate the sodium pump in increased longevity. *Hum. Genet.* 126: 431–447.
- Aso, Y., and G. M. Rubin, 2016 Dopaminergic neurons write and update memories with cell-type-specific rules. *Elife* 5.:
- Astigarraga, S., K. Hofmeyer, R. Farajian, and J. E. Treisman, 2010 Three *Drosophila* liprins interact to control synapse formation. *J. Neurosci.* 30: 15358–15368.
- Awata, H., R. Wakuda, Y. Ishimaru, Y. Matsuoka, K. Terao *et al.*, 2016 Roles of OA1 octopamine receptor and Dop1 dopamine receptor in mediating appetitive and aversive reinforcement revealed by RNAi studies. *Sci. Rep.* 6: 29696.
- Aylor, D. L., W. Valdar, W. Foulds-Mathes, R. J. Buus, R. A. Verdugo *et al.*, 2011 Genetic analysis of complex traits in the emerging Collaborative Cross. *Genome Res.* 21: 1213–1222.
- Baggett, V., A. Mishra, A. L. Kehrer, A. O. Robinson, P. Shaw *et al.*, 2018 Place learning overrides innate behaviors in *Drosophila*. *Learn. Mem.* 25: 122–128.
- Balfanz, S., T. Strünker, S. Frings, and A. Baumann, 2005 A family of

- octopamine [corrected] receptors that specifically induce cyclic AMP production or Ca²⁺ release in *Drosophila melanogaster*. *J. Neurochem.* 93: 440–451.
- Baracchi, D., A. Cabirol, J.-M. Devaud, A. Haase, P. d'Ettorre *et al.*, 2020
Pheromone components affect motivation and induce persistent modulation of associative learning and memory in honey bees. *Commun Biol* 3: 447.
- Barghi, N., R. Tobler, V. Nolte, and C. Schlötterer, 2017 *Drosophila simulans*: A Species with Improved Resolution in Evolve and Resequencing Studies. *G3* 7: 2337–2343.
- Barrett, T., S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim *et al.*, 2013 NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* 41: D991–5.
- Barton, N. H., A. M. Etheridge, and A. Véber, 2016 The infinitesimal model. *Cold Spring Harbor Laboratory* 039768.
- Barton, N. H., A. M. Etheridge, and A. Véber, 2017 The infinitesimal model: Definition, derivation, and implications. *Theor. Popul. Biol.* 118: 50–73.
- Bassett, A. R., and J.-L. Liu, 2014 CRISPR/Cas9 and genome editing in *Drosophila*. *J. Genet. Genomics* 41: 7–19.
- Batabyal, A., and M. Thaker, 2019 Lizards from suburban areas learn faster to stay safe. *Biol. Lett.* 15: 20190009.
- Baudier, K. M., C. L. D'Amelio, R. Malhotra, M. P. O'Connor, and S. O'Donnell, 2018 Extreme Insolation: Climatic Variation Shapes the Evolution of Thermal Tolerance at Multiple Scales. *Am. Nat.* 192: 347–359.

- Beavis, W. D., 1994 The power and deceit of QTL experiments: lessons from comparative QTL studies, pp. 266 in *Proceedings of the forty-ninth annual corn and sorghum industry research conference*, Chicago, IL.
- Beavis, W. D., D. Grant, M. Albertsen, and R. Fincher, 1991 Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theor. Appl. Genet.* 83: 141–145.
- Beavis, W. D., and A. H. Paterson, 1998 QTL analyses: power, precision, and accuracy. *Molecular dissection of complex traits* 1998: 145–162.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57: 289–300.
- Benjamini, Y., and D. Yekutieli, 2001 The control of the false discovery rate in multiple testing under dependency. *aos* 29: 1165–1188.
- Bennett, J. M., P. Calosi, S. Clusella-Trullas, B. Martínez, J. Sunday *et al.*, 2018 GlobTherm, a global database on thermal tolerances for aquatic and terrestrial organisms. *Sci Data* 5: 180022.
- Binenbaum, I., H. A.-T. Atamni, G. Fotakis, G. Kontogianni, T. Koutsandreas *et al.*, 2020 Container-aided integrative QTL and RNA-seq analysis of Collaborative Cross mice supports distinct sex-oriented molecular modes of response in obesity. *BMC Genomics* 21: 761.
- Blanco, G., and R. W. Mercer, 1998 Isozymes of the Na-K-ATPase: heterogeneity in structure, diversity in function. *Am. J. Physiol.* 275: F633–50.

- Blokland, G. A. M., K. L. McMahon, P. M. Thompson, N. G. Martin, G. I. de Zubicaray *et al.*, 2011a Heritability of working memory brain activation. *J. Neurosci.* 31: 10882–10890.
- Blokland, G. A. M., K. L. McMahon, P. M. Thompson, N. G. Martin, G. I. de Zubicaray *et al.*, 2011b Heritability of working memory brain activation. *J. Neurosci.* 31: 10882–10890.
- Boher, F., R. Godoy-Herrera, and F. Bozinovic, 2010 The interplay between thermal tolerance and life history is associated with the biogeography of *Drosophila* species. *Evol. Ecol. Res.* 12: 973–986.
- Bordenstein, S. R., and J. H. Werren, 1998 Effects of A and B *Wolbachia* and host genotype on interspecies cytoplasmic incompatibility in *Nasonia*. *Genetics* 148: 1833–1844.
- Bourtchuladze, R., B. Frenguelli, J. Blendy, D. Cioffi, G. Schutz *et al.*, 1994 Deficient long-term memory in mice with a targeted mutation of the cAMP-responsive element-binding protein. *Cell* 79: 59–68.
- Boyle, E. A., Y. I. Li, and J. K. Pritchard, 2017 An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* 169: 1177–1186.
- Bozinovic, F., P. Calosi, and J. I. Spicer, 2011 Physiological Correlates of Geographic Range in Animals. *Annu. Rev. Ecol. Evol. Syst.* 42: 155–179.
- Bozinovic, F., N. R. Medina, J. M. Alruiz, G. Cavieres, and P. Sabat, 2016 Thermal tolerance and survival responses to scenarios of experimental climatic change: changing thermal variability reduces the heat and cold tolerance in a fly. *J. Comp. Physiol. B* 186: 581–587.

- Brandes, C., 1988 Estimation of heritability of learning behavior in honeybees (*Apis mellifera capensis*). *Behav. Genet.* 18: 119–132.
- Brand, A. H., and N. Perrimon, 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118: 401–415.
- Broman, K. W., and S. Sen, 2009 *A Guide to QTL Mapping with R/qtl*. Springer New York: 1–399.
- Brown, J. B., N. Boley, R. Eisman, G. E. May, M. H. Stoiber *et al.*, 2014 Diversity and dynamics of the *Drosophila* transcriptome. *Nature* 512: 393–399.
- Buckley, L. B., and R. B. Huey, 2016 How Extreme Temperatures Impact Organisms and the Evolution of their Thermal Tolerance. *Integr. Comp. Biol.* 56: 98–109.
- Burdon, R. H., 1986 Heat shock and the heat shock proteins. *Biochem. J* 240: 313–324.
- Burke, C. J., W. Huetteroth, D. Oswald, E. Perisse, M. J. Krashes *et al.*, 2012 Layered reward signalling through octopamine and dopamine in *Drosophila*. *Nature* 492: 433–437.
- Burke, M. K., E. G. King, P. Shahrestani, M. R. Rose, and A. D. Long, 2014a Genome-wide association study of extreme longevity in *Drosophila melanogaster*. *Genome Biol. Evol.* 6: 1–11.
- Burke, M. K., G. Liti, and A. D. Long, 2014b Standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. *Mol. Biol. Evol.* 31: 3228–3239.

- Busto, G. U., I. Cervantes-Sandoval, and R. L. Davis, 2010 Olfactory learning in *Drosophila*. *Physiology* 25: 338–346.
- Butala, M., D. Zgur-Bertok, and S. J. W. Busby, 2009 The bacterial LexA transcriptional repressor. *Cell. Mol. Life Sci.* 66: 82–93.
- Byers, D., R. L. Davis, and J. A. Kiger Jr, 1981 Defect in cyclic AMP phosphodiesterase due to the dunce mutation of learning in *Drosophila melanogaster*. *Nature* 289: 79–81.
- Byrne, P. F., and M. D. McMullen, 1996 Defining genes for agricultural traits: QTL analysis and the candidate gene approach. *Probe*.
- Cahill, A. E., M. E. Aiello-Lammens, M. C. Fisher-Reid, X. Hua, C. J. Karanewsky *et al.*, 2013 How does climate change cause extinction? *Proc. Biol. Sci.* 280: 20121890.
- Calosi, P., D. T. Bilton, and J. I. Spicer, 2008 Thermal tolerance, acclimatory capacity and vulnerability to global climate change. *Biol. Lett.* 4: 99–102.
- Cancer Genome Atlas Network, 2015 Genomic Classification of Cutaneous Melanoma. *Cell* 161: 1681–1696.
- Carlson, C. S., M. A. Eberle, L. Kruglyak, and D. A. Nickerson, 2004 Mapping complex disease loci in whole-genome association studies. *Nature* 429: 446–452.
- Cavallaro, S., V. D'Agata, P. Manickam, F. Dufour, and D. L. Alkon, 2002 Memory-specific temporal profiles of gene expression in the hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* 99: 16279–16284.
- Chakraborty, M., J. J. Emerson, S. J. Macdonald, and A. D. Long, 2019

- Structural variants exhibit widespread allelic heterogeneity and shape variation in complex traits. *Nat. Commun.* 10: 4872.
- Chakraborty, M., N. W. VanKuren, R. Zhao, X. Zhang, S. Kalsow *et al.*, 2018 Hidden genetic variation shapes the structure of functional elements in *Drosophila*. *Nat. Genet.* 50: 20–25.
- Chakravarti, L. J., P. Buerger, R. A. Levin, and M. J. H. van Oppen, 2020 Gene regulation underpinning increased thermal tolerance in a laboratory-evolved coral photosymbiont. *Mol. Ecol.* 29: 1684–1703.
- Chakravarti, A., and T. N. Turner, 2016 Revealing rate-limiting steps in complex disease biology: The crucial importance of studying rare, extreme-phenotype families. *Bioessays* 38: 578–586.
- Chandra, S. B., G. J. Hunt, S. Cobey, and B. H. Smith, 2001 Quantitative trait loci associated with reversal learning and latent inhibition in honeybees (*Apis mellifera*). *Behav. Genet.* 31: 275–285.
- Charmantier, A., R. H. McCleery, L. R. Cole, C. Perrins, L. E. B. Kruuk *et al.*, 2008 Adaptive Phenotypic Plasticity in Response to Climate Change in a Wild Bird Population. *Science* 320: 800–803.
- Charoensuk, K., T. Sakurada, A. Tokiyama, M. Murata, T. Kosaka *et al.*, 2017 Thermotolerant genes essential for survival at a critical high temperature in thermotolerant ethanogenic *Zymomonas mobilis* TISTR 548. *Biotechnol. Biofuels* 10: 204.
- Chen, Z., 2019 The formation of the Thickveins (Tkv) gradient in *Drosophila* wing discs: A theoretical study. *J. Theor. Biol.* 474: 25–41.

