DISSECTING THE GENETIC ARCHITECTURE OF COMPLEX TRAITS IN THE NEMATODE *CAENORHABDITIS REMANEI*

by

CHRISTINE H. O'CONNOR

A DISSERTATION

Presented to the Department of Biology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2018

DISSERTATION APPROVAL PAGE

Student: Christine H. O'Connor

Title: Dissecting the Genetic Architecture of Complex Traits in the Nematode *Caenorhabditis remanei*

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology by:

Matthew Streisfeld	Chairperson
Patrick Phillips	Advisor
John Conery	Core Member
William Cresko	Core Member
Michael Harms	Institutional Representative

and

Sara D. Hodges	Interim Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2018

DISSERTATION ABSTRACT

Christine H. O'Connor

Doctor of Philosophy

Department of Biology

June 2018

Title: Dissecting the Genetic Architecture of Complex Traits in the Nematode *Caenorhabditis remanei*

A central problem in evolutionary quantitative genetics has been to attempt to dissect the genetic basis of complex traits. A variety of inferential methods have been developed to probe this issue. Here, I use experimental evolution, next generation sequencing and standing genetic variation in the nematode *Caenorhabditis remanei* to dissect the genetic basis of two model complex traits: oxidative and heat stress response.

Pleiotropy, when one gene affects more than one trait, is an important phenomenon to understand when attempting to understand the genetic architecture of a complex trait. Previous work in the nematode *C. elegans* found that abiotic stress response is controlled by a handful of genes of major effect, and that mutations in one gene can affect the ability of the organism to respond to multiple types of stressors. I used experimental evolution to probe the extent of pleiotropy between the genes selected for resistance to one of two abiotic stressors: acute heat and oxidative. In contrast to expectations, I find that acute heat stress response and acute oxidative response are polygenic, complex traits. Additionally, I find that the evolved responses do not share a genetic basis. This lack of correlation is reflected at the levels of phenotype, gene expression and genomic response to selection.

In addition to the complex interactions within an organism, the genetic architecture

of complex traits and response to selection are affected by population dynamics. Here, I investigate the effect of gene flow on patterns and extent of phenotypic and genetic divergence between populations in distinct environments – a standard lab environment and a chronic heat stress environment. Gene flow of lab-adapted individuals into chronic heat stress adapted populations did not affect phenotypic adaptation, but greatly decreased the number of genomic sites that responded to selection. These results fit predictions that gene flow of non-locally adapted individuals will create an additional barrier for local adaptation, and the strength of selection of locally adapted alleles must not only be greater than the strength of random effects, but also be stronger than the effects of gene flow.

This work includes unpublished co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Christine H. O'Connor

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
University of Illinois, Urbana, Illinois

DEGREES AWARDED:

Doctor of Philosophy, Biology, 2018 University of Oregon, Eugene
Bachelor of Science, Integrative Biology and Anthropology, 2011, University of Illinois

AREAS OF SPECIAL INTEREST:

Evolutionary Genetics
Quantitative Genetics

PROFESSIONAL EXPERIENCE:

Graduate Employee, Department of Biology, University of Oregon, Eugene September 2017 – present

Genomics and Data Science Intern, Monsanto Company, St. Louis, MO May 2017 – August 2017

Graduate Teaching Fellow, Department of Biology, University of Oregon, Eugene September 2011 – June 2012

GRANTS, AWARDS, AND HONORS:

Doctoral Dissertation Improvement Grant, National Science Foundation, 2016 – 2018

Graduate Research Fellow, National Science Foundation, 2014 – 2017

Genetics Training Grant Fellow, National Institutes of Health, 2013 – 2014

Magna cum Laude, University of Illinois 2011

ACKNOWLEDGMENTS

This work would not have been possible without the help of many people. I would first like to thank my advisor Patrick Phillips for his guidance, advice and support throughout the entire graduate school process. I especially appreciate his enthusiasm for complex projects and excitement about results I originally found disappointing and/or confusing. I would also like to thank the other members of my committee: Matt Streisfeld, Bill Cresko, John Conery and Mike Harms for their support and guidance over the years.

I am grateful to all the undergraduate volunteers and Phillips lab members without whom this project would not have been possible. Sally Claridge, Paula Berry Megan Wyatt, Alexis Engelbrecht, and Audrey Hernan were instrumental to creating and maintaining experimental evolution lines and assisting me in data collection. Gavin Woodruff was a fount of ideas and data analysis scripts. Anna Crist and Anna Coleman-Hulbert answered all my basic worm biology questions and accommodated a number of last minute, large plate orders. John Willis prepared many of the next-generations sequencing libraries used in this research. I would also like to thank past and present members of the Phillips lab for their support over the years: Anna Coleman-Hulbert, Ben Blue, Christine Sedore, Erik Johnson, Kristin Robinson, Megan Moerdyk-Shauwecker, Anna Crist, Jailynn Harke, Taylor Wilson, Chadwick Smith, Kristin Sikkink, Heather Archer, Alex De Verteuil, Katja Kasimatis, John Willis, Nadine Timmermeyer, Janna Fierst, Gavin Woodruff, Stephen Banse, Jessica Preston and Anastasia Teterina. I also need to thank research community in the Institute of Ecology and Evolution and in the Department of Biology, in particular Allison Fuiten, Braedan McCluskey, Pete Batzel,

Jason Sydes, Stacey Wagner, Andrea Loes and Madeline Chase. I would also like to thank Doug Turnbull and Maggie Weitzman at the Genomics and cell Characterization Core Facility, for doing all the sequencing described in this dissertation.

Finally, this work would not have been possible without multiple sources of generous support: grants from the National Science Foundation (DEB-1210922 to W. A. Cresko and K. L. Sikkink.; DEB-1607194 to P. C. Phillips and C. H. O'Connor), grants from the National Institutes of Health (GM096008 and GM102511 to P.C. Phillips.; RR032670 to W. A. Cresko), and pre-doctoral fellowships to C. H. O'Connor from the National Science Foundation and National Institutes of Health (T32 GM007413).

To my parents, who fostered my love of learning and exploration in all forms.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

UNDERSTANDING THE GENETIC ARCHITECTURE OF COMPLEX TRAITS

A long-standing goal of evolutionary and quantitative genetics has been to understand the genetic architecture of phenotypes of interest. Here genetic architecture is defined as the number of genes that underlying heritable variation in a trait, their location in the genome and the size of their effects (Cheverud 2006). By understanding the genetic architecture of a trait, or phenotype, of interest, we can begin to understand how a trait can or might respond to selection pressures.

However, this is hard to do (Cheverud 2006; Mackay, Stone, and Ayroles 2009; Rockman 2012; Lewontin 1974). Variation for many phenotypes of interest is quantitative, that is individuals vary almost continuously within a population. (Mackay, Stone, and Ayroles 2009; Falconer and Mackay 1996). Attempting to understand the genetic basis of a trait that is due to variation in tens or hundreds of genes is impossible to do directly via a cross, so indirect methods were developed (Falconer and Mackay 1996). The field of quantitative genetics developed to understand the genetic basis of such traits, and has been widely successful in this endeavor, as has been most evident in plant and animal breeding (Falconer and Mackay 1996; Hill 2014).

The re-discovery of Mendel's work in the early 20th century and the formation of the modern synthesis formed the origin of modern evolutionary and quantitative genetics. Scientists soon connected Mendelian variation with seemingly continuous traits (East 1910; Fisher 1918; Visscher, Wray, Zhang, Sklar, McCarthy, Brown, and Yang 2017b)

using math and inference to understand that inheritance for traits such as height or maize kernel color must be Mendelian, but that multiple factors (i.e. genes) must be involved, unlike with the traits characterized by Mendel's original crosses. Building on this early work, evolutionary quantitative genetics has developed a powerful statistical framework for analyzing complex traits, even without knowing the genes that affect variation for that trait (Barton and Turelli 1989).

INCORPORATING SEQUENCING TECHNOLOGIES

Detecting signatures of selection in polygenic traits can be difficult because signatures of selection over tens or hundreds of genes are subtle (Kemper et al. 2014; Pritchard, Pickrell, and Coop 2010). The advent of sequencing technology also meant the development of a variety of techniques to identify loci that underlie variation in polygenic traits (Wellenreuther and Hansson 2016). These techniques can be broadly categorized into forward and reverse genetic methods. Two of the most widely used forward genetics approaches are Quantitative Trail Loci (QTL) mapping and Genome Wide Association Studies (GWAS), both workhorses in the field of quantitative genetics. QTL mapping and GWAS attempt to map phenotypic variation to genetic variation using genomic markers/single-nucleotide polymorphisms (SNPs), although in different ways. QTL mapping typically uses a population created from a cross of two inbred lines or a population with a detailed pedigree (Slate 2005; Wellenreuther and Hansson 2016). As such QTL mapping has some inherent shortcomings: most studies use inbred populations that only represent a subset of the total genetic variation in a population, as using outbred populations is more difficult (Slate 2005; Lynch and Walsh 1998) but QTL effects can

depend on genetic background, environment and sex (Mackay 2004; Wellenreuther and Hansson 2016), which means that any single QTL mapping study can paint an incomplete picture of the genetic architecture of the trait in question. Genome wide association studies, in contrast, do not require a genetic cross or pedigree information and can be done using highly genetically diverse populations, but they do require more extensive genetic data than QTL mapping studies (Wellenreuther and Hansson 2016), and very large sample sizes: the 'missing heritability' problem identified by early GWAS results turns out to be most likely due to underpowered GWAS (Gibson 2018).