- Chen, I.-C., J. K. Hill, R. Ohlemüller, D. B. Roy, and C. D. Thomas, 2011a Rapid range shifts of species associated with high levels of climate warming. *Science* 333: 1024–1026.
- Chen, A., E. F. Kramer, L. Purpura, J. L. Krill, T. Zars *et al.*, 2011b The influence of natural variation at the foraging gene on thermotolerance in adult *Drosophila* in a narrow temperature range. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 197: 1113–1118.
- Chevin, L.-M., S. Collins, and F. Lefèvre, 2013a Phenotypic plasticity and evolutionary demographic responses to climate change: taking theory out to the field. *Funct. Ecol.* 27: 967–979.
- Chevin, L.-M., R. Gallet, R. Gomulkiewicz, R. D. Holt, and S. Fellous, 2013b Phenotypic plasticity in evolutionary rescue experiments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368: 20120089.
- Chien, S., L. T. Reiter, E. Bier, and M. Gribskov, 2002 Homophila: human disease gene cognates in *Drosophila*. *Nucleic Acids Res.* 30: 149–151.
- Chow, C. Y., M. F. Wolfner, and A. G. Clark, 2013 Using natural variation in *Drosophila* to discover previously unknown endoplasmic reticulum stress genes. *Proc. Natl. Acad. Sci. U. S. A.* 110: 9013–9018.
- Churchill, G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. *Genetics* 138: 963–971.
- Clapham, D. E., 2003 TRP channels as cellular sensors. *Nature* 426: 517–524.
- Cleves, P. A., A. I. Tinoco, J. Bradford, D. Perrin, L. K. Bay *et al.*, 2020 Reduced thermal tolerance in a coral carrying CRISPR-induced mutations in the gene

- for a heat-shock transcription factor. *Proc. Natl. Acad. Sci. U. S. A.* 117: 28899–28905.
- Collier, R. J., K. Gebremedhin, A. R. Macko, and K. S. Roy, 2012 Genes Involved in the Thermal Tolerance of Livestock, pp. 379–410 in *Environmental Stress and Amelioration in Livestock Production*, edited by V. Sejian, S. M. K. Naqvi, T. Ezeji, J. Lakritz, and R. Lal. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Crawford, J. E., and R. Nielsen, 2013 Detecting adaptive trait loci in nonmodel systems: divergence or admixture mapping? *Mol. Ecol.* 22: 6131–6148.
- Croset, V., C. D. Treiber, and S. Waddell, 2018 Cellular diversity in the *Drosophila* midbrain revealed by single-cell transcriptomics. *Elife* 7.:
- Croston, R., D. Y. Kozlovsky, C. L. Branch, T. L. Parchman, E. S. Bridge *et al.*, 2016 Individual variation in spatial memory performance in wild mountain chickadees from different elevations. *Anim. Behav.* 111: 225–234.
- Cubillos, F. A., C. Brice, J. Molinet, S. Tisné, V. Abarca *et al.*, 2017 Identification of Nitrogen Consumption Genetic Variants in Yeast Through QTL Mapping and Bulk Segregant RNA-Seq Analyses. *G3* 7: 1693–1705.
- Cubillos, F. A., L. Parts, F. Salinas, A. Bergström, E. Scovacricchi *et al.*, 2013 High-resolution mapping of complex traits with a four-parent advanced intercross yeast population. *Genetics* 195: 1141–1155.
- Cuculescu, M., D. Hyde, and K. Bowler, 1998 Thermal tolerance of two species of marine crab, *Cancer pagurus* and *Carcinus maenas*. *J. Therm. Biol.* 23: 107–110.

- Cunningham, J. P., C. J. Moore, M. P. Zalucki, and S. A. West, 2004 Learning, odour preference and flower foraging in moths. *J. Exp. Biol.* 207: 87–94.
- Davies, H., 2010 A role for “omics” technologies in food safety assessment. *Food Control* 21: 1601–1610.
- Davis, M. W., D. Somerville, R. Y. Lee, S. Lockery, L. Avery *et al.*, 1995 Mutations in the *Caenorhabditis elegans* Na,K-ATPase alpha-subunit gene, *eat-6*, disrupt excitable cell function. *J. Neurosci.* 15: 8408–8418.
- Dell’Acqua, M., D. M. Gatti, G. Pea, F. Cattonaro, F. Coppens *et al.*, 2015 Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *Zea mays*. *Genome Biol.* 16: 167.
- Denny, M. W., and W. W. Dowd, 2012 Biophysics, environmental stochasticity, and the evolution of thermal safety margins in intertidal limpets. *J. Exp. Biol.* 215: 934–947.
- Dhaka, A., V. Viswanath, and A. Patapoutian, 2006 Trp ion channels and temperature sensation. *Annu. Rev. Neurosci.* 29: 135–161.
- Díaz, F., R. F. Orobio, P. Chavarriaga, and N. Toro-Perea, 2015 Differential expression patterns among heat-shock protein genes and thermal responses in the whitefly *Bemisia tabaci* (MEAM 1). *J. Therm. Biol.* 52: 199–207.
- Díaz-Ricaurte, J. C., F. C. Serrano, E. C. Guevara-Molina, C. Araujo, and M. Martins, 2021 Correction: Does behavioral thermal tolerance predict distribution pattern and habitat use in two sympatric Neotropical frogs? *PLoS One* 16: e0246851.

- Diederich, C. M., and J. A. Pechenik, 2013 Thermal tolerance of *Crepidula fornicata* (Gastropoda) life history stages from intertidal and subtidal subpopulations. *Mar. Ecol. Prog. Ser.* 486: 173–187.
- Dillon, M. E., G. Wang, P. A. Garrity, and R. B. Huey, 2009 Review: Thermal preference in *Drosophila*. *J. Therm. Biol.* 34: 109–119.
- Ding, Y., A. Berrocal, T. Morita, K. D. Longden, and D. L. Stern, 2016 Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* 536: 329–332.
- Dixon, G. B., S. W. Davies, G. V. Aglyamova, E. Meyer, L. K. Bay *et al.*, 2015 Genomic determinants of coral heat tolerance across latitudes. *Science* 348: 1460–1462.
- Dobson, S. L., C. W. Fox, and F. M. Jiggins, 2002 The effect of *Wolbachia*-induced cytoplasmic incompatibility on host population size in natural and manipulated systems. *Proc. Biol. Sci.* 269: 437–445.
- Douglas, A. E., 2018 The *Drosophila* model for microbiome research. *Lab Anim.* 47: 157–164.
- Dubnau, J., A.-S. Chiang, L. Grady, J. Barditch, S. Gossweiler *et al.*, 2003 The *staufen/pumilio* pathway is involved in *Drosophila* long-term memory. *Curr. Biol.* 13: 286–296.
- Dudai, Y., Y. N. Jan, D. Byers, W. G. Quinn, and S. Benzer, 1976 *dunce*, a mutant of *Drosophila* deficient in learning. *Proc. Natl. Acad. Sci. U. S. A.* 73: 1684–1688.
- Dudai, Y., and S. Zvi, 1985 Multiple defects in the activity of adenylate cyclase

- from the *Drosophila* memory mutant *rutabaga*. *J. Neurochem.* 45: 355–364.
- Duffy, J. B., 2002 GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. *Genesis* 34: 1–15.
- Dukas, R., 2004 Evolutionary Biology of Animal Cognition. *Annu. Rev. Ecol. Evol. Syst.*
- Dukas, R., 2008a Evolutionary biology of insect learning. *Annu. Rev. Entomol.* 53: 145–160.
- Dukas, R., 2008b Evolutionary biology of insect learning. *Annu. Rev. Entomol.* 53: 145–160.
- DuRant, S. E., W. A. Hopkins, A. W. Carter, C. M. Stachowiak, and G. R. Hepp, 2013 Incubation conditions are more important in determining early thermoregulatory ability than posthatch resource conditions in a precocial bird. *Physiol. Biochem. Zool.* 86: 410–420.
- DuRant, S. E., W. A. Hopkins, A. F. Wilson, and G. R. Hepp, 2012 Incubation temperature affects the metabolic cost of thermoregulation in a young precocial bird. *Funct. Ecol.* 26: 416–422.
- Eichler, E. E., J. Flint, G. Gibson, A. Kong, S. M. Leal *et al.*, 2010 Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* 11: 446–450.
- Falconer, D. S., 1996 Introduction to quantitative genetics: Prentice Hall,.
- Falconer, D. S., and T. F. C. Mackay, 1983 *Quantitative genetics*. Longman.
- Fangue, N. A., M. Hofmeister, and P. M. Schulte, 2006 Intraspecific variation in thermal tolerance and heat shock protein gene expression in common

- killifish, *Fundulus heteroclitus*. *J. Exp. Biol.* 209: 2859–2872.
- Feng, Y., L. Huynh, K. Takeyasu, and D. M. Fambrough, 1997 The *Drosophila* Na,K-ATPase alpha-subunit gene: gene structure, promoter function and analysis of a cold-sensitive recessive-lethal mutation. *Genes Funct.* 1: 99–117.
- Finka, A., A. F. H. Cuendet, F. J. M. Maathuis, Y. Saidi, and P. Goloubinoff, 2012 Plasma membrane cyclic nucleotide gated calcium channels control land plant thermal sensing and acquired thermotolerance. *Plant Cell* 24: 3333–3348.
- Fowler, M. A., and C. Montell, 2013 *Drosophila* TRP channels and animal behavior. *Life Sci.* 92: 394–403.
- Fox, C. W., and J. B. Wolf, 2006 *Evolutionary Genetics: Concepts and Case Studies*. Oxford University Press.
- Frank, G., E. Pressman, R. Ophir, L. Althan, R. Shaked *et al.*, 2009 Transcriptional profiling of maturing tomato (*Solanum lycopersicum* L.) microspores reveals the involvement of heat shock proteins, ROS scavengers, hormones, and sugars in the heat stress response. *J. Exp. Bot.* 60: 3891–3908.
- Fraser, H. B., and X. Xie, 2009 Common polymorphic transcript variation in human disease. *Genome Res.* 19: 567–575.
- Frazer, K. A., S. S. Murray, N. J. Schork, and E. J. Topol, 2009 Human genetic variation and its contribution to complex traits. *Nat. Rev. Genet.* 10: 241–251.