In addition, there are reverse genetics methods. Reverse genetics include the many genome scan and outlier detection methods that have been developed (Wellenreuther and Hansson 2016; Luikart et al. 2003), such as F_{ST} outlier tests (Wright 1951; Lewontin and Krakauer 1973). At their core, all these methods rely on being able to correctly distinguish neutrally evolving from non-neutrally evolving loci, and work on the premise that non-neutrally evolving loci are relevant to the phenotype in question (Przeworski, Coop, and Wall 2005; Bailey and Bataillon 2016; Vitti, Grossman, and Sabeti 2013). However, these methods can be very sensitive to assumptions about selection and demographic history, and may miss many loci, depending on the selection and demographic history of the population in question (Teshima, Coop, and Przeworski 2006; Lotterhos and Whitlock 2015).

Another technique that has become more prominent recently is experimental evolution, or evolve and re-sequence. Briefly, in an experimental evolution project the researcher starts with a genetically diverse population, selects for a trait of interest over multiple generations and then compares the ancestor and descendent populations. As the

proportion of individuals with the trait of interest will have risen over the course of the experiment, assuming the trait is heritable, any genetic changes between the ancestor and descendent populations should identify regions of the genome associated with the selected for trait (Kofler et al. 2015).

As a method for dissecting the genetic basis of a trait of interest, experimental evolution has many strengths over QTL mapping and GWAS. Unlike QTL mapping, experimental evolution can be done in populations with natural levels of standing genetic variation, and does not require pedigree information or creating inbred lines. Experimental evolution can be done with much smaller sample sizes than GWAS – GWAS sample sizes can number in the tens of thousands and the genes identified can still only explain a small proportion of the total variation (Consortium et al. 2014). Finally, with an experimental evolution population, you have near complete knowledge of the selection and demographic history of the population in question. This means that there is no need to make assumptions concerning the demographic history of the population in questions, which is important as demography can affect allele frequency distributions (Hoban et al. 2016), and interpreting results does not rely on assumptions as to how selection will affect the genome – which is good as selection for complex traits may leave ambiguous signatures on the genome (Kemper et al. 2014).

Experimental evolution, like the other methods described above, has pitfalls and potential biases. A good reference genome is needed for data analysis (Kofler et al. 2015), and genomic information can easily be lost during genome alignment or through reduced representation sequencing methods, and errors made during genome assembly can bias results (Hoban et al. 2016). The combination of effective population sizes,

strength of selection, number of replicates and length of experiment necessary for sufficient statistical power can be resource intensive and limiting, depending on the model system used (Taus, Futschik, and Schlötterer 2017; Kofler et al. 2015).

ORGANISMS (AND POPULATIONS) ARE COMPLEX AS WELL

An added complication trying to understand the genetic architecture of a complex trait is that biological organisms are highly integrated systems, and genes may affect variation in more than one trait, the reciprocal side of one trait being affected by multiple genes. Pleiotropy, when one gene affects more than one trait, is an important phenomenon when trying to understand the genetic architecture of complex traits and the ability of a phenotype to respond to selection. Reviews have found that some level of pleiotropy is widespread in a number of different organisms (Paaby and Rockman 2013; Wagner and Zhang 2011). It is easy to see how pleiotropy can constrain evolution – the more interconnected a gene is, the more difficult it will be for that gene to respond to selection on just one trait, since so many others will be affected. However, most of the studies reviewed in Paaby and Rockman (2013) and Wagner and Zhange (2011) focused on what Paaby and Rockman (2013) call molecular gene pleiotropy: the functions a gene has as determined by a knockout study. However, evolution rarely proceeds via knockouts (Wagner and Zhang 2011; Phillips and McGuigan 2006), so understanding the pleiotropic effects of allelic variants or mutations that change gene function may be more relevant to understanding the effects of pleiotropy on an evolutionary timescale.

To further complicate the picture, genetic architecture is not shaped by selection (be it natural or artificial) alone. Population dynamics, such as population size and

population and selection history can also shape the genome and the amount of genetic diversity present in a population. Another important part of understanding the genetic architecture of complex traits is understanding how it, and response to selection, can be affected by these dynamics. While theory and artificial selection experiments often assume otherwise, selection pressures are rarely uniform across space and time. Although classic theories of adaptation and speciation have assumed that divergence in the presence of gene flow would be unlikely (Felsenstein 1981; Ehrlich and Raven 1969; Haldane 1930; Coyne and Orr 2004) and only occur under very strict conditions (Maynard Smith 1966), there are a number of examples of genetic and phenotypic divergence occurring in the face of gene flow, including in white sands lizards in New Mexico (Rosenblum 2006), apple maggot flies (Feder et al. 2003) and maize (Ross-Ibarra, Tenaillon, and Gaut 2009).

Sympatric speciation, speciation in absence of geographic isolation (Coyne and Orr 2004), or divergence in the presence of gene flow, is a perennially controversial topic (Foote 2018). One area of focus in the study of speciation genomics has been on 'genomic islands of divergence' – genomic regions that are more divergent between two populations than would be expected by chance. Genomic islands of divergence are hypothesized to be 'protected' from being broken up by gene flow and thus important for catalyzing the process of reproductive isolation (Via 2012). Specifically, these 'islands' artificially lower gene flow between two populations in certain regions of the genome, increasing divergence. Hitchhiking and physical linkage causes these regions to grow, becoming 'continents of divergence' and leading to reproductive isolation between two populations (Feder et al. 2014; Feder et al. 2013; Via 2012).

However, this theory is controversial. A reanalysis of published data concluded that previously identified 'islands of divergence' were located in regions of reduced genetic diversity, and the increased signal of divergence was due to that reduction in genetic diversity, not an increase in differentiation (Cruickshank and Hahn 2014). This pattern has been seen in a number of different organisms (Han et al. 2017; Renaut et al. 2013), including a group of flycatchers where heterogeneous genomic divergence originally attributed to the interplay between divergent selection and gene flow (Ellegren et al. 2012) was later found to instead be more likely be due to variation in genomic architecture (Burri et al. 2015). Furthermore, there is another explanation to patterns of divergence with gene flow: cause and effect are reversed, i.e. divergence constrains gene flow. Indeed, negative correlations between gene flow and adaptive divergence do not imply causation one way or the other (Räsänen and Hendry 2008). Finally, it can be difficult to robustly test hypotheses concerning divergence and speciation with gene flow, as important parameters including effective population size, population history, strength of selection and migration must be inferred using the same data that is being used to measure divergence.

DISSERTATION OUTLINE

Research described in this dissertation was done using the nematode *Caenorhabditis remanei* as a model organism. *C. remanei* is an ideal model animal system for evolutionary and quantitative genetics. Unlike its more well-known model system cousin, *C. elegans*, *C. remanei* is an obligate sexual reproducer (Diaz, Lindström, and Hayden 2008) and highly genetically diverse (Cutter, Baird, and Charlesworth 2006;

Jovelin, Ajie, and Phillips 2003). *C. remanei* also has a four day generation time (Diaz, Lindström, and Hayden 2008) and large populations can easily be maintained in a small space, both of which are traits that make it ideal for lab based experiments. Like *C. elegans*, *C. remanei* can be cryogenically frozen and revived (Brenner 1974), which means that phenotypic and genetic data can be collected simultaneously from ancestor and descendent populations.

Both projects described in this dissertation investigated the genetic architecture of abiotic stress response, specifically oxidative stress and heat stress. Oxidative stress response and heat stress response are ideal complex traits to study in an experimental evolution context. Abiotic stress response has been found to have pleiotropic effects on aging and longevity in a number of molecular genetics screens (Kenyon 2010; Rodriguez et al. 2013) and previous work in the Phillips lab has shown that oxidative and heat stress response are heritable traits (Reynolds and Phillips 2013), making them tractable targets for experimental evolution. Finally, they are also both relatively easy stressors to create in a laboratory environment.

Experimental evolution was used in both projects described in this dissertation. As described above, experimental evolution is a powerful experimental technique, and *C. remanei*, with its four day generation time and ease of maintenance in a lab environment, is an ideal organism for experimental evolution. Here, experimental evolution allows us to have near perfect knowledge of population history, strength of selection and population size, among other important parameters. There are also ample genomic resources available, including assembled genomes for multiple strains of *C. remanei* (Fierst et al. 2015; A. Teterina and P.C. Phillips in prep).

In Chapter II I use experimental evolution and next generation sequencing to dissect the genetic architecture of an evolved acute heat stress response and acute oxidative stress response. In particular, I explore the role of pleiotropy in the context of trying to dissect the genetic architecture of complex traits. While data from gene knockout studies suggest that abiotic stress response is a controlled by a handful of highly pleiotropic genes and some level of pleiotropy is widespread throughout the genome, less is known about patterns mutational pleiotropy and, importantly, evolution does not often proceed by gene knockout. I use transcriptomics and whole genome sequencing to identify regions of the genome that responded to selection for acute heat stress and acute oxidative stress response, allowing me to identify the number of genes that respond to selection and the extent of overlap between acute heat and acute oxidative stress response.

Chapter II includes unpublished co-authored material. In addition to myself, Kristin L. Sikkink, Thomas C. Nelson, Janna L. Fierst and Patrick C. Phillips contributed significantly to the work described in this chapter.

In Chapter III I expand our understanding of the genetic architecture of complex traits by incorporating population dynamics, specifically reciprocal gene flow of non-locally adapted individuals between two populations. I experimentally test theory concering the effects of gene flow of non-locally adapted individuals on the genetic architecture of adaptation to a novel stress environment, using extensive phenotypic and genomic data. There is a rich field of theory, models and examples concerning the effects of gene flow between two populations on adaptation and divergence at both the level of the phenotype and the genotype, but these theories can be difficult to rigorously test using

examples from the natural world. By using experimental evolution, phenotyping and whole genome sequencing I am able to directly test hypotheses on the effects of gene flow on the extent and patterns of adaptation to a novel, stressful environment and of divergence between two populations.

In Chapter IV, I summarize results from Chapters II and III and discuss how they contribute to our understanding of the genetic architecture and evolution of complex traits and how gene flow affects divergence and the genetic architecture of adaptation.

CHAPTER II

COMPLEX PLEIOTROPIC GENETIC ARCHITECTURE OF EVOLVED HEAT STRESS AND OXIDATIVE STRESS RESISTANCE IN THE NEMATODE *Caenorhabditis remanei*

This chapter includes unpublished, coauthored material. Kristin L. Sikkink created the experimental evolution lines and collected and analyzed the data shown in Figure 1. She also collected and prepared all RNA sequencing libraries. Thomas C. Nelson analyzed data shown in Figures S2.4–S2.7. Janna L. Fierst assisted with original development of the genomic data analysis pipeline and with processing genomic data. I prepared the DNA sequencing libraries and analyzed all transcriptomic and whole genome sequencing data. Patrick C. Phillips was the principle investigator for the work; he and I wrote the manuscript.

INTRODUCTION

Biological organisms are complex, integrated systems. From an evolutionary point of view, a central consequence of this integration is that natural selection acting on one feature of an organism has the potential to have cascading effects across the whole organism (Cheverud 1984; Lande 1979; Phillips and McGuigan 2006). The thought that genetic coupling across functional systems should be the rule rather than exception led Sewall Wright (1968) to posit the idea of "universal pleiotropy" for genetics systems. At its limit, extensive pleiotropy can serve to "lock" the response to selection, creating a genetic constraint that limits the rate of evolutionary change (Arnold 1992; Lande 1979;

B. Charlesworth 1990). Evolutionary quantitative genetics has created a powerful statistical framework for analyzing complex phenotypes, even in the absence of knowledge of seemingly important factors such as the exact number or identity of genes or the distribution of their phenotypic effects (Barton and Turelli 1989). The advent of whole genome analyses would seem to herald the advent of a new era in which we have the power to understand the genetic basis of complex, multifactorial traits. However, even now, the field has not had much success identifying the genetic basis of most adaptive phenotypes within natural populations, aside from traits where a majority of the variation is due to a handful of well characterized mutations, such as coat color and oxygen affinity in deer mice (Linnen et al. 2013; Natarajan et al. 2013; Steiner, Weber, and Hoekstra 2007) or body size variation in horses (Makvandi-Nejad et al. 2012). There may be a small handful of cases in which natural variation is generated by simple genetics, but it is clear that in many, if not most, traits this will not be the case (Rockman 2012). For example, human height is a well-studied, highly heritable, quantitative trait and yet despite sample sizes that now stretch into the millions, mapped loci still only account for a small fraction of the total genetic variance (Nolte et al. 2017; Lango Allen et al. 2010). Similarly, studies of complex diseases such as schizophrenia are able to map hundreds of genes associated with disease prevalence but still account for less than 10% of variation in disease risk (Consortium et al. 2014).

The picture grows even more complicated when considering multiple traits. Several studies have found median levels of pleiotropy to be between 4–7 traits affected per gene (Wagner and Zhang 2011; Paaby and Rockman 2013), depending on the organism, but there are many cases in which a single mutation can affect dozens of traits

(Knight et al. 2006). These studies tend to focus on what Paaby and Rockman (2013) defined as molecular gene pleiotropy: how many functions does a given locus affect when that gene is knocked out. There is substantially less information on the extent of pleiotropy within populations generated by naturally segregating variation, or "mutational pleiotropy" (Paaby and Rockman 2013; Wagner and Zhang 2011). Here the focus is on the entire spectrum of pleiotropy generated by segregating alleles as opposed to knockouts per se. The distinction is critical because, for evolutionary change, the allele and not the gene is the appropriate unit of variation since alleles themselves can have variation in pleiotropic effects (Wagner and Zhang 2011; Phillips and McGuigan 2006; Wright 1968; Lande 1984). Studies that used forward genetic screens may find a large fraction of potential targets of those genes, but that does not necessarily inform us as to how a trait of interest would respond to selection pressures in a population over an evolutionary time-scale.

How can we understand the balance between molecular and mutational pleiotropy, especially in the context of naturally segregating variation influencing complex correlated traits? Experimental evolution coupled with whole genome sequencing, or evolve and re-sequence, has emerged as a powerful method for studying the genetic architecture of complex traits (Kofler and Schlötterer 2013). Selection on complex traits can leave weak or ambiguous signals of selection on the genome (Kemper et al. 2014), but evolve and re-sequence allows us to directly compare ancestral and evolved populations to identify divergent regions of the genome. In particular, replicated experiments within a well-defined hypothesis testing framework allow sampling variance to be accounted for in a way that historical contingency renders impossible within natural

populations. While we have learned a great deal from experimental evolution in microbial populations initiated from a single fixed genetic background, questions related to the influence of complex segregating variation on the response to selection can only be addressed by capturing a broad array of natural genetic variation in the initial ancestor population (Kofler et al. 2015).

The nematode *Caenorhabditis remanei* is particularly ideal for addressing questions relating to complex trait via experimental evolution (Sikkink, Reynolds, et al. 2014). It is a highly genetically diverse, sexually reproducing organism (Cutter, Baird, and Charlesworth 2006; Jovelin, Ajie, and Phillips 2003). Large populations can easily be maintained in a lab, with upwards of 2000 individuals living on one 10cm agar plate, so an entire experiment can live one shelf in an incubator. And of course, its close relationship to the model nematode *C. elegans* allows both genomic and functional information to be applied across species.

Indeed, abiotic stress response in nematodes provides an interesting model system of interacting complex traits. In particular, many of the important stress response pathways are well conserved throughout animals and have interesting pleiotropic effects on aging and longevity (Kenyon 2010), with a handful of stress response pathways being know to affect the response to many different abiotic stressors (Rodriguez et al. 2013; Murphy and Hu 2013). For example, as elucidated within numerous studies with *C. elegans*, mutations in one well-studied pathway, the insulin/insulin-like growth factor signaling pathway, affect resistance to a stressors such as heat stress, oxidative stress, ultraviolet radiation and pathogen stress (Murphy and Hu 2013).

Here, we use experimental evolution and whole genome sequencing to dissect the genetic architecture of acute heat stress and acute oxidative stress response in *C. remanei*. In previous work, we saw a strong phenotypic response to selection for both acute heat stress and acute oxidative stress, but in contrast to expectations set by previous research in *C. elegans* (Rodriguez et al. 2013), there was little to no correlated phenotypic response between selection for acute heat stress and acute oxidative stress resistance (Sikkink et al. 2015). Previous work with these populations also indicates that there are complex changes in gene regulation patterns in acute heat stress evolved populations (Sikkink, Reynolds, et al. 2014; Sikkink, Ituarte, et al. 2014). We seek to align these global phenotypic patterns with a detailed analysis of the genomic response to selection at single nucleotide resolution in order to test complexity and coherence of genotypic and phenotypic responses to selection.

MATERIALS AND METHODS

Experimental evolution

As described in (Sikkink, Reynolds, et al. 2014), the ancestor population was created using 26 isofemale strains that were crossed to create a population representative of the naturally segregating genetic variation of a single population of *C. remanei* collected in Ontario, Canada. This population (PX443) was used to establish each of the replicate selection and control lines and was frozen soon after creation to prevent adaptation to the lab environment. As determined used the methods described below, this population had high initial levels of genetic diversity (genome wide average $\pi = 0.0178$, Figure S2.1). All natural isolates and experimental lines were raised on Nematode

Growth Medium-lite (NGM-lite, U.S. Biological) seeded with *Escherichia coli* strain OP50 (Brenner 1974).

Populations of *C. remanei* were evolved to one of three acute stress environments: control, acute heat stress or acute oxidative stress as previously described in Sikkink *et al* 2014b and Sikkink *et al* 2015. Here, an acute stress is one that the worms experienced for four hours as L1 juveniles before returning to the laboratory control environment. The acute heat stress assay induced ~70% mortality in the ancestor population and the acute oxidative stress assay induced ~80% mortality in the ancestor population (Sikkink et al. 2015). Populations were subjected to acute stress conditions every two generations or whenever the population produced ≥ 24,000 eggs, whichever happened later. Control populations were randomly culled to 1000 worms during selective generations. For all acute stress events, each population (control evolved, acute heat stress evolved and acute oxidative stress evolved) was age synchronized via a bleach treatment (Stiernagle 2006). All experimentally evolved populations were subject to a total of ten acute stress events. As described in (Sikkink, Reynolds, et al. 2014) each population was frozen after about every second generation of acute stress selection to ensure that the worms did not lose the ability to survive a freeze and to provide a record of evolutionary changes. There were a total of eight experimentally evolved populations: four control replicates, two acute heat stress replicates and two oxidative stress replicates.

Preparation of whole genome sequencing libraries

We determined the genetic response to selection using a population-based pool-seq approach (Schlötterer et al. 2014). Ancestor and acute evolved populations were sequenced via whole genome shotgun sequencing. Approximately 10,000 L1 individuals

(~24 hours old juvenile) were pooled together for DNA extraction and sequencing library preparation. Samples were prepared using the Nextera DNA Sample Preparation Kit (Illumina). Multiple sequencing libraries were prepared from the original DNA sample as needed to ensure adequate mean sequencing coverage (Table S1). Whole genome sequencing libraries were obtained for the ancestor and all eight evolved populations. Samples were sequenced on an Illumina HiSeq 2000 or on an Illumina HiSeq 2500. All sequences were deposited on NCBI's Sequence Read Archive, accession number SRP126594 (https://www.ncbi.nlm.nih.gov/sra/SRP126594).

Whole genome sequencing analysis

Raw sequence reads were quality filtered using the 'process_shortreads' component of *Stacks* (Catchen et al. 2013). Low quality reads were trimmed or discarded; all reads that passed quality filtering were aligned to the *C. remanei* reference genome assembled from the PX356 *C. remanei* reference strain (Fierst et al. 2015) using the alignment program *GSNAP* (Wu and Nacu 2010). Single-nucleotide polymorphism (SNP) tolerant alignment parameters were used to ensure that divergent SNPs would align to the reference genome. The software *OrthoFinder* (Emms and Kelly 2015) was used to identify *C. elegans* orthologs of *C. remanei* genes and protein domains. The *C. elegans* protein annotations were obtained from Wormbase ParaSite (Lee et al. 2018), version WS258.