- Frazier, M. R., R. B. Huey, and D. Berrigan, 2006 Thermodynamics constrains the evolution of insect population growth rates: “warmer is better.” *Am. Nat.* 168: 512–520.
- Frydenberg, J., A. A. Hoffmann, and V. Loeschcke, 2003 DNA sequence variation and latitudinal associations in hsp23, hsp26 and hsp27 from natural populations of *Drosophila melanogaster*. *Mol. Ecol.* 12: 2025–2032.
- Fry, A. J., M. R. Palmer, and D. M. Rand, 2004 Variable fitness effects of *Wolbachia* infection in *Drosophila melanogaster*. *Heredity* 93: 379–389.
- Galton, F., 1875 The history of twins, as a criterion of the relative powers of nature and nurture. *Fraser’s Magazine* 12: 566–576.
- Gao, J., W. Zhang, W. Dang, Y. Mou, Y. Gao *et al.*, 2014 Heat shock protein expression enhances heat tolerance of reptile embryos. *Proc. Biol. Sci.* 281: 20141135.
- Garrity, P. A., M. B. Goodman, A. D. Samuel, and P. Sengupta, 2010 Running hot and cold: behavioral strategies, neural circuits, and the molecular machinery for thermotaxis in *C. elegans* and *Drosophila*. *Genes Dev.* 24: 2365–2382.
- Gaspar, P., S. Arif, L. Sumner-Rooney, M. Kittelmann, A. J. Bodey *et al.*, 2020 Characterization of the genetic architecture underlying eye size variation within *Drosophila melanogaster* and *Drosophila simulans*. *G3: Genes, Genomes, Genetics* 10: 1005–1018.
- Gaudet, P., M. S. Livstone, S. E. Lewis, and P. D. Thomas, 2011 Phylogenetic-based propagation of functional annotations within the Gene Ontology

- consortium. *Brief. Bioinform.* 12: 449–462.
- Gibson, G., and B. Weir, 2005 The quantitative genetics of transcription. *Trends Genet.* 21: 616–623.
- Gilad, Y., S. A. Rifkin, and J. K. Pritchard, 2008 Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet.* 24: 408–415.
- Gilchrist, G. W., and R. B. Huey, 2004 Plastic and Genetic Variation in Wing Loading as a Function of Temperature Within and Among Parallel Clines in *Drosophila subobscura*¹. *Integr. Comp. Biol.* 44: 461–470.
- Giraud, H., C. Lehermeier, E. Bauer, M. Falque, V. Segura *et al.*, 2014 Linkage disequilibrium with linkage analysis of multiline crosses reveals different multiallelic QTL for hybrid performance in the flint and dent heterotic groups of maize. *Genetics* 198: 1717–1734.
- Godenschwege, T. A., D. Reisch, S. Diegelmann, K. Eberle, N. Funk *et al.*, 2004 Flies lacking all synapsins are unexpectedly healthy but are impaired in complex behaviour. *Eur. J. Neurosci.* 20: 611–622.
- Goldstein, D. B., 2009 Common Genetic Variation and Human Traits. *N. Engl. J. Med.* 360: 1696–1698.
- Gomez, M., E. De Castro, E. Guarin, H. Sasakura, A. Kuhara *et al.*, 2001 Ca²⁺ signaling via the neuronal calcium sensor-1 regulates associative learning and memory in *C. elegans*. *Neuron* 30: 241–248.
- Gould, J. L., 1987 Honey bees store learned flower-landing behaviour according to time of day. *Anim. Behav.* 35: 1579–1581.
- Goulson, D., and J. S. Cory, 1993 Flower constancy and learning in foraging

- preferences of the green-veined white butterfly *Pieris napi*. *Ecol. Entomol.* 18: 315–320.
- Grainger, T. N., S. M. Rudman, P. Schmidt, and J. M. Levine, 2021 Competitive history shapes rapid evolution in a seasonal climate. *Proc. Natl. Acad. Sci. U. S. A.* 118.:
- Gratz, S. J., C. D. Rubinstein, M. M. Harrison, J. Wildonger, and K. M. O'Connor-Giles, 2015 CRISPR-Cas9 genome editing in *Drosophila*. *Curr. Protoc. Mol. Biol.* 111: 31.2.1–31.2.20.
- Grinder, R. M., R. D. Bassar, and S. K. Auer, 2020 Upper thermal limits are repeatable in Trinidadian guppies. *J. Therm. Biol.* 90: 102597.
- Gunderson, A. R., and J. H. Stillman, 2015 Plasticity in thermal tolerance has limited potential to buffer ectotherms from global warming. *Proc. Biol. Sci.* 282: 20150401.
- Gurley, W. B., 2000 HSP101: a key component for the acquisition of thermotolerance in plants. *Plant Cell* 12: 457–460.
- Haley, C. S., and S. A. Knott, 1992 A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69: 315–324.
- Hallauer, A. R., M. J. Carena, and J. B. Miranda Filho, 2010 *Quantitative Genetics in Maize Breeding*. Springer Science & Business Media.
- Hamada, H., M. G. Petrino, and T. Kakunaga, 1982 A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. U. S. A.* 79: 6465–6469.

- Hamada, F. N., M. Rosenzweig, K. Kang, S. R. Pulver, A. Ghezzi *et al.*, 2008 An internal thermal sensor controlling temperature preference in *Drosophila*. *Nature* 454: 217–220.
- Hammer, M., and R. Menzel, 1995 Learning and memory in the honeybee. *J. Neurosci.* 15: 1617–1630.
- Hammond, L. M., and G. E. Hofmann, 2010 Thermal tolerance of *Strongylocentrotus purpuratus* early life history stages: mortality, stress-induced gene expression and biogeographic patterns. *Mar. Biol.* 157: 2677–2687.
- Harley, C. D. G., and R. T. Paine, 2009 Contingencies and compounded rare perturbations dictate sudden distributional shifts during periods of gradual climate change. *Proc. Natl. Acad. Sci. U. S. A.* 106: 11172–11176.
- Healy, S. D., and T. A. Hurly, 2004 Spatial learning and memory in birds. *Brain Behav. Evol.* 63: 211–220.
- Heck, A., M. Fastenrath, S. Ackermann, B. Auschra, H. Bickel *et al.*, 2014 Converging genetic and functional brain imaging evidence links neuronal excitability to working memory, psychiatric disease, and brain activity. *Neuron* 81: 1203–1213.
- Hertz, P. E., R. B. Huey, and R. D. Stevenson, 1993 Evaluating temperature regulation by field-active ectotherms: the fallacy of the inappropriate question. *Am. Nat.* 142: 796–818.
- Hige, T., Y. Aso, M. N. Modi, G. M. Rubin, and G. C. Turner, 2015 Heterosynaptic Plasticity Underlies Aversive Olfactory Learning in

- Drosophila*. *Neuron* 88: 985–998.
- Highfill, C. A., J. H. Tran, S. K. T. Nguyen, T. R. Moldenhauer, X. Wang *et al.*, 2017 Naturally Segregating Variation at *Ugt86Dd* Contributes to Nicotine Resistance in *Drosophila melanogaster*. *Genetics* 207: 311–325.
- Hindorff, L. A., P. Sethupathy, H. A. Junkins, E. M. Ramos, J. P. Mehta *et al.*, 2009 Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U. S. A.* 106: 9362–9367.
- Hirschhorn, J. N., and M. J. Daly, 2005 Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* 6: 95–108.
- Hoffmann, A. A., 2010 Physiological climatic limits in *Drosophila*: patterns and implications. *J. Exp. Biol.* 213: 870–880.
- Hoffmann, A. A., A. Anderson, and R. Hallas, 2002 Opposing clines for high and low temperature resistance in *Drosophila melanogaster*. *Ecol. Lett.* 5: 614–618.
- Hoffmann, A. A., S. L. Chown, and S. Clusella-Trullas, 2013 Upper thermal limits in terrestrial ectotherms: how constrained are they? *Funct. Ecol.* 27: 934–949.
- Hoffmann, A. A., and C. M. Sgrò, 2011 Climate change and evolutionary adaptation. *Nature* 470: 479–485.
- Hoffmann, A. A., J. G. Sørensen, and V. Loeschcke, 2003 Adaptation of *Drosophila* to temperature extremes: bringing together quantitative and molecular approaches. *J. Therm. Biol.* 28: 175–216.

- Hoffmann, A. A., M. Turelli, and G. M. Simmons, 1986 UNIDIRECTIONAL INCOMPATIBILITY BETWEEN POPULATIONS OF DROSOPHILA SIMULANS. *Evolution* 40: 692–701.
- Hong, S. W., and E. Vierling, 2000 Mutants of *Arabidopsis thaliana* defective in the acquisition of tolerance to high temperature stress. *Proc. Natl. Acad. Sci. U. S. A.* 97: 4392–4397.
- Huang, L.-H., B. Chen, and L. Kang, 2007 Impact of mild temperature hardening on thermotolerance, fecundity, and Hsp gene expression in *Liriomyza huidobrensis*. *J. Insect Physiol.* 53: 1199–1205.
- Huey, R. B., and D. Berrigan, 2001 Temperature, demography, and ectotherm fitness. *Am. Nat.* 158: 204–210.
- Huey, R. B., C. A. Deutsch, J. J. Tewksbury, L. J. Vitt, P. E. Hertz *et al.*, 2009 Why tropical forest lizards are vulnerable to climate warming. *Proc. Biol. Sci.* 276: 1939–1948.
- Huey, R. B., M. R. Kearney, A. Krockenberger, J. A. M. Holtum, M. Jess *et al.*, 2012 Predicting organismal vulnerability to climate warming: roles of behaviour, physiology and adaptation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367: 1665–1679.
- Hu, H., M. Li, J.-P. Labrador, J. McEwen, E. C. Lai *et al.*, 2005 Cross GTPase-activating protein (CrossGAP)/Vilse links the Roundabout receptor to Rac to regulate midline repulsion. *Proc. Natl. Acad. Sci. U. S. A.* 102: 4613–4618.
- Human Genome Structural Variation Working Group, E. E. Eichler, D. A. Nickerson, D. Altshuler, A. M. Bowcock *et al.*, 2007 Completing the map of

- human genetic variation. *Nature* 447: 161–165.
- Hu, K.-M., D.-Y. Qiu, X.-L. Shen, X.-H. Li, and S.-P. Wang, 2008 Isolation and manipulation of quantitative trait loci for disease resistance in rice using a candidate gene approach. *Mol. Plant* 1: 786–793.
- Hutchison, V. H., and J. D. Maness, 1979 The role of behavior in temperature acclimation and tolerance in ectotherms. *Am. Zool.* 19: 367–384.
- Ito, I., R. C.-Y. Ong, B. Raman, and M. Stopfer, 2008 Sparse odor representation and olfactory learning. *Nat. Neurosci.* 11: 1177–1184.
- Iwama, G. K., P. T. Thomas, R. B. Forsyth, and M. M. Vijayan, 1998 Heat shock protein expression in fish. *Rev. Fish Biol. Fish.* 8: 35–56.
- Jennings, B. H., 2011 *Drosophila* – a versatile model in biology & medicine. *Mater. Today* 14: 190–195.
- Jensen, L. T., F. E. Cockerell, T. N. Kristensen, L. Rako, V. Loeschcke *et al.*, 2010 Adult heat tolerance variation in *Drosophila melanogaster* is not related to Hsp70 expression. *J. Exp. Zool. A Ecol. Genet. Physiol.* 313: 35–44.
- Jha, A. R., C. M. Miles, N. R. Lippert, C. D. Brown, K. P. White *et al.*, 2015 Whole-Genome Resequencing of Experimental Populations Reveals Polygenic Basis of Egg-Size Variation in *Drosophila melanogaster*. *Mol. Biol. Evol.* 32: 2616–2632.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna *et al.*, 2012 A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337: 816–821.
- Jordán-Álvarez, S., E. Santana, S. Casas-Tintó, Á. Acebes, and A. Ferrús, 2017

- The equilibrium between antagonistic signaling pathways determines the number of synapses in *Drosophila*. PLoS One 12: e0184238.
- Jørgensen, L. B., R. M. Robertson, and J. Overgaard, 2020 Neural dysfunction correlates with heat coma and CTmax in *Drosophila* but does not set the boundaries for heat stress survival. J. Exp. Biol. 223.:
- Jostins, L., S. Ripke, R. K. Weersma, R. H. Duerr, D. P. McGovern *et al.*, 2012 Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 491: 119–124.
- Kahsai, L., and T. Zars, 2011 Learning and Memory in *Drosophila*: Behavior, Genetics, and Neural Systems, pp. 139–167 in *International Review of Neurobiology*, edited by N. Atkinson. Academic Press.
- Kang, M. S., 2020 *Quantitative Genetics, Genomics and Plant Breeding, 2nd Edition*. CABI.
- Kan, L., S. Ott, B. Joseph, E. S. Park, C. Dai *et al.*, 2020 A neural m6A/YTHDF pathway is required for learning and memory in *Drosophila*. bioRxiv.
- Kasuya, J., H. Ishimoto, and T. Kitamoto, 2009 Neuronal mechanisms of learning and memory revealed by spatial and temporal suppression of neurotransmission using shibirets1, a temperature-sensitive dynamin mutant gene in *Drosophila melanogaster*. Front. Mol. Neurosci. 2: 11.
- Kawecki, T. J., 2010 Evolutionary ecology of learning: insights from fruit flies. Popul. Ecol. 52: 15–25.
- Kearsey, M. J., 1998 The principles of QTL analysis (a minimal mathematics approach). J. Exp. Bot. 49: 1619–1623.