We used the R (R Core Team 2017) package *Nest* (Jonas et al. 2016) to estimate effective population size (N_e) of each experimental replicate using temporal changes in allele frequency based on existing theory (R. K. Waples, Larson, and Waples 2016; Jorde and Ryman 2007; Krimbas and Tsakas 1971), but modified for samples taken from

pooled population data. The input data, allele frequency and coverage data for each SNP at generation 0 (ancestor population) and generation 30 (each evolved population) was obtained from the 'snp-frequency-diff.pl' program from *Popoolation2* (Kofler, Pandey, and Schlötterer 2011). Every SNP had to have a minimum coverage of 10X in each population to be included in the effective population size estimator. Adding a maximum coverage threshold did not affect the estimated N_e. We used a pool sample size (the number of individuals who went into the original sequencing sample) of 1000 but increasing the pool size from 1000 to 10,000 individuals had a negligible effect on the estimated N_e. Expected N_e based on census population sizes for the duration of the experiment were calculated based on Equation 3.5 in (Kimura 1983) using the 'harmonic.mean' function in the R (R Core Team 2017) package *psych* (https://CRAN.R-project.org/package=psych).

Nucleotide diversity (π) in the ancestor population was calculated over 1000 base pair (bp) windows using a program in the software *Popoolation* (Kofler et al. 2011). Allele frequency differences were analyzed on a site by site basis between the ancestor and control, heat stress evolved or oxidative stress evolved populations using the Fisher's exact test program from the software *Popoolation2* (Kofler, Pandey, and Schlötterer 2011). Each SNP had to have a minimum of 10X coverage, a maximum coverage of 98% of total sequence coverage and the minor allele had to have at least two copies at a site in order to be retained for further analysis. Allele frequency differences were also analyzed over a range of sliding window sizes, with each sliding window overlapping with one third of the length of the previous window. For each sliding window 75% or more of the bases in the sliding window had to meet the minimum and maximum coverage thresholds

used for the single locus Fisher's exact test comparison in order to be included. The geometric mean *p*-value of Fisher's exact test across the window was calculated for each interval.

We used a permutation test to determine whether or not a sliding window was significantly diverged from the ancestor population using a random sampling program in which windows were created at random by calculating the geometric mean *p*-value for a set of Fisher's exact test outcomes from the pool of single-locus results. The mean number of SNPs in an empirical sliding window of a particular size was used to determine the number of SNPs to be used for each random sample window. Random sampling was repeated 1,600,000 times for each ancestor-experimentally evolved population pair. Random windows equivalent to empirical 1000–20,000 bp windows were created, with the 5000 bp window size chosen for the reported results because it is approximately the same size or slightly larger than nearly all genes in *C. remanei*, ensuring that genic and near genic regions could be captured in one window (see Figures S2.2 and S2.3 for distributions of empirical versus random sample windows). Every ancestor-evolved population pair produced roughly 20,000 5000 bp windows. We used the Bonferroni corrected *p*-value for 20,000 tests (2.5×10^{-6}) as the cut-off for significance in each random sample window file. Any empirical sliding window that had a mean Fisher's exact test value greater than or equal to the 99.99975% of a random sample window was considered to be significantly differentiated from the ancestral population.

A given window was defined as being significantly divergent within a given selection regime if it did not overlap with divergent windows in the control populations and the window was classified as significant within each replicate. *BEDTools* 'intersect'

(Quinlan and Hall 2010) was used to identify overlapping genomic windows and genes under significantly divergent genomic windows.

Haplotype phasing using single worm RAD sequencing

We tested for the potential buildup of linkage disequilibrium within selection lines by crossing worms from the reference genome strain (PX356) with the strain used in the experimental evolution project (PX443), thereby generating F1s heterozygous for a single ancestral haplotype. Single worm RAD sequencing libraries were prepared for 79 male F1 worms. Because the genotype of the reference strain is known, the haplotype of the target crossed individual could be directly determined in the F1 heterozygotes.

The RAD capture (Ali et al. 2016) method was used to genotype F1 worms. RAD capture is a flexible reduced-representation sequencing method which allowed us to multiplex and economically sequence all 79 genomes in a single Illumina lane. We digested genomic DNA with the restriction enzyme *EcoRI*; ligated digested fragments to barcoded, biotinylated adaptors; and sheared fragments with a BioRuptor. Samples were sequenced on an Illumina HiSeq 4000 at the University of Oregon Genomics Core Facility.

Raw fastq files were reoriented using flip2BeRAD (https://github.com/tylerhether/Flip2BeRAD), which reorients read pairs such that all RAD tags and inline barcodes are present as read 1. Retained reads were filtered and demulitplexed with 'process_radtags' from *Stacks* using standard stringencies. Reads were aligned to the *C. remanei* PX356 reference genome (Fierst et al. 2015) using *GSNAP* ultrafast settings. We only used read-pairs that mapped to the reference in the expected orientation (pairs facing inward toward each other) in downstream analyses. Individuals were genotyped using the

'HaplotypeCaller' and 'GenotypeGVCF' programs from *GATK* (https://software.broadinstitute.org/gatk/) (Van der Auwera et al. 2013), and *VCFtools* (Danecek et al. 2011) was used to filter sites based on missing data, minor allele frequencies and genomic position. *PLINK* (Purcell et al. 2007) was used to calculate linkage disequilibrium on the 13 largest contigs, which account for ~54% of total genome length.

Transcriptional profiling of evolved populations

RNA sequencing libraries were prepared as described in (Sikkink, Reynolds, et al. 2014). Briefly, total RNA was isolated using standard TRIzol methods from tissue samples of the ancestor population, one of the control populations from replicate two (hereafter referred to as control – two), and the acute heat stress and acute oxidative stress evolved populations from the second evolutionary replicate. All tissue samples were collected from L1 larval worms. We collected 6–8 replicate samples from each evolved line from at least two independently thawed populations from each evolved line, except for the oxidative stress evolved line, where 6 replicate samples were collected from one thaw. Samples were sequenced on an Illumina HiSeq 2000 at the University of Oregon Genomics Core Facility.

Raw sequence reads were quality filtered as described in (Sikkink, Ituarte, et al. 2014). Briefly, raw sequencing reads were quality filtered using the 'process_shortreads' component of Stacks (Catchen et al. 2013) and all reads that passed quality filtering were aligned to *a C. remanei* reference genome assembled from the PX356 *C. remanei* reference strain (Fierst et al. 2015) using *GSNAP* (Wu and Nacu 2010). We used the 'htseq-count' tool from the Python *HTSeq* package (Anders, Pyl, and Huber 2014) to

count all reads that aligned to protein coding genes, using the criteria described in (Sikkink, Reynolds, et al. 2014).

Differential gene expression analysis was conducted using the R package *DESeq2* (Love, Huber, and Anders 2014). Low variance genes were filtered using *DESeq2*'s automatic filter. We tested for differences in gene expression between the ancestor and three experimentally evolved populations (control, acute heat stress evolved and acute oxidative stress evolved). Genes were only called as differentially expressed in each comparison if the *p*-value was less than 0.05 after false discovery correction. We also tested for a relationship between log2 fold changes in gene expression between acute heat stress–ancestor and acute oxidative stress–ancestor comparisons using Model II linear regression. This analysis was implemented in the *lmodel2* package (Legendre 2018) in R (R Core Team 2017).

RESULTS

Genetic architecture of acute stress response is complex

As demonstrated previously, selection for acute heat stress and acute oxidative stress resistance leads to the evolution of significantly increased survival within the acute heat stress and acute oxidative stress environments (Figure 2.1) (Sikkink et al. 2015). Using whole-genome resequencing of the selected lines, we find that this phenotypic response is strongly echoed across the entire genome, with 13–27% of genomic windows within each population more divergent than expected by chance (Figure 2.2). There was a large amount of genetic divergence between not only the ancestor and stress evolved populations, but also between the ancestor and control evolved populations (Figure 2.2).

Because there was essentially no response to either an acute heat stress or an acute oxidative stress in any of the control evolved populations at the phenotypic level (Figure 2.1), the genetic divergence is most likely due to the effects of genetic drift and off-target selection, in particular adaptation to the lab environment, since this population had very recently been brought into the lab from nature (Sikkink, Reynolds, et al. 2014). For this reason, we focused on divergent genomic regions in the acute stress populations that were not divergent in any of the control populations (that is, unique) and that were shared among the evolutionary replicates (Figure 2.2). Those regions are referred to "divergent windows" in what follows. After this filtering, we were able to identify hundreds of reproducibly divergent windows: 495 and 491 divergent windows in the acute heat stress evolved and acute oxidative stress evolved populations, respectively.

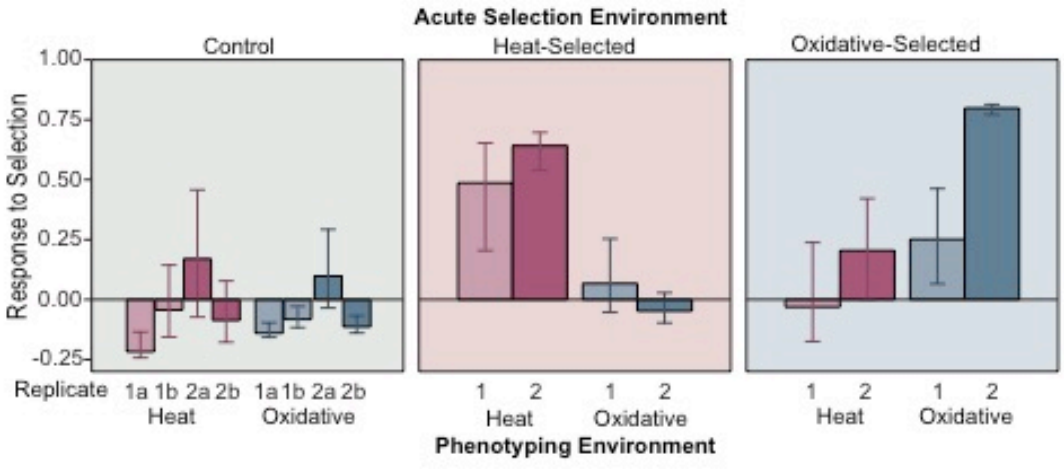
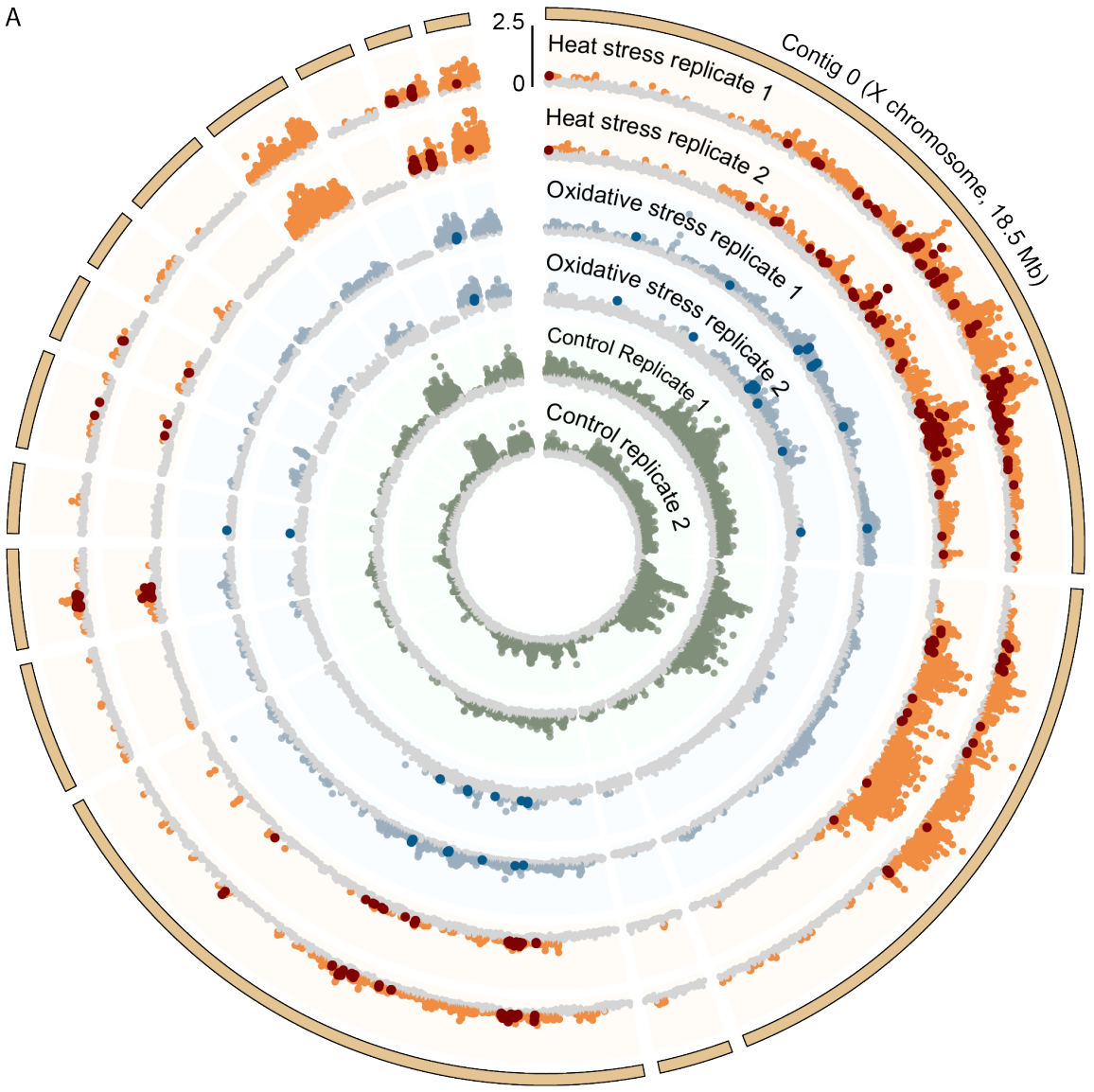


Figure 2.1: Direct and correlated response to selection for each replicate of experimentally evolved lines. A direct response to selection is measured by the response to selection where the phenotyping environment matches the acute selection environment and a correlated response to selection is measured by the response to selection where the phenotyping environment does not match the acute selection environment. There is a strong direct response to selection seen in both acute heat and oxidative selected populations but no correlated response to selection. Modified from Supplementary Figure S1 in Sikkink et al. 2015.


A
2.5
0
Heat stress replicate 1
Heat stress replicate 2
Oxidative stress replicate 1
Oxidative stress replicate 2
Control Replicate 1
Control replicate 2
Contig 0 (X chromosome, 18.5 Mb)
Divergent window in a heat stress evolved population
Divergent window in an oxidative stress evolved population
Divergent window in a control evolved population
Divergent window in both heat stress evolved populations not present in a control population
Divergent window in both oxidative stress evolved populations not present in a control population

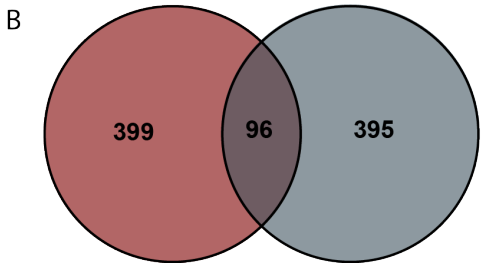

B
399
96
395

Figure 2.2: Genomic response for selection to either an acute heat stress or acute oxidative stress. (a) Location of significantly divergent 5000 bp genomic windows. Mean $-\log_{10}$ of the Fisher's Exact Test p-value for each window shown. The two control populations for each replicate are plotted on same ring. Widespread divergence from the ancestor (non-grey dots) is seen in all evolved populations, including the control populations. (b) Overlap between unique divergent windows between heat stress and oxidative stress evolved populations. Only around one fifth of the unique divergent windows overlap in location.

There is little overlap of divergent genomic windows between heat stress and oxidative stress evolved populations

Stress resistance, including resistance to heat and oxidative stress, is expected have a common genetic basis in nematodes (Kenyon 2005; Rodriguez et al. 2013). In contrast to these expectations, we observed little correlated phenotypic response to either acute heat stress resistance or acute oxidative stress, even though there was a strong evolutionary response within a given selective environment (Figure 2.1). This phenotypic prediction of little between-stress pleiotropy is strongly reflected in the pattern of evolutionary response at the genetic level. Of the 491 acute oxidative stress evolved and 495 acute heat stress evolved divergent windows, only 96 (~19%) overlap (Figure 2.3). This indicates a general lack of pleiotropy in the genes underlying the response to selection for heat and oxidative stress.

Linkage disequilibrium as a possible source of complexity in genetic response

The large number of genomic regions that we saw responding to selection could in fact be due to a large number of genes underlying the traits or interest or could potentially be due to an apparent response to selection generated by allele frequency changes at linked loci caused by limited sampling of haplotypes during the process of adaptation (Franssen et

al. 2015). We used single individual SNP data in the ancestral population to estimate the degree of linkage disequilibrium within these populations. We found that linkage disequilibrium decays to background levels within 200 bp (Figures S2.4 and S2.5), very similar to the pattern observed within natural populations (Cutter, Baird, and Charlesworth 2006), and found no evidence of unusual long-distance linkage disequilibrium in the ancestor population beyond that (Figures S2.6 and S2.7). Overall, then, it is unlikely that the response to selection that we observe was generated by pairwise linkage disequilibrium or cryptic haplotype structure within the population.

Effective population size

Estimates of effective population size (N_e) for each of the evolved populations range from 379-1602 (Table 2.1). This is much smaller than the expected N_e given our census population sizes (Table 2.2), which fluctuated between 1,000–24,000 worms (Sikkink et al. 2015). We also obtained N_e estimates using SNPs found only on the X chromosome. *C. remanei* has an XX/X0 sex determination system (Thomas, Woodruff, and Haag 2012; Hodgkin 1987) so we would expect N_e on the X chromosome to be around 75% of the N_e on the autosomes. However, our estimates show X chromosome N_e values that range from 60% to 146% of the all SNP N_e estimates.