- Keele, G. R., W. L. Crouse, S. N. P. Kelada, and W. Valdar, 2018 Determinants of QTL mapping power in the realized Collaborative Cross. *bioRxiv* 459966.
- Kellett, M., A. A. Hoffmann, and S. W. Mckechnie, 2005 Hardening capacity in the *Drosophila melanogaster* species group is constrained by basal thermotolerance. *Funct. Ecol.* 19: 853–858.
- Kesner, R. P., R. O. Hopkins, and A. A. Chiba, 1992 Learning and memory in humans, with an emphasis on the role of the hippocampus. *Neuropsychology of memory.*, 2nd ed. 2: 106–121.
- Khaliq, I., C. Hof, R. Prinzinger, K. Böhning-Gaese, and M. Pfenninger, 2014 Global variation in thermal tolerances and vulnerability of endotherms to climate change. *Proc. Biol. Sci.* 281: 20141097.
- Kim, D., B. Langmead, and S. L. Salzberg, 2015 HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12: 357.
- King, E. G., and A. D. Long, 2017 The Beavis Effect in Next-Generation Mapping Panels in *Drosophila melanogaster*. *G3* 7: 1643–1652.
- King, E. G., S. J. Macdonald, and A. D. Long, 2012a Properties and power of the *Drosophila* Synthetic Population Resource for the routine dissection of complex traits. *Genetics* 191: 935–949.
- King, E. G., C. M. Merkes, C. L. McNeil, S. R. Hoofer, S. Sen *et al.*, 2012b Genetic dissection of a model complex trait using the *Drosophila* Synthetic Population Resource. *Genome Res.* 22: 1558–1566.
- King, E. G., B. J. Sanderson, C. L. McNeil, A. D. Long, and S. J. Macdonald, 2014 Genetic dissection of the *Drosophila melanogaster* female head

- transcriptome reveals widespread allelic heterogeneity. *PLoS Genet.* 10: e1004322.
- Kingsolver, J. G., and L. B. Buckley, 2015 Climate variability slows evolutionary responses of *Colias* butterflies to recent climate change. *Proc. Biol. Sci.* 282.:
- Kitamoto, T., 2001 Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive shibire allele in defined neurons. *J. Neurobiol.* 47: 81–92.
- Kjærsgaard, A., D. Demontis, T. N. Kristensen, N. Le, S. Faurby *et al.*, 2010 Locomotor activity of *Drosophila melanogaster* in high temperature environments: plastic and evolutionary responses. *Clim. Res.* 43: 127–134.
- Knoll, A. T., L. R. Halladay, A. J. Holmes, and P. Levitt, 2016 Quantitative Trait Loci and a Novel Genetic Candidate for Fear Learning. *J. Neurosci.* 36: 6258–6268.
- Kobayashi, M., H. Katoh, T. Takayanagi, and S. Suzuki, 2010 Characterization of thermotolerance-related genes in grapevine (*Vitis vinifera*). *J. Plant Physiol.* 167: 812–819.
- Ko, Y.-W., C.-P. Liao, R. W. Clark, J.-Y. Hsu, H.-Y. Tseng *et al.*, 2020 Aposematic coloration of prey enhances memory retention in an agamid lizard. *Anim. Behav.* 161: 1–13.
- Königer, A., and S. Grath, 2018 Transcriptome Analysis Reveals Candidate Genes for Cold Tolerance in *Drosophila ananassae*. *Genes* 9.:
- de Koning, D.-J., and L. M. McIntyre, 2017a Back to the Future: Multiparent

- Populations Provide the Key to Unlocking the Genetic Basis of Complex Traits. *G3: Genes, Genomes, Genetics* 7: 1617–1618.
- de Koning, D.-J., and L. M. McIntyre, 2017b Back to the Future: Multiparent Populations Provide the Key to Unlocking the Genetic Basis of Complex Traits. *Genetics* 206: 527–529.
- Koon, A. C., and V. Budnik, 2012 Inhibitory control of synaptic and behavioral plasticity by octopaminergic signaling. *J. Neurosci.* 32: 6312–6322.
- Kramer, J. M., K. Kochinke, M. A. W. Oortveld, H. Marks, D. Kramer *et al.*, 2011 Epigenetic regulation of learning and memory by *Drosophila* EHMT/G9a. *PLoS Biol.* 9: e1000569.
- Krashes, M. J., A. C. Keene, B. Leung, J. D. Armstrong, and S. Waddell, 2007 Sequential use of mushroom body neuron subsets during *drosophila* odor memory processing. *Neuron* 53: 103–115.
- Krashes, M. J., and S. Waddell, 2008 Rapid consolidation to a radish and protein synthesis-dependent long-term memory after single-session appetitive olfactory conditioning in *Drosophila*. *J. Neurosci.* 28: 3103–3113.
- Krishnan, M., H. T. Nguyen, and J. J. Burke, 1989 Heat shock protein synthesis and thermal tolerance in wheat. *Plant Physiol.* 90: 140–145.
- Kruuk, L. E. B., J. Slate, and A. J. Wilson, 2008 New Answers for Old Questions: The Evolutionary Quantitative Genetics of Wild Animal Populations. *Annu. Rev. Ecol. Evol. Syst.*
- Kuhnlein, R. P., 2012 Lipid droplet-based storage fat metabolism in *Drosophila*. *J. Lipid Res.*

- Kurusu, M., A. Cording, M. Taniguchi, K. Menon, E. Suzuki *et al.*, 2008 A screen of cell-surface molecules identifies leucine-rich repeat proteins as key mediators of synaptic target selection. *Neuron* 59: 972–985.
- Kuwabara, P. E., and A. Coulson, 2000 RNAi--prospects for a general technique for determining gene function. *Parasitol. Today* 16: 347–349.
- LaDage, L. D., T. C. Roth, A. M. Cerjanic, B. Sinervo, and V. V. Pravosudov, 2012 Spatial memory: are lizards really deficient? *Biol. Lett.* 8: 939–941.
- Ladner, J. T., D. J. Barshis, and S. R. Palumbi, 2012 Protein evolution in two co-occurring types of Symbiodinium: an exploration into the genetic basis of thermal tolerance in Symbiodinium clade D. *BMC Evol. Biol.* 12: 217.
- LaFerriere, H., D. J. Guarnieri, D. Sitaraman, S. Diegelmann, U. Heberlein *et al.*, 2008 Genetic dissociation of ethanol sensitivity and memory formation in *Drosophila melanogaster*. *Genetics* 178: 1895–1902.
- LaFerriere, H., K. Speichinger, A. Stromhaug, and T. Zars, 2011 The radish gene reveals a memory component with variable temporal properties. *PLoS One* 6: e24557.
- Lancaster, L. T., and A. M. Humphreys, 2020 Global variation in the thermal tolerances of plants. *Proc. Natl. Acad. Sci. U. S. A.* 117: 13580–13587.
- Lander, E. S., and D. Botstein, 1989 Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185–199.
- Langham, G. M., 2004 Specialized avian predators repeatedly attack novel color morphs of *Heliconius* butterflies. *Evolution* 58: 2783–2787.
- Lansing, E., J. Justesen, and V. Loeschcke V., 2000 Variation in the expression

- of Hsp70, the major heat-shock protein, and thermotolerance in larval and adult selection lines of *Drosophila melanogaster*. *J. Therm. Biol.* 25: 443–450.
- Lau, H. L., T. A. Timbers, R. Mahmoud, and C. H. Rankin, 2013 Genetic dissection of memory for associative and non-associative learning in *Caenorhabditis elegans*: Short- and long-term memory for context. *Genes Brain Behav.* 12: 210–223.
- Lecheta, M. C., D. N. Awde, T. S. O’Leary, L. N. Unfried, N. A. Jacobs *et al.*, 2020 Integrating GWAS and Transcriptomics to Identify the Molecular Underpinnings of Thermal Stress Responses in *Drosophila melanogaster*. *Front. Genet.* 11: 658.
- Leek, J. T., W. E. Johnson, H. S. Parker, E. J. Fertig, A. E. Jaffe *et al.*, 2019 *sva: Surrogate Variable Analysis*.
- Leek, J. T., W. E. Johnson, H. S. Parker, A. E. Jaffe, and J. D. Storey, 2012 The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28: 882–883.
- Leinonen, R., H. Sugawara, M. Shumway, and International Nucleotide Sequence Database Collaboration, 2011 The sequence read archive. *Nucleic Acids Res.* 39: D19–21.
- Lemos, B., L. O. Araripe, P. Fontanillas, and D. L. Hartl, 2008 Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 105: 14471–14476.
- Letcher, T. M., 2015 *Climate Change: Observed Impacts on Planet Earth*.

Elsevier.

- Levin, L. R., P. L. Han, P. M. Hwang, P. G. Feinstein, R. L. Davis *et al.*, 1992 The *Drosophila* learning and memory gene *rutabaga* encodes a Ca^{2+} /Calmodulin-responsive adenylyl cyclase. *Cell* 68: 479–489.
- Li, Y. I., B. van de Geijn, A. Raj, D. A. Knowles, A. A. Petti *et al.*, 2016 RNA splicing is a primary link between genetic variation and disease. *Science* 352: 600–604.
- Li, Q.-L., Z.-H. Ju, J.-M. Huang, J.-B. Li, R.-L. Li *et al.*, 2011 Two novel SNPs in HSF1 gene are associated with thermal tolerance traits in Chinese Holstein cattle. *DNA Cell Biol.* 30: 247–254.
- Little, J. W., and D. W. Mount, 1982 The SOS regulatory system of *Escherichia coli*. *Cell* 29: 11–22.
- Livingstone, M. S., P. P. Sziber, and W. G. Quinn, 1984 Loss of calcium/calmodulin responsiveness in adenylate cyclase of *rutabaga*, a *Drosophila* learning mutant. *Cell* 37: 205–215.
- Locke, A. E., B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers *et al.*, 2015 Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518: 197–206.
- Lockwood, B. L., C. R. Julick, and K. L. Montooth, 2017 Maternal loading of a small heat shock protein increases embryo thermal tolerance in *Drosophila melanogaster*. *J. Exp. Biol.* 220: 4492–4501.
- Lofdahl, K. L., M. Holliday, and J. Hirsch, 1992 Selection for conditionability in *Drosophila melanogaster*. *J. Comp. Psychol.* 106: 172–183.