Pleiotropic structure of gene expression

The lack of correlated response to selection in the stress resistance phenotypes is echoed at the level of the molecular phenotype of gene expression as well. In this case, however, many thousand responses can be interrogated simultaneously. Here, changes in expression are based on the comparison of baseline expression in the evolved populations

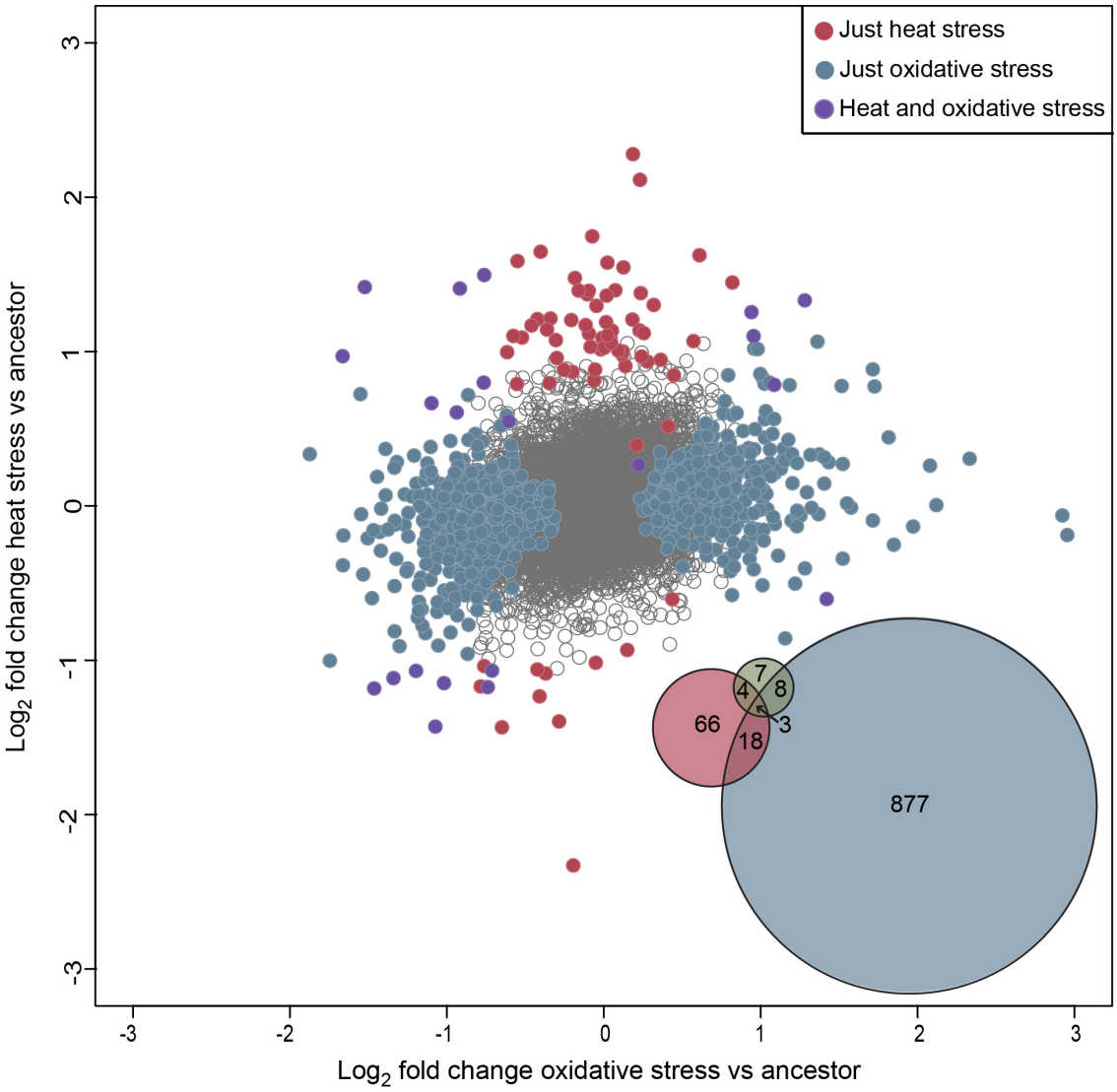


Figure 2.3: Correlation between log2 fold changes of gene expression in heat stress evolved vs. ancestor and oxidative stress evolved vs. ancestor tests for differential gene expression. The venn diagram in the lower right shows the overlap in identity of significantly differentially expressed genes in the heat stress evolved, oxidative stress evolved and control evolved populations.

Population	N_e	X Chromosome N_e
Replicate 1 Control – 1	650	619
Replicate 1 Control – 2	781	479
Replicate 1 Heat stress	978	640
Replicate 1 Oxidative stress	1602	1449
Replicate 2 Control – 1	1051	1507
Replicate 2 Control – 2	821	805
Replicate 2 Heat stress	654	528
Replicate 2 Oxidative stress	379	557

Table 2.1: Estimated effective population (N_e) size for each experimentally evolved population. N_e estimated for all eight evolved populations using allele frequency and coverage data estimated using *Popoolation2* (Kofler, Pandey, and Schlötterer 2011). N_e for all SNPs and just SNPs on the assembled X chromosome shown.

Population	Expected N_e
Control evolved	~2,500 – ~2,800
Heat stress evolved	~5,000 – ~6100
Oxidative stress evolved	~3,500 – ~4600

Table 2.2: Expected effective population sizes for the three experimental evolution conditions given the cyclical census population sizes described in Sikkink et al. (2015). Ranges come from different intermediate generation census population sizes. N_e values calculated using equation 3.5 in Kimura (1983).

versus those in the ancestor population. Not only do the numbers of differentially expressed genes greatly differ across selective environments (906 in the oxidative stress evolved population versus 91 in the heat stress evolved populations), but also the genes that are differentially expressed are largely non-overlapping – only 21 in common (Figure 2.3). In addition, we observed an interesting pattern in the directionality of the evolution of gene expression. Genes differentially expressed in the heat stress evolved population were more likely to be up-regulated than down-regulated compared to both the oxidative stress evolved population (Fisher's exact test, $p = 1.7 \times 10^{-8}$) and the control population

(Fisher's exact test, $p = 1.9x10^{-4}$), while the proportion of up- versus down-regulated genes in the oxidative stress evolved population was not significantly different from the control population (Fisher's exact test, $p = 0.29$). There was also a weak correlation between changes in gene expression between oxidative-stress evolved and heat-stress evolved populations ($r^2 = 0.049$, $p = 0.01$,), but this correlation is appears to be largely driven by a relationship between the genes that have not changed in expression very much (i.e., it is a weak aggregate response over many genes).

Candidate genes under selection for acute heat stress and acute oxidative stress

Genes implicated as potentially being involved in the phenotypic response to selection, either via differences in gene expression or divergence due to allele frequency changes tended to create non-overlapping sets among the stress-response environments. Specifically, only three genes were found to be in common for the differential expression and genomic divergence gene sets within the heat stress evolved populations and 16 in common for the oxidative stress evolved populations.

Divergent genes due to allele frequency changes

There were 591 genes contained within significantly divergent genomic windows in the heat stress evolved populations and 590 genes contained within significantly divergent genomic regions in the oxidative stress evolved populations. Of those divergent genes, 126 are shared in common between the heat stress and oxidative stress evolved populations (again, roughly 20%). There are a only handful of obvious candidate genes found within the divergent genomic regions in the heat stress evolved populations: a gene with a heat shock chaperonin-binding motif, *daf-8*, three insulin-like growth factor binding proteins, and *Ras*. The heat shock chaperonin-binding motif is found on the

stress-inducible phosphoprotein STI1, known to bind heat stress proteins (Song et al. 2009). Daf-8 (Figure 2.4A) is a member of the TGF-β pathway, one of the signaling response pathways involved in heat shock response in *C. elegans* (Rodriguez et al. 2013). Ras is a member of the MAPK signaling pathway, which has been found to be responsive to pathogen (Ewbank 2006) and endoplasmic reticulum stress (Darling and Cook 2014). Finally, insulin like growth factor binding proteins control the distribution and activity of insulin like growth factor receptor, an important stress response transmembrane receptor protein (Murphy and Hu 2013; Rodriguez et al. 2013). There is one divergent gene in the oxidative stress evolved populations with a relevant stress response function: SKI-interacting protein, which is a member of the TGF- β signaling pathway that promotes dauer formation when active (Savage-Dunn 2005). There are hundreds of other genes in these regions and, perhaps most noticeably, for the most part they do not contain members of the canonical stress response pathways that have been determined via mutagenesis approaches.

Differentially expressed genes

Similar to the pattern of divergent genes due to changes in allele frequency, there were a handful of significantly differentially expressed genes that are likely to be functionally important for abiotic stress response. There were constitutively up-regulated heat shock proteins (HSPs) in both the heat shock evolved and oxidative shock evolved populations, and constitutively up-regulated insulin like growth factor binding proteins in the oxidative shock evolved population. As was found in previous work (Sikkink, Ituarte, et al. 2014), we find that up-regulated HSPs are significantly enriched in the heat stress

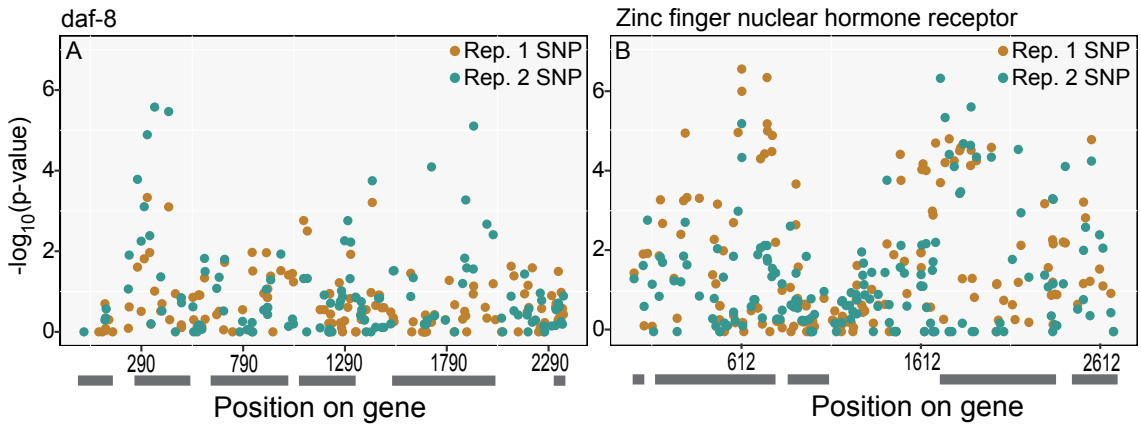


Figure 2.4: Fisher's exact test results on a SNP by SNP basis for heat stress evolved populations for two genes. (a) –log10 *p*-values of Fisher's Exact Test for each SNP in *daf-8* in the heat stress evolved populations. daf-8 affects dauer development and is a component of the TGF-β signaling pathway (Savage-Dunn 2005). Grey bars indicate exons. (b) –log10 *p*-values of Fisher's Exact Test for each SNP in a zinc finger nuclear hormone receptor. This gene has lowest mean Fisher's Exact Test *p*-value of the divergent genes in the heat stress evolved populations. Grey bars indicate exons.

evolved population (Fisher's exact test, $p = 0.00146$), as well as an enrichment of HSPs that are constitutively up-regulated in the oxidative stress evolved population (Fisher's exact test, $p = 0.029$). In addition to up-regulated HSPs, SKN-1, a transcription factor whose regulatory targets are important for oxidative stress resistance (An and Blackwell 2003; Murphy and Hu 2013) was also significantly up-regulated in oxidative stress evolved population.