- Long, A., G. Liti, A. Luptak, and O. Tenaillon, 2015 Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nat. Rev. Genet.* 16: 567–582.
- Long, A. D., S. J. Macdonald, and E. G. King, 2014 Dissecting complex traits using the *Drosophila* Synthetic Population Resource. *Trends Genet.* 30: 488–495.
- Love, M. I., W. Huber, and S. Anders, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15: 550.
- Lundström, A., M. Gallio, C. Englund, P. Steneberg, J. Hemphälä *et al.*, 2004 Vilse, a conserved Rac/Cdc42 GAP mediating Robo repulsion in tracheal cells and axons. *Genes Dev.* 18: 2161–2171.
- Lynch, M., B. Walsh, and Others, 1998 *Genetics and analysis of quantitative traits*. Sinauer Sunderland, MA.
- Mackay, T. F. C., 2001 The genetic architecture of quantitative traits. *Annu. Rev. Genet.* 35: 303–339.
- Mackay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2012 The *Drosophila melanogaster* Genetic Reference Panel. *Nature* 482: 173–178.
- MacLean, H. J., J. Overgaard, T. N. Kristensen, C. Lyster, L. Hessner *et al.*, 2019a Temperature preference across life stages and acclimation temperatures investigated in four species of *Drosophila*. *J. Therm. Biol.* 86: 102428.
- MacLean, H. J., J. G. Sørensen, T. N. Kristensen, V. Loeschcke, K. Beedholm *et*

- al.*, 2019b Evolution and plasticity of thermal performance: an analysis of variation in thermal tolerance and fitness in 22 *Drosophila* species. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 374: 20180548.
- Magyar, J. P., U. Bartsch, Z. Q. Wang, N. Howells, A. Aguzzi *et al.*, 1994 Degeneration of neural cells in the central nervous system of mice deficient in the gene for the adhesion molecule on Glia, the beta 2 subunit of murine Na,K-ATPase. *J. Cell Biol.* 127: 835–845.
- Maher, B., 2008 Personal genomes: The case of the missing heritability. *Nature* 456: 18–21.
- Manichaikul, A., J. Dupuis, S. Sen, and K. W. Broman, 2006 Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* 174: 481–489.
- Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff *et al.*, 2009 Finding the missing heritability of complex diseases. *Nature* 461: 747–753.
- Marella, S., W. Fischler, P. Kong, S. Asgarian, E. Rueckert *et al.*, 2006 Imaging taste responses in the fly brain reveals a functional map of taste category and behavior. *Neuron* 49: 285–295.
- Marriage, T. N., E. G. King, A. D. Long, and S. J. Macdonald, 2014 Fine-mapping nicotine resistance loci in *Drosophila* using a multiparent advanced generation inter-cross population. *Genetics* 198: 45–57.
- Martin, M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17: 10–12.
- McCain, C. M., and R. K. Colwell, 2011 Assessing the threat to montane

- biodiversity from discordant shifts in temperature and precipitation in a changing climate. *Ecol. Lett.* 14: 1236–1245.
- McFaline-Figueroa, J. L., C. Trapnell, and J. T. Cuperus, 2020 The promise of single-cell genomics in plants. *Curr. Opin. Plant Biol.* 54: 114–121.
- McGuire, T. R., and J. Hirsch, 1977a Behavior-genetic analysis of *Phormia regina*: conditioning, reliable individual differences, and selection. *Proc. Natl. Acad. Sci. U. S. A.* 74: 5193–5197.
- McGuire, T. R., and J. Hirsch, 1977b Behavior-genetic analysis of *Phormia regina*: conditioning, reliable individual differences, and selection. *Proc. Natl. Acad. Sci. U. S. A.* 74: 5193–5197.
- McGuire, T. R., and J. Hirsch, 1977c General intelligence (g) and heritability (H^2), pp. 25–72 in *The structuring of experience*, Springer.
- McKechnie, A. E., and B. O. Wolf, 2010 Climate change increases the likelihood of catastrophic avian mortality events during extreme heat waves. *Biol. Lett.* 6: 253–256.
- McQuillan, M. A., T. C. Roth 2nd, A. V. Huynh, and A. M. Rice, 2018 Hybrid chickadees are deficient in learning and memory. *Evolution* 72: 1155–1164.
- Meltzer, H., E. Marom, I. Alyagor, O. Mayseless, V. Berkun *et al.*, 2019 Tissue-specific (ts)CRISPR as an efficient strategy for in vivo screening in *Drosophila*. *Nat. Commun.* 10: 2113.
- Mery, F., 2013a Natural variation in learning and memory. *Curr. Opin. Neurobiol.* 23: 52–56.
- Mery, F., 2013b Natural variation in learning and memory. *Curr. Opin. Neurobiol.*

23: 52–56.

Mery, F., A. T. Belay, A. K.-C. So, M. B. Sokolowski, and T. J. Kawecki, 2007a

Natural polymorphism affecting learning and memory in *Drosophila*. *Proc.*

Natl. Acad. Sci. U. S. A. 104: 13051–13055.

Mery, F., A. T. Belay, A. K.-C. So, M. B. Sokolowski, and T. J. Kawecki, 2007b

Natural polymorphism affecting learning and memory in *Drosophila*. *Proc.*

Natl. Acad. Sci. U. S. A. 104: 13051–13055.

Mery, F., and T. J. Kawecki, 2003 A fitness cost of learning ability in *Drosophila*

melanogaster. *Proc. Biol. Sci.* 270: 2465–2469.

Mery, F., and T. J. Kawecki, 2002a Experimental evolution of learning ability in

fruit flies. *Proc. Natl. Acad. Sci. U. S. A.* 99: 14274–14279.

Mery, F., and T. J. Kawecki, 2002b Experimental evolution of learning ability in

fruit flies. *Proc. Natl. Acad. Sci. U. S. A.* 99: 14274–14279.

Mery, F., J. Pont, T. Preat, and T. J. Kawecki, 2007c Experimental evolution of

olfactory memory in *Drosophila melanogaster*. *Physiol. Biochem. Zool.* 80:

399–405.

Mi, H., X. Huang, A. Muruganujan, H. Tang, C. Mills *et al.*, 2017 PANTHER

version 11: expanded annotation data from Gene Ontology and Reactome

pathways, and data analysis tool enhancements. *Nucleic Acids Res.* 45:

D183–D189.

Miles, C. M., S. E. Lott, C. L. L. Hendriks, M. Z. Ludwig, Manu *et al.*, 2011

Artificial selection on egg size perturbs early pattern formation in *Drosophila*

melanogaster. *Evolution* 65: 33–42.

- Mobasheri, A., J. Avila, I. Czar-Castellano, M. D. Brownleader, M. Trevan *et al.*, 2000 Na⁺, K⁺-ATPase isozyme diversity; comparative biochemistry and physiological implications of novel functional interactions. *Biosci. Rep.* 20: 51–91.
- Molthagen, M., M. Schachner, and U. Bartsch, 1996 Apoptotic cell death of photoreceptor cells in mice deficient for the adhesion molecule on glia (AMOG, the beta 2- subunit of the Na, K-ATPase). *J. Neurocytol.* 25: 243–255.
- Morand-Ferron, J., and J. L. Quinn, 2015a The evolution of cognition in natural populations. *Trends Cogn. Sci.* 19: 235–237.
- Morand-Ferron, J., and J. L. Quinn, 2015b The evolution of cognition in natural populations. *Trends Cogn. Sci.* 19: 235–237.
- Morgan, T. J., and T. F. C. Mackay, 2006 Quantitative trait loci for thermotolerance phenotypes in *Drosophila melanogaster*. *Heredity* 96: 232–242.
- Mori, I., 1999 Genetics of chemotaxis and thermotaxis in the nematode *Caenorhabditis elegans*. *Annu. Rev. Genet.* 33: 399–422.
- Morin, P. A., G. Luikart, R. K. Wayne, and the SNP workshop group, 2004 SNPs in ecology, evolution and conservation. *Trends Ecol. Evol.* 19: 208–216.
- Mousseau, T. A., and D. A. Roff, 1987 Natural selection and the heritability of fitness components. *Heredity* 59 (Pt 2): 181–197.
- Najarro, M. A., J. L. Hackett, B. R. Smith, C. A. Highfill, E. G. King *et al.*, 2015 Identifying Loci Contributing to Natural Variation in Xenobiotic Resistance in

- Drosophila*. PLoS Genet. 11: e1005663.
- Neargarder, G., E. P. Dahlhoff, and N. E. Rank, 2003 Variation in thermal tolerance is linked to phosphoglucose isomerase genotype in a montane leaf beetle. *Funct. Ecol.* 17: 213–221.
- Neely, G. G., A. Hess, M. Costigan, A. C. Keene, S. Goulas *et al.*, 2010 A genome-wide *Drosophila* screen for heat nociception identifies $\alpha 2\delta 3$ as an evolutionarily conserved pain gene. *Cell* 143: 628–638.
- Ng'oma, E., P. A. Williams-Simon, A. Rahman, and E. G. King, 2020 Diverse biological processes coordinate the transcriptional response to nutritional changes in a *Drosophila melanogaster* multiparent population. *BMC Genomics* 21: 84.
- Nguyen, C., M. H. Bahar, G. Baker, and N. R. Andrew, 2014 Thermal tolerance limits of diamondback moth in ramping and plunging assays. *PLoS One* 9: e87535.
- Nishi, M., and K. Nanjo, 2011 Insulin gene mutations and diabetes. *J. Diabetes Investig.* 2: 92–100.
- Nord, A., and J.-Å. Nilsson, 2011 Incubation temperature affects growth and energy metabolism in blue tit nestlings. *Am. Nat.* 178: 639–651.
- Norry, F. M., J. Dahlgaard, and V. Loeschcke, 2004 Quantitative trait loci affecting knockdown resistance to high temperature in *Drosophila melanogaster*. *Mol. Ecol.* 13: 3585–3594.
- Norry, F. M., F. H. Gomez, and V. Loeschcke, 2007a Knockdown resistance to heat stress and slow recovery from chill coma are genetically associated in a

- quantitative trait locus region of chromosome 2 in *Drosophila melanogaster*.
Mol. Ecol. 16: 3274–3284.
- Norry, F. M., P. Sambucetti, A. C. Scannapieco, F. H. Gomez, and V. Loeschcke,
2007b X-linked QTL for knockdown resistance to high temperature in
Drosophila melanogaster. Insect Mol. Biol. 16: 509–513.
- Norry, F. M., A. C. Scannapieco, P. Sambucetti, C. I. Bertoli, and V. Loeschcke,
2008 QTL for the thermotolerance effect of heat hardening, knockdown
resistance to heat and chill-coma recovery in an intercontinental set of
recombinant inbred lines of *Drosophila melanogaster*. Mol. Ecol. 17: 4570–
4581.
- Oleksiak, M. F., G. A. Churchill, and D. L. Crawford, 2002 Variation in gene
expression within and among natural populations. Nat. Genet. 32: 261–266.
- Ørsted, M., A. A. Hoffmann, P. D. Rohde, P. Sørensen, and T. N. Kristensen,
2019 Strong impact of thermal environment on the quantitative genetic basis
of a key stress tolerance trait. Heredity 122: 315–325.
- Ostrowski, D., L. Kahsai, E. F. Kramer, P. Knutson, and T. Zars, 2015 Place
memory retention in *Drosophila*. Neurobiol. Learn. Mem. 123: 217–224.
- Ostrowski, D., and T. Zars, 2014 Place memory, pp. 125–134 in *Handbook of
behavior genetics of Drosophila melanogaster: Behavioral phenotypes and
models of neurobehavioral disorders*, edited by J. Dubnau.
books.google.com.
- Overgaard, J., T. N. Kristensen, K. A. Mitchell, and A. A. Hoffmann, 2011
Thermal tolerance in widespread and tropical *Drosophila* species: does