DISCUSSION

Within each organism, tens of thousands of genes interact within one another to generate hundreds of thousands of potential phenotypes. The classic assumption within the field is that majority of phenotypic variation found within natural populations is quantitative, with the phenotypic differences among individuals likely caused by a large

number of genes operating within a complex environmental context (Hill 2009). Yet over the last two decades a small controversy has emerged about whether we expect the genetic basis of adaptation to be dominated by a few genes of major effect or whether it will be highly polygenic (H. A. Orr and Coyne 1992; H. A. Orr 2009; Rockman 2012). While this question is obviously most relevant in the context of evolution within specific natural populations, there is an unavoidable conflation of adaptation, demographic history, and chance events for any particular realization of an evolutionary lineage that makes it impossible to test specific hypotheses within a single instance. For example, even in the case of putatively adaptive bone loss in sticklebacks, in which it initially appeared that the same syndrome had been replicated many times independently during adaptation to fresh water, we now know that the genetic basis of this adaptation in each case is generated by a complex set of alleles that have likely been forged by thousands of generations of migration-selection balance (Nelson and Cresko 2018). Experimental evolution provides a means of addressing these questions in a controlled, repeatable manner. Our previous work has demonstrated that the nematode *C. remanei* is a powerful system for experimental evolution of stress resistance phenotypes (Sikkink, Reynolds, et al. 2014; Sikkink et al. 2015). Here, we use an evolve and resequence approach to probe the genetic basis of these phenotypic changes.

Our data demonstrate that both heat stress response and oxidative stress response are complex, polygenic traits, where hundreds of regions in the genome respond to selection (Figure 2.2) and hundreds of genes show constitutive changes in gene expression (Figure 2.3). This fits the pattern shown by many other experimental evolution and artificial selection studies (Pettersson et al. 2013; Turner et al. 2011; Orozco-

terWengel et al. 2012; Hirsch et al. 2014), in which researchers have found hundreds of divergent genes and/or genomic regions after selection for traits as varied and seed size, novel temperature regimes and body size, although there are a few notable exceptions (Makvandi-Nejad et al. 2012).

The infinitesimal model, which allows researchers to predict the effects of selection on a trait without knowing the identity of the genes that underlie variation in that trait (Hill 2009; Fisher 1918), forms the cornerstone of quantitative genetics (Falconer and Mackay 1996) and has been very successful in plant and animal breeding (Hill 2014). Recent work in human genetics has found that that the vast majority of human traits and diseases are caused by variation in hundreds if not thousands of loci (Wood et al. 2014; Visscher, Wray, Zhang, Sklar, McCarthy, Brown, and Yang 2017a). Taken in the context of this larger body of work, our study supports the view that much variation is quantitative in nature and responses to selection are generated by changes in hundreds of genes, even when selective itself is extremely strong (Hirsch et al. 2014; Pettersson et al. 2013; Johansson et al. 2010).

The genomic basis of mutational pleiotropy

Despite the expectation set by many molecular biology studies of stress response pathways in *C. elegans* (Murphy and Hu 2013; Rodriguez et al. 2013; Kenyon 2010), we found that there was little to no correlated response to selection for heat stress or oxidative stress for the other stress at either the level of survival (Figure 2.1) (Sikkink et al. 2015) or gene expression (Figure 2.3). And at the level of genomic response to selection we also find little evidence for pleiotropy between the heat stress evolved and oxidative stress evolved populations at the level of individual genes (Figure 2.2). This of

course does not mean that pleiotropy is not a characteristic of many, or even most, of the genes involved in stress response network, simply just for those that respond to selection here. First, a correlated response to selection is generated by the alleles that actually respond to selection, not all possible genes (Lande 1979). It is possible that many of the stress response genes that have been generated via mutagenesis have broadly pleiotropic effects such that they can not respond to direct selection for increased stress resistance per se because of potential negative pleiotropic effects on other traits (e.g., developmental rate, reproduction). Second, molecular geneticists and quantitative geneticists focus on related but subtlety different definitions of the term pleiotropy. Most fundamentally, pleiotropy is a property of alleles, not genes (Phillips and McGuigan 2006), and so it is possible for there to be variation in pleiotropy at a single locus such that some segregating alleles will lead to a correlated response to selection while other will not. For stress resistance, we do not find a clear signal for canonical stress response pathway members in the first place, and so variance in pleiotropy at genes defined via measures of molecular pleiotropy still does not appear to be the primary driver here, regardless of any potential variation in allelic effects.

Finally, on the timescale of this study, selection can only realistically operate on segregating variation. For the reasons described above, it is simply possible that there is more segregating variation at sites with fewer pleiotropic effects than more. The overall picture that emerges from this point of view is that variation in pleiotropy may be as important for structuring the evolutionary response to selection on complex traits as many other classical parameters such allele frequency and average effects on single traits

(Sikkink et al. 2015). Here, pleiotropy acts as a functional sieve that determines how the genetic system as a whole structures the response to selection on complex traits.

The genetic and genomic basis of the response to selection

While there are a handful of genetically divergent and differentially expressed genes with functions that are directly relate to abiotic stress response, it is clear that the response to acute heat stress and acute oxidative stress cannot be explained by these genes alone. We do not find evidence that many of the canonical stress response genes, including *hsf-1* or *daf-16* (Murphy and Hu 2013; Rodriguez et al. 2013) for heat stress and *pink-1, lrk-1* and *sod-1*, *-2* or *-3* for oxidative stress (Rodriguez et al. 2013) responded to selection for acute heat stress or acute oxidative stress survival, although *skn-1* is constitutively up-regulated in the oxidative stress evolved population. This is further evidence that gene knockouts can paint an incomplete picture of the genetic architecture of a phenotype of interest and the fact that the genes that respond to selective pressures in a population are unlikely to be highly pleiotropic genes of large effect.

Some clustering of the genomic response can be clearly seen on Figure 2.2. For example, 22% of the divergent windows in the two heat stress evolved populations are found on the X chromosome. This raises the possibility that linkage disequilibrium is responsible for some of the response to selection in the evolved populations. However, we found that linkage disequilibrium decays to background levels within 200 bp within the ancestor population (Figures S2.4 and S2.5) Additionally, the same clustering of genomic response is not seen in either of the oxidative stress evolved populations (Figure 2.2), where at most 10% of divergent windows were found on one contig (not shown in Figure 2.2). However, when we look at strongly selected genes we do see evidence of

local selective sweeps, where multiple SNPs have similarly highly significant Fisher's exact test *p*-values and changes in allele frequency (Figure 2.4B, Figure S2.8).

We find that most our estimated effective population sizes are about an order of magnitude smaller than the post-selection census population size (Sikkink et al. 2015)). Fluctuating population sizes will reduce N_e (Kimura 1983), and in our case very strong selection has the potential strongly reduce the level of variation at linked sites via the Hill-Robertson effect (Hill and Robertson 1966; B. Charlesworth 1996; B. Charlesworth 2012; Comeron, Williford, and Kliman 2007). Even so, the reduction in N_e seems to be particularly severe relative to the very large population sizes (for an animal) carefully maintained throughout the experiment. For example, in spite of having a much larger census population size than similar experiments in *Drosophila melanogaster*, our estimated N_e values are only two–three times higher than estimated N_e values for those populations, despite us maintaining census sizes that are 10-20X larger (Jonas et al. 2016; Orozco-terWengel et al. 2012). Differences in overall strength of selection may play a role here (strength of selection is unknown in Jonas et al. 2016), but one additional source of reduction in N_e is potential variance in mating success in which male mating success and/or female reproductive output displays a greater than Poisson variance (Kimura 1983; Wright 1931). This seems particularly likely because the control populations have similar effective population sizes to the selected lines. Variation in male mating ability has been previously documented within *C. remanei* (Palopoli et al. 2015), and bears further investigation within the context of experimental evolution within this species.

CONCLUSION

While organisms are complex, integrated systems and the manner in which natural variation percolates though that system is likely to be highly structured by functional interactions. Although pleiotropy may be ubiquitous throughout the genome, individuals with mutations in highly pleiotropic genes are likely to suffer from reduced fitness since those mutations will most likely affect more traits than those under selection. As such, seeking to understand the genetic architecture of traits solely via an analysis of gene knockouts is likely to leave an incomplete picture of the overall structure of the pleiotropic network, especially with respect to segregating variation within natural populations. Using standing genetic variation, experimental evolution and whole genome sequencing therefore provides powerful tool for understanding the genetic architecture of complex traits, no matter how complex they ultimately prove to be.

BRIDGE

In Chapter II I demonstrated that over an evolutionary timescale, acute heat stress and acute oxidative stress are both complex traits with a unique genetic basis. Both of these results are in contrast to expectations set by molecular biology studies, and demonstrate the importance of incorporating standing genetic variation when trying to understand the genetic architecture of complex traits. However, organisms exist within populations; and it is the population, not the individual organism, that is the unit of evolution. In this next chapter I expand my focus from complex interactions between genes within an organisms to complex interactions between two populations. In Chapter III I investigate how gene flow of non-locally adapted individuals affects genetic and phenotypic patterns of adaptation to a novel, chronically stressful environment.