- phenotypic plasticity increase with latitude? *Am. Nat.* 178 Suppl 1: S80–96.
- Overgaard, J., and J. G. Sørensen, 2008 Rapid thermal adaptation during field temperature variations in *Drosophila melanogaster*. *Cryobiology* 56: 159–162.
- Overgaard, J., A. Tomcala, J. G. Sørensen, M. Holmstrup, P. H. Krogh *et al.*, 2008 Effects of acclimation temperature on thermal tolerance and membrane phospholipid composition in the fruit fly *Drosophila melanogaster*. *J. Insect Physiol.* 54: 619–629.
- Owaid, D., S. Lin, and S. Waddell, 2015 Light, heat, action: neural control of fruit fly behaviour. *Transactions of the Royal*
- Palladino, M. J., J. E. Bower, R. Kreber, and B. Ganetzky, 2003 Neural Dysfunction and Neurodegeneration in *Drosophila* Na⁺/K⁺ ATPase Alpha Subunit Mutants. *J. Neurosci.* 23: 1276–1286.
- Palladino, M. J., T. J. Hadley, and B. Ganetzky, 2002 Temperature-sensitive paralytic mutants are enriched for those causing neurodegeneration in *Drosophila*. *Genetics* 161: 1197–1208.
- Pampouille, E., C. Berri, S. Boitard, C. Hennequet-Antier, S. A. Beauclercq *et al.*, 2018 Mapping QTL for white striping in relation to breast muscle yield and meat quality traits in broiler chickens. *BMC Genomics* 19: 202.
- Pan, Y., Y. Zhou, C. Guo, H. Gong, Z. Gong *et al.*, 2009 Differential roles of the fan-shaped body and the ellipsoid body in *Drosophila* visual pattern memory. *Learn. Mem.* 16: 289–295.
- Papassotiropoulos, A., K. Henke, E. Stefanova, A. Aerni, A. Müller *et al.*, 2011 A

- genome-wide survey of human short-term memory. *Mol. Psychiatry* 16: 184–192.
- Papassotiropoulos, A., K. Henke, E. Stefanova, A. Aerni, A. Müller *et al.*, 2009 A genome-wide survey of human short-term memory. *Mol. Psychiatry* 16: 184.
- Parker, C. C., P. Carbonetto, G. Sokoloff, Y. J. Park, M. Abney *et al.*, 2014 High-resolution genetic mapping of complex traits from a combined analysis of F2 and advanced intercross mice. *Genetics* 198: 103–116.
- Parker, C. C., G. Sokoloff, R. Cheng, and A. A. Palmer, 2012 Genome-wide association for fear conditioning in an advanced intercross mouse line. *Behav. Genet.* 42: 437–448.
- Parmesan, C., and G. Yohe, 2003 A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.
- Partridge, L., B. Barrie, K. Fowler, and V. French, 1994 EVOLUTION AND DEVELOPMENT OF BODY SIZE AND CELL SIZE IN *DROSOPHILA MELANOGASTER* IN RESPONSE TO TEMPERATURE. *Evolution* 48: 1269–1276.
- Pascual, A., and T. Prat, 2001 Localization of long-term memory within the *Drosophila* mushroom body. *Science* 294: 1115–1117.
- Pearce, J. M., 2008a Animal learning and cognition: An introduction, 3rd ed. psycnet.apa.org › record › 2008-01980-000 psycnet.apa.org › record › 2008-01980-000 3: 420.
- Pearce, J. M., 2008b *Animal Learning & Cognition: An Introduction*. Psychology Press.

- Peleg, S., F. Sananbenesi, A. Zovoilis, S. Burkhardt, S. Bahari-Javan *et al.*, 2010
Altered histone acetylation is associated with age-dependent memory
impairment in mice. *Science* 328: 753–756.
- Perkins, L. A., L. Holderbaum, R. Tao, Y. Hu, R. Sopko *et al.*, 2015 The
Transgenic RNAi Project at Harvard Medical School: Resources and
Validation. *Genetics* 201: 843–852.
- Perrimon, N., J.-Q. Ni, and L. Perkins, 2010 In vivo RNAi: today and tomorrow.
Cold Spring Harb. Perspect. Biol. 2: a003640.
- Pertea, M., D. Kim, G. M. Pertea, J. T. Leek, and S. L. Salzberg, 2016 Transcript-
level expression analysis of RNA-seq experiments with HISAT, StringTie
and Ballgown. *Nat. Protoc.* 11: 1650–1667.
- Pertea, M., G. M. Pertea, C. M. Antonescu, T.-C. Chang, J. T. Mendell *et al.*,
2015 StringTie enables improved reconstruction of a transcriptome from
RNA-seq reads. *Nat. Biotechnol.* 33: 290–295.
- Pfeiffer, B. D., T.-T. B. Ngo, K. L. Hibbard, C. Murphy, A. Jenett *et al.*, 2010
Refinement of tools for targeted gene expression in *Drosophila*. *Genetics*
186: 735–755.
- von Philipsborn, A. C., T. Liu, J. Y. Yu, C. Masser, S. S. Bidaye *et al.*, 2011
Neuronal control of *Drosophila* courtship song. *Neuron* 69: 509–522.
- Pickrell, J. K., 2014 Joint analysis of functional genomic data and genome-wide
association studies of 18 human traits. *Am. J. Hum. Genet.* 94: 559–573.
- Pimm, S. L., 2009 Climate disruption and biodiversity. *Curr. Biol.* 19: R595–601.
- Pinheiro, J. C., D. M. Bates, S. DebRoy, D. Sarkar, and

- RDevelopmentCoreTeam, 2013 nlme: Linear and nonlinear mixed effects models. R package version 3.1-109.
- Pintanel, P., M. Tejedo, S. R. Ron, G. A. Llorente, and A. Merino-Viteri, 2019 Elevational and microclimatic drivers of thermal tolerance in Andean *Pristimantis* frogs. *J. Biogeogr.* 46: 1664–1675.
- Pitman, J. L., S. DasGupta, M. J. Krashes, B. Leung, P. N. Perrat *et al.*, 2009 There are many ways to train a fly. *Fly* 3: 3–9.
- Poinsot, D., and H. Merçot, 1997 WOLBACHIA INFECTION IN DROSOPHILA SIMULANS: DOES THE FEMALE HOST BEAR A PHYSIOLOGICAL COST? *Evolution* 51: 180–186.
- Pörtner, H. O., 2002 Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 132: 739–761.
- Pravosudov, V. V., S. P. Mendoza, and N. S. Clayton, 2003 The relationship between dominance, corticosterone, memory, and food caching in mountain chickadees (*Parus gambeli*). *Horm. Behav.* 44: 93–102.
- Pritchard, J. K., J. K. Pickrell, and G. Coop, 2010 The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Curr. Biol.* 20: R208–R215.
- Putz, G., and M. Heisenberg, 2002 Memories in drosophila heat-box learning. *Learn. Mem.* 9: 349–359.
- Qiu, Y., and R. L. Davis, 1993 Genetic dissection of the learning/memory gene

- dunce of *Drosophila melanogaster*. *Genes Dev.* 7: 1447–1458.
- Quinn, W. G., W. A. Harris, and S. Benzer, 1974 Conditioned behavior in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 71: 708–712.
- Rampino, P., G. Mita, S. Pataleo, M. De Pascali, N. Di Fonzo *et al.*, 2009 Acquisition of thermotolerance and HSP gene expression in durum wheat (*Triticum durum* Desf.) cultivars. *Environ. Exp. Bot.* 66: 257–264.
- Ramsey, I. S., M. Delling, and D. E. Clapham, 2006 An introduction to TRP channels. *Annu. Rev. Physiol.* 68: 619–647.
- Rand, D. M., D. M. Weinreich, D. Lerman, D. Folk, and G. W. Gilchrist, 2010 Three selections are better than one: clinal variation of thermal QTL from independent selection experiments in *Drosophila*. *Evolution* 64: 2921–2934.
- Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott *et al.*, 2013 Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8: 2281–2308.
- Rapanelli, M., L. R. Frick, and B. S. Zanutto, 2009 Differential gene expression in the rat hippocampus during learning of an operant conditioning task. *Neuroscience* 163: 1031–1038.
- R Core Team, 2018 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Redon, R., S. Ishikawa, K. R. Fitch, L. Feuk, G. H. Perry *et al.*, 2006 Global variation in copy number in the human genome. *Nature* 444: 444–454.
- Reiter, L. T., L. Potocki, S. Chien, M. Gribskov, and E. Bier, 2001 A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res.* 11: 1114–1125.

- Remolina, S. C., P. L. Chang, J. Leips, S. V. Nuzhdin, and K. A. Hughes, 2012
GENOMIC BASIS OF AGING AND LIFE-HISTORY EVOLUTION IN
DROSOPHILA MELANOGASTER. *Evolution* 66: 3390–3403.
- Renkawek, K., W. O. Renier, J. J. de Pont, O. J. Vogels, and F. J. Gabreëls,
1992 Neonatal status convulsivus, spongiform encephalopathy, and low
activity of Na⁺/K⁺-ATPase in the brain. *Epilepsia* 33: 58–64.
- Rifkin, S. A., J. Kim, and K. P. White, 2003 Evolution of gene expression in the
Drosophila melanogaster subgroup. *Nat. Genet.* 33: 138–144.
- Rinehart, J. P., G. D. Yocum, and D. L. Denlinger, 2000 Thermotolerance and
rapid cold hardening ameliorate the negative effects of brief exposures to
high or low temperatures on fecundity in the flesh fly, *Sarcophaga*
crassipalpis. *Physiol. Entomol.* 25: 330–336.
- Risch, N. J., 2000 Searching for genetic determinants in the new millennium.
Nature 405: 847–856.
- Risch, N., and K. Merikangas, 1996 The future of genetic studies of complex
human diseases. *Science* 273: 1516–1517.
- Rockman, M. V., and L. Kruglyak, 2006 Genetics of global gene expression. *Nat.*
Rev. Genet. 7: 862–872.
- Roff, D. A., 2008 Comparing sire and dam estimates of heritability: jackknife and
likelihood approaches. *Heredity* 100: 32–38.
- Roff, D. A., 1996 THE EVOLUTION OF GENETIC CORRELATIONS: AN
ANALYSIS OF PATTERNS. *Evolution* 50: 1392–1403.
- Roff, D. A., and T. A. Mousseau, 1987 Quantitative genetics and fitness: lessons

- from *Drosophila*. *Heredity* 58 (Pt 1): 103–118.
- Roff, D. A., and R. F. Preziosi, 1994 The estimation of the genetic correlation: the use of the jackknife. *Heredity* 73: 544–548.
- Rolandi, C., J. R. B. Lighton, G. J. de la Vega, P. E. Schilman, and J. Mensch, 2018 Genetic variation for tolerance to high temperatures in a population of *Drosophila melanogaster*. *Ecol. Evol.* 8: 10374–10383.
- Román-Palacios, C., and J. J. Wiens, 2020 Recent responses to climate change reveal the drivers of species extinction and survival. *Proc. Natl. Acad. Sci. U. S. A.* 117: 4211–4217.
- Sandoz, J.-C., and R. Menzel, 2001 Side-Specificity of Olfactory Learning in the Honeybee: Generalization between Odors and Sides. *Learn. Mem.* 8: 286–294.
- Schaid, D. J., W. Chen, and N. B. Larson, 2018 From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat. Rev. Genet.* 19: 491–504.
- Schneider, M. V., and S. Orchard, 2011 Omics technologies, data and bioinformatics principles. *Methods Mol. Biol.* 719: 3–30.
- Schrooten, C., H. Bovenhuis, W. Coppieters, and J. A. Van Arendonk, 2000 Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. *J. Dairy Sci.* 83: 795–806.
- Schubiger, M., Y. Feng, D. M. Fambrough, and J. Palka, 1994 A mutation of the *Drosophila* sodium pump alpha subunit gene results in bang-sensitive paralysis. *Neuron* 12: 373–381.

- Schulte, P. M., T. M. Healy, and N. A. Fangue, 2011 Thermal performance curves, phenotypic plasticity, and the time scales of temperature exposure. *Integr. Comp. Biol.* 51: 691–702.
- Sebat, J., B. Lakshmi, J. Troge, J. Alexander, J. Young *et al.*, 2004 Large-scale copy number polymorphism in the human genome. *Science* 305: 525–528.
- Seeholzer, L. F., M. Seppo, D. L. Stern, and V. Ruta, 2018 Evolution of a central neural circuit underlies *Drosophila* mate preferences. *Nature* 559: 564–569.
- Senan, S., D. Kizhakayil, B. Sasikumar, and T. E. Sheeja, 2014 Methods for Development of Microsatellite Markers: An Overview. *Not. Sci. Biol.* 6: 1–13.
- Seuront, L., K. R. Nicastro, G. I. Zardi, and E. Goberville, 2019 Decreased thermal tolerance under recurrent heat stress conditions explains summer mass mortality of the blue mussel *Mytilus edulis*. *Sci. Rep.* 9: 1–14.
- Sgro, C. M., J. Overgaard, T. N. Kristensen, K. A. Mitchell, F. E. Cockerell *et al.*, 2010 A comprehensive assessment of geographic variation in heat tolerance and hardening capacity in populations of *Drosophila melanogaster* from eastern Australia. *J. Evol. Biol.* 23: 2484–2493.
- Shi, H., G. Kichaev, and B. Pasaniuc, 2016 Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *Am. J. Hum. Genet.* 99: 139–153.
- Shima, Y., Y. Tada, M. Furuki, Y. Hara, and H. Ohta, 1998 A missense mutation of the gene for Na⁺,K⁺-ATPase alpha-subunit causes abnormal feeding behavior in *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.* 248: 778–782.

- Shriner, D., A. Adeyemo, E. Ramos, G. Chen, and C. N. Rotimi, 2011 Mapping of disease-associated variants in admixed populations. *Genome Biol.* 12: 223.
- Shumake, J., S. Furgeson-Moreira, and M. H. Monfils, 2014 Predictability and heritability of individual differences in fear learning. *Anim. Cogn.* 17: 1207–1221.
- Silva, A. J., P. W. Frankland, Z. Marowitz, E. Friedman, G. S. Laszlo *et al.*, 1997 A mouse model for the learning and memory deficits associated with neurofibromatosis type I. *Nat. Genet.* 15: 281–284.
- Silva, A. J., R. Paylor, J. M. Wehner, and S. Tonegawa, 1992a Impaired spatial learning in alpha-calcium-calmodulin kinase II mutant mice. *Science* 257: 206–211.
- Silva, A. J., C. F. Stevens, S. Tonegawa, and Y. Wang, 1992b Deficient hippocampal long-term potentiation in alpha-calcium-calmodulin kinase II mutant mice. *Science* 257: 201–206.
- Sinclair, B. J., C. M. Williams, and J. S. Terblanche, 2012 Variation in thermal performance among insect populations. *Physiol. Biochem. Zool.* 85: 594–606.
- Sitaraman, D., Y. Aso, G. M. Rubin, and M. N. Nitabach, 2015 Control of Sleep by Dopaminergic Inputs to the *Drosophila* Mushroom Body. *Front. Neural Circuits* 9: 73.
- Sitaraman, D., E. F. Kramer, L. Kahsai, D. Ostrowski, and T. Zars, 2017 Discrete Serotonin Systems Mediate Memory Enhancement and Escape Latencies after Unpredicted Aversive Experience in *Drosophila* Place Memory. *Front.*

- Syst. Neurosci. 11: 92.
- Sitaraman, D., H. LaFerriere, S. Birman, and T. Zars, 2012 Serotonin is critical for rewarded olfactory short-term memory in *Drosophila*. *J. Neurogenet.* 26: 238–244.
- Sitaraman, D., M. Zars, H. Laferriere, Y.-C. Chen, A. Sable-Smith *et al.*, 2008 Serotonin is necessary for place memory in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 105: 5579–5584.
- Sitaraman, D., M. Zars, and T. Zars, 2010 Place memory formation in *Drosophila* is independent of proper octopamine signaling. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 196: 299–305.
- Siviter, H., J. Koricheva, M. J. F. Brown, and E. Leadbeater, 2018 Quantifying the impact of pesticides on learning and memory in bees. *J. Appl. Ecol.* 55: 2812–2821.
- Smid, H. M., and L. E. Vet, 2016 The complexity of learning, memory and neural processes in an evolutionary ecological context. *Curr Opin Insect Sci* 15: 61–69.
- Smith, J. M., and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. *Genet. Res.* 23: 23–35.
- Smith, B. R., and S. J. Macdonald, 2020 Dissecting the Genetic Basis of Variation in *Drosophila* Sleep Using a Multiparental QTL Mapping Resource. *Genes* 11.:
- Snell-Rood, E. C., G. Davidowitz, and D. R. Papaj, 2011a Reproductive tradeoffs of learning in a butterfly. *Behav. Ecol.* 22: 291–302.

- Snell-Rood, E. C., G. Davidowitz, and D. R. Papaj, 2011b Reproductive tradeoffs of learning in a butterfly. *Behav. Ecol.* 22: 291–302.
- Snell-Rood, E. C., and D. R. Papaj, 2009 Patterns of phenotypic plasticity in common and rare environments: a study of host use and color learning in the cabbage white butterfly *Pieris rapae*. *Am. Nat.* 173: 615–631.
- Solovieff, N., C. Cotsapas, P. H. Lee, S. M. Purcell, and J. W. Smoller, 2013 Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* 14: 483–495.
- Somero, G. N., 2011 Comparative physiology: a “crystal ball” for predicting consequences of global change. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 301: R1–R14.
- Somero, G. N., 2010 The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine “winners” and “losers.” *J. Exp. Biol.* 213: 912–920.
- Sørensen, J. G., J. Dahlgaard, and V. Loeschcke, 2001 Genetic variation in thermal tolerance among natural populations of *Drosophila buzzatii*: down regulation of Hsp70 expression and variation in heat stress resistance traits. *Funct. Ecol.* 15: 289–296.
- Sørensen, J. G., M. P. Giribets, R. Tarrío, F. Rodríguez-Trelles, M. F. Schou *et al.*, 2019 Expression of thermal tolerance genes in two *Drosophila* species with different acclimation capacities. *J. Therm. Biol.* 84: 200–207.
- Stanley, P. D., E. Ng’oma, S. O’Day, and E. G. King, 2017 Genetic Dissection of Nutrition-Induced Plasticity in Insulin/Insulin-Like Growth Factor Signaling

- and Median Life Span in a *Drosophila* Multiparent Population. *Genetics* 206: 587–602.
- Stark, R., M. Grzelak, and J. Hadfield, 2019 RNA sequencing: the teenage years. *Nat. Rev. Genet.* 20: 631–656.
- Steinberger, D., D. S. Reynolds, P. Ferris, R. Lincoln, S. Datta *et al.*, 2003 Genetic mapping of variation in spatial learning in the mouse. *J. Neurosci.* 23: 2426–2433.
- Stern, D. L., 2014 Identification of loci that cause phenotypic variation in diverse species with the reciprocal hemizygosity test. *Trends Genet.* 30: 547–554.
- Stevenson, R. D., 1985 Body Size and Limits to the Daily Range of Body Temperature in Terrestrial Ectotherms. *Am. Nat.* 125: 102–117.
- Stratman, R., and T. A. Markow, 1998 Resistance to thermal stress in desert *Drosophila*. *Funct. Ecol.* 12: 965–970.
- Sunday, J. M., A. E. Bates, and N. K. Dulvy, 2011 Global analysis of thermal tolerance and latitude in ectotherms. *Proc. Biol. Sci.* 278: 1823–1830.
- Sunday, J. M., A. E. Bates, and N. K. Dulvy, 2012 Thermal tolerance and the global redistribution of animals. *Nat. Clim. Chang.* 2: 686–690.
- Sunday, J., J. M. Bennett, P. Calosi, S. Clusella-Trullas, S. Gravel *et al.*, 2019 Thermal tolerance patterns across latitude and elevation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 374: 20190036.
- Takahashi, T., A. Hamada, K. Miyawaki, Y. Matsumoto, T. Mito *et al.*, 2009 Systemic RNA interference for the study of learning and memory in an insect. *J. Neurosci. Methods* 179: 9–15.

- Takemura, S.-Y., Y. Aso, T. Hige, A. Wong, Z. Lu *et al.*, 2017 A connectome of a learning and memory center in the adult *Drosophila* brain. *Elife* 6.:
- Tang, Y. P., E. Shimizu, G. R. Dube, C. Rampon, G. A. Kerchner *et al.*, 1999 Genetic enhancement of learning and memory in mice. *Nature* 401: 63–69.
- Tello-Ramos, M. C., C. L. Branch, A. M. Pitera, D. Y. Kozlovsky, E. S. Bridge *et al.*, 2018 Memory in wild mountain chickadees from different elevations: comparing first-year birds with older survivors. *Anim. Behav.* 137: 149–160.
- Tello-Ramos, M. C., T. A. Hurly, C. Higgott, and S. D. Healy, 2015 Time–place learning in wild, free-living hummingbirds. *Anim. Behav.* 104: 123–129.
- Thomas, C. D., 2010 Climate, climate change and range boundaries: Climate and range boundaries. *Divers. Distrib.* 16: 488–495.
- Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont *et al.*, 2004 Extinction risk from climate change. *Nature* 427: 145–148.
- Threlkeld, S. W., M. M. McClure, J. Bai, Y. Wang, J. J. LoTurco *et al.*, 2007 Developmental disruptions and behavioral impairments in rats following in utero RNAi of *Dyx1c1*. *Brain Res. Bull.* 71: 508–514.
- Thurmond, J., J. L. Goodman, V. B. Strelets, H. Attrill, L. S. Gramates *et al.*, 2019 FlyBase 2.0: the next generation. *Nucleic Acids Res.* 47: D759–D765.
- Tison, L., S. Holtz, A. Adeoye, Ö. Kalkan, N. S. Irmisch *et al.*, 2017 Effects of sublethal doses of thiacloprid and its formulation Calypso® on the learning and memory performance of honey bees. *J. Exp. Biol.* 220: 3695–3705.
- Tissières, A., H. K. Mitchell, and U. M. Tracy, 1974 Protein synthesis in salivary glands of *Drosophila melanogaster*: relation to chromosome puffs. *J. Mol.*

- Biol. 84: 389–398.
- Tjota, M., S.-K. Lee, J. Wu, J. A. Williams, M. R. Khanna *et al.*, 2011 Annexin B9 binds to β (H)-spectrin and is required for multivesicular body function in *Drosophila*. *J. Cell Sci.* 124: 2914–2926.
- Tolman, E. C., 1924 The Inheritance of Maze-Learning Ability in Rats. *J. Comp. Psychol.* 4: 1.
- Tolwinski, N. S., 2017 Introduction: *Drosophila*—A Model System for Developmental Biology. *Journal of Developmental Biology* 5: 9.
- Tonegawa, S., K. Nakazawa, and M. A. Wilson, 2003 Genetic neuroscience of mammalian learning and memory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358: 787–795.
- Tosini, G., and M. Menaker, 1995 Circadian rhythm of body temperature in an ectotherm (*Iguana iguana*). *J. Biol. Rhythms* 10: 248–255.
- Toure, M. W., F. J. Young, W. O. McMillan, and S. H. Montgomery, 2020 Heliconiini butterflies can learn time-dependent reward associations. *Biol. Lett.* 16: 20200424.
- Turner, T. L., 2014 Fine-mapping natural alleles: quantitative complementation to the rescue. *Mol. Ecol.* 23: 2377–2382.
- Turner, T. L., and P. M. Miller, 2012 Investigating natural variation in *Drosophila* courtship song by the evolve and resequence approach. *Genetics* 191: 633–642.
- Turner, T. L., A. D. Stewart, A. T. Fields, W. R. Rice, and A. M. Tarone, 2011 Population-based resequencing of experimentally evolved populations

- reveals the genetic basis of body size variation in *Drosophila melanogaster*.
PLoS Genet. 7: e1001336.
- Turner, H. N., and S. S. Y. Young, 1969 Quantitative genetics in sheep breeding.
Quantitative genetics in sheep breeding.
- Ueda, A., and C.-F. Wu, 2012 Cyclic adenosine monophosphate metabolism in synaptic growth, strength, and precision: neural and behavioral phenotype-specific counterbalancing effects between dnc phosphodiesterase and rut adenylyl cyclase mutations. J. Neurogenet. 26: 64–81.
- Urban, M. C., 2015 Accelerating extinction risk from climate change. Science 348: 571–573.
- Valdes, A. M., D. Glass, and T. D. Spector, 2013 Omics technologies and the study of human ageing. Nat. Rev. Genet. 14: 601–607.
- VanToai, T. T., S. K. St. Martin, K. Chase, G. Boru, V. Schnipke *et al.*, 2001 Identification of a QTL associated with tolerance of soybean to soil waterlogging. Crop Sci. 41: 1247–1252.
- Vazquez, G., B. J. Wedel, O. Aziz, M. Trebak, and J. W. Putney Jr, 2004 The mammalian TRPC cation channels. Biochim. Biophys. Acta 1742: 21–36.
- Velazquez, J. M., S. Sonoda, G. Bugaisky, and S. Lindquist, 1983 Is the major *Drosophila* heat shock protein present in cells that have not been heat shocked? J. Cell Biol. 96: 286–290.
- Venkatachalam, K., and C. Montell, 2007 TRP channels. Annu. Rev. Biochem. 76: 387–417.
- Verble-Pearson, R. M., M. E. Gifford, and S. P. Yanoviak, 2015 Variation in

- thermal tolerance of North American ants. *J. Therm. Biol.* 48: 65–68.
- Vermeulen, C. J., R. Bijlsma, and V. Loeschcke, 2008a A major QTL affects temperature sensitive adult lethality and inbreeding depression in life span in *Drosophila melanogaster*. *BMC Evol. Biol.* 8: 297.
- Vermeulen, C. J., R. Bijlsma, and V. Loeschcke, 2008b QTL mapping of inbreeding-related cold sensitivity and conditional lethality in *Drosophila melanogaster*. *J. Evol. Biol.* 21: 1236–1244.
- Visscher, P. M., 2008 Sizing up human height variation. *Nat. Genet.* 40: 489–490.
- Visscher, P. M., and J. Yang, 2016 A plethora of pleiotropy across complex traits. *Nat. Genet.* 48: 707–708.
- Vuilleumier, R., T. Lian, S. Flibotte, Z. N. Khan, A. Fuchs *et al.*, 2019 Retrograde BMP signaling activates neuronal gene expression through widespread deployment of a conserved BMP-responsive cis-regulatory activation element. *Nucleic Acids Res.* 47: 679–699.
- Waddell, S., and W. G. Quinn, 2001 Flies, genes, and learning. *Annu. Rev. Neurosci.* 24: 1283–1309.
- Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan *et al.*, 2002 Ecological responses to recent climate change. *Nature* 416: 389–395.
- Wang, T., A. G. Hauswirth, A. Tong, D. K. Dickman, and G. W. Davis, 2014 Endostatin is a trans-synaptic signal for homeostatic synaptic plasticity. *Neuron* 83: 616–629.
- Warren, R., J. VanDerWal, J. Price, J. A. Welbergen, I. Atkinson *et al.*, 2013

- Quantifying the benefit of early climate change mitigation in avoiding biodiversity loss. *Nat. Clim. Chang.* 3: 678–682.
- Webster, N. S., E. S. Botté, R. M. Soo, and S. Whalan, 2011 The larval sponge holobiont exhibits high thermal tolerance. *Environ. Microbiol. Rep.* 3: 756–762.
- Welter, D., J. MacArthur, J. Morales, T. Burdett, P. Hall *et al.*, 2014 The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 42: D1001–6.
- Wen, J., F. Jiang, Y. Weng, M. Sun, X. Shi *et al.*, 2019 Identification of heat-tolerance QTLs and high-temperature stress-responsive genes through conventional QTL mapping, QTL-seq and RNA-seq in tomato. *BMC Plant Biol.* 19: 398.
- Wertheim, B., R. Allemand, L. E. M. Vet, and M. Dicke, 2006 Effects of aggregation pheromone on individual behaviour and food web interactions: a field study on *Drosophila*. *Ecol. Entomol.* 31: 216–226.
- Wheeler, H. E., S. Ploch, A. N. Barbeira, R. Bonazzola, A. Andaleon *et al.*, 2019 Imputed gene associations identify replicable trans-acting genes enriched in transcription pathways and complex traits. *Genet. Epidemiol.* 43: 596–608.
- Widmann, A., M. Artinger, L. Biesinger, K. Boepple, C. Peters *et al.*, 2016 Genetic Dissection of Aversive Associative Olfactory Learning and Memory in *Drosophila* Larvae. *PLoS Genet.* 12: e1006378.
- Williams-Simon, P. A., C. Posey, S. Mitchell, E. Ng'oma, J. A. Mrkvicka *et al.*, 2019 Multiple genetic loci affect place learning and memory performance in

- Drosophila melanogaster*. Genes Brain Behav. 18: e12581.
- Winkler, C. A., G. W. Nelson, and M. W. Smith, 2010 Admixture mapping comes of age. Annu. Rev. Genomics Hum. Genet. 11: 65–89.
- Wood, A. R., T. Esko, J. Yang, S. Vedantam, T. H. Pers *et al.*, 2014 Defining the role of common variation in the genomic and biological architecture of adult human height. Nat. Genet. 46: 1173–1186.
- Wray, N. R., S. Ripke, M. Mattheisen, M. Trzaskowski, E. M. Byrne *et al.*, 2018 Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat. Genet. 50: 668–681.
- Wustmann, G., K. Rein, R. Wolf, and M. Heisenberg, 1996 A new paradigm for operant conditioning of *Drosophila melanogaster*. J. Comp. Physiol. A 179: 429–436.
- Xu, S., 2003 Theoretical basis of the Beavis effect. Genetics 165: 2259–2268.
- Yamaguchi, M., and H. Yoshida, 2018 *Drosophila* as a Model Organism. Adv. Exp. Med. Biol. 1076: 1–10.
- Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher, 2011 GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88: 76–82.
- Yang, J., K. Sun, D. Li, L. Luo, Y. Liu *et al.*, 2019 Identification of stable QTLs and candidate genes involved in anaerobic germination tolerance in rice via high-density genetic mapping and RNA-Seq. BMC Genomics 20: 355.
- Yu, Y., A. P. Hill, and D. A. McCormick, 2012a Warm body temperature facilitates energy efficient cortical action potentials. PLoS Comput. Biol. 8: e1002456.
- Yurkovic, A., O. Wang, A. C. Basu, and E. A. Kravitz, 2006 Learning and memory

- associated with aggression in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 103: 17519–17524.
- Yu, H., F. H. Wan, and J. Y. Guo, 2012b Different thermal tolerance and hsp gene expression in invasive and indigenous sibling species of *Bemisia tabaci*. *Biol. Invasions* 14: 1587–1595.
- Zars, T., M. Fischer, R. Schulz, and M. Heisenberg, 2000a Localization of a short-term memory in *Drosophila*. *Science* 288: 672–675.
- Zars, T., R. Wolf, R. Davis, and M. Heisenberg, 2000b Tissue-specific expression of a type I adenylyl cyclase rescues the rutabaga mutant memory defect: in search of the engram. *Learn. Mem.* 7: 18–31.
- Zars, T., R. Wolf, R. Davis, and M. Heisenberg, 2000c Tissue-specific expression of a type I adenylyl cyclase rescues the rutabaga mutant memory defect: in search of the engram. *Learn. Mem.* 7: 18–31.
- Zhang, Y., L. Huang, J. Zhang, D. Moskophidis, and N. F. Mivechi, 2002 Targeted disruption of hsf1 leads to lack of thermotolerance and defines tissue-specific regulation for stress-inducible Hsp molecular chaperones. *J. Cell. Biochem.* 86: 376–393.
- Zhang, L., A. Martin, M. W. Perry, K. R. L. van der Burg, Y. Matsuoka *et al.*, 2017 Genetic Basis of Melanin Pigmentation in Butterfly Wings. *Genetics* 205: 1537–1550.
- Zhu, L., J. Wilken, N. B. Phillips, U. Narendra, G. Chan *et al.*, 2000 Sexual dimorphism in diverse metazoans is regulated by a novel class of intertwined zinc fingers. *Genes Dev.* 14: 1750–1764.

Zhu, M., and S. Zhao, 2007 Candidate gene identification approach: progress and challenges. *Int. J. Biol. Sci.* 3: 420–427.

VITA

Patricka was born in 1984 on a tiny Caribbean island called Grenada to Patrick Williams and Patricia Grant. She was curious about how things worked throughout her childhood and thus developed an affinity for science earlier in life. She graduated from high school in 2001, then enrolled at the T.A. Marryshow Community College, where she completed two years of business courses. At the age of 19, Patricka met her now-husband, Terry. Shortly after, she migrated to the US to join her parents in Brooklyn, NY. A few years later, she became a mom of a

beautiful baby girl, Terrin. At the age of 25, she enrolled in night classes at Medgar Evers College in pursuit of becoming a physical therapist. At Medgar Evers College, Patricka joined the research team of Dr. Edward Catapane and Dr. Margaret Carroll as an undergraduate researcher in 2012. There, she examined the underlying mechanisms of sensory-motor processes in the eastern oyster. In July of 2013, she moved to Columbia, Missouri, to join the Post-Baccalaureate Research Program (PREP) at the University of Missouri. As a PREP scholar, she worked in Dr. Chris Lorson's laboratory, investigating different therapeutics in a spinal muscular atrophy mouse model. In 2015, Patricka joined the Division of Biological Sciences at the University of Missouri as a graduate student. Shortly after, she joined the laboratories of Dr. Elizabeth King and Dr. Troy Zars. She dissected the genetic basis of learning, memory, and thermal tolerance in a fruit fly model. From this work, she published several papers in peer-reviewed journals and earned several prestigious fellowships and awards. In 2019, Patricka gave birth to a baby boy, Jace. Patricka will be joining the laboratory of Dr. Paul Schmidt as a postdoctoral fellow in the fall of 2021. She will continue to develop as an independent scientist with the future goal of running her laboratory.