CHAPTER III

HOW SELECTION AND GENE FLOW INTERACT TO AFFECT THE GENETIC ARCHITECTURE OF ADAPTATION TO A NOVEL, STRESSFUL ENVIRONMENT

INTRODUCTION

In the wild, species often exist in subpopulations spread across heterogeneous environments. While classic theories of adaptation and speciation have assumed that divergence in the presence of gene flow of non-locally adapted individuals should be very unlikely (Felsenstein 1981; Ehrlich and Raven 1969; Haldane 1930; Coyne and Orr 2004) and only proceed under very stringent conditions (Maynard Smith 1966), observations of natural populations reveal that such divergence does indeed occur, perhaps frequently. For example, as has been seen in oceanic and freshwater stickleback (Barrett, Rogers, and Schluter 2008), Alaskan salmon (Larson et al. 2016), white sands lizards in New Mexico (Rosenblum 2006), maize (Ross-Ibarra, Tenaillon, and Gaut 2009) or pea aphids in alfalfa and clover fields (Via 1991).

Gene flow is not without consequence of course. Early theoretical work on divergence under migration-selection dynamics demonstrated that locally adapted alleles could be maintained, as long as the migration rate does not pass a critical threshold relative to the strength of selection (Levene 1953; Maynard Smith 1970; Bulmer 1972; Felsenstein 1976). However, it was also shown that ongoing migration of non-adapted individuals will slow down or even prevent adaptation to a novel environment because of the continuous influx of non-adapted individuals, or even lead to the loss of locally advantageous alleles if the migration rate is stronger than selection for that allele

(Haldane 1930; Felsenstein 1976). Although, if selection is strong enough to overcome the homogenizing effects of migration, divergence can still occur (Yeaman and Whitlock 2011; Flaxman et al. 2014; Felsenstein 1976; Haldane 1930).

In particular, it is predicted that alleles of small effect will be swamped by the influx of non-adaptive individuals under migration (Haldane 1930). Under standard conditions of one interbreeding population in a homogenous environment, the strength of selection at an individual locus necessary to overcome random processes such as genetic drift is directly proportional to the effective population size (N_e); that is the selection coefficient at an individual locus only needs to be greater than $1/2N_e$ in diploids (Kimura 1983; Falconer and Mackay 1996). With gene flow of non-locally adapted individuals, however, the strength of selection (s) at an individual locus would also need to be greater that the effect of migration (m) at that locus, i.e. $s > m > 1/2N_e$ (Haldane 1930). This additional filter of migration should then lead to only loci or haplotypes of large effect for selected traits becoming divergent between populations (Kirkpatrick and Barton 2006; Yeaman and Whitlock 2011; Yeaman 2015; Yeaman and Otto 2011). Migration is also predicted to affect divergence in other ways. For example, in simulations, Griswold (2006) found that migration leads alleles of large effect to be responsible for a larger proportion of differences between two populations as compared to the situation where there is no migration between the two populations.

While reciprocal gene flow should increase the correlated response to selection in a novel environment, the effect of migration of non-locally adapted individuals on adaptation to this novel environment will depend on the genetic architecture of the trait that confers adaptation to the novel environment (here we assume fitness can be captured

in one trait). Imagine a two-population system, where one population is formed when a subset of the original population moved into a novel environment from an ancestral environment, generating novel selection pressures. If there is no gene flow there should be no correlation for fitness in the novel environment between the ancestral and novel environment adapted populations (Figure 3.1A). However, with reciprocal gene flow, the type and extent of a correlated response will depend on the nature of potential genetic constraints present.

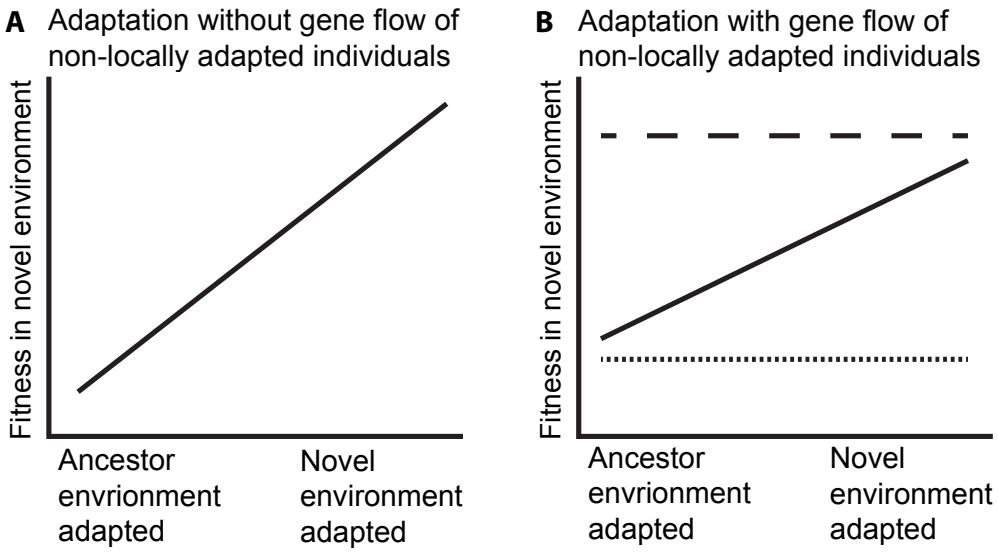


Figure 3.1: Possible phenotypic adaptation scenarios for adaptation to a novel environment with and without gene flow. (A) Expected pattern of adaptation without reciprocal gene flow of non-locally adapted individuals. (B) Possibly patterns of adaptation to a novel environment with gene flow of non-locally adapted individuals. Solid line: genetic basis of adaptation to the novel environment is highly additive, and interactions such as pleiotropy, epistasis or dominance are minimal. Dotted line: genetic basis of adaptation to the novel environment is constrained; for example the strength of selection against alleles that positively affect adaptation in the novel environment in the ancestral environment is stronger than positive selection in the novel environment so no adaptation occurs. Dashed line: There is no constraint on alleles that are positively selected in the novel environment so they spread to ancestor environment and both populations show high fitness in the novel environment.

If the trait underlying fitness is highly additive with many alleles of small effect, we would expect to see a dampened response in adaptation to the novel environment within the population adapted to the novel environment and an increased level of fitness in the novel environment in the population adapted to the ancestral environment (solid line, Figure 3.1B). However, if there are strong genetic constraints on adaptation, such as the alleles that positively affect adaptation to the novel environment have an even stronger negative pleiotropic effect on fitness in the ancestral environment, another possibility is that migration of non-locally adapted individuals to the novel environment will prevent adaptation to the novel environment (dotted line, Figure 3.1B). A third possibility is that the alleles positively affecting adaptation to the novel environment are selectively neutral to alleles locally adapted to the ancestral environment; that is there is no cost to adaptation. Positively selected alleles for the novel environment could then be preserved in high frequency in both environments and both populations could have high fitness levels in the novel environment (dashed line, Figure 3.1B).

Testing hypotheses about patterns of genomic divergence under divergent selection with gene flow is difficult because it is usually only possible to look at a single snapshot of an ongoing process. For example, the current rate of gene flow may not reflect historical rates that actually led to currently seen patterns of genomic divergence and/or the true history of population divergence may be incomplete, incorrect or impossible to measure indirectly (Nosil and Feder 2012). Important variables, including population history, population size and migration rate, are usually unknown and must be estimated from the same data being used to analyze divergence itself. While there are numerous examples of putative divergence with migration in the literature, (Larson et al.

2016; Graham et al. 2017; Pfeifer et al. 2018) such natural experiments have limited predictive capacity.

Many of these issues can be addressed by studying migration-selection dynamics within an experimental framework, which provides the ability to simultaneously control population history, level of genetic variation, strength of selection, population size, and migration rate in a replicated experimental design. In this study we use experimental evolution and whole genome sequencing of the nematode *Caenorhabditis remanei* to investigate the effects of migration on the ability of a population to adapt to a novel, stressful environment – chronic heat stress. Our goal is to test theory concerning the patterns of genetic divergence seen in diverging populations in allopatry versus with gene flow and to better isolate the genetic basis of heat stress resistance. Studying the interplay between migration and selection can offer a novel approach to studying the genetics of complex traits, as migration is predicted to filter out background genetic noise (Nosil, Funk, and Ortiz-Barrientos 2009) as well as alleles with small effects, leading to a more clarified genetic architecture consisting of alleles of large effect (Yeaman and Whitlock 2011; Flaxman et al. 2014).

MATERIALS AND METHODS

Creation of experimental populations

The initial base population of *C. remanei* was established from a natural collection of individuals from Ontario, Canada, as described in (Sikkink, Reynolds, et al. 2014). Briefly, strain PX443 was created from 26 isofemale strains that were crossed to generate a population with a high amount of naturally segregating variation. The baseline

experimental evolution population used in this study was then derived from this population after it had been adapted to laboratory conditions for 30 generations (Sikkink et al. 2015) and (Sikkink, Reynolds, et al. 2014). One of the control evolved populations from this earlier work was further adapted to the laboratory for an additional 45 generations, generating a total of 75 generations in the laboratory environment before any additional selection was imposed. For this phase of lab adaptation, the population was kept at a census population size of 50,000–60,000 individuals and was frozen every 10-12 generations to generate a population record and so that individuals would not lose the ability to recover from freezing.

The generation 75 population was frozen and used as the ancestor population for all subsequent work. For all phases of this project worms were frozen using the Soft Agar Freezing Solution protocol described in Stiernagle (2006). All natural isolates and experimental evolution lines described were raised on Nematode Growth Medium-lite (NGM-lite, U.S. Biological) seeded with *Escherichia coli* strain OP50 (Brenner 1974).

Experimental evolution

Populations of the generation 75 lab adapted ancestral population were evolved to one of two selection environments (ancestral control lab environment of 20°C or chronic heat stress environment of 31°C) and one of three migration treatments (no migration, 5% between ancestral environment and chronic heat stress environment population pairs, or 20% between ancestral environment and chronic heat stress environment population pairs), resulting in a total of six possible experimental treatments (Figure 3.2). The chronic heat stress temperature of 31°C was chosen after pilot studies showed that this strain of *C. remanei* is sterile at temperatures of 31.5°C and above (Figure S3.1). Each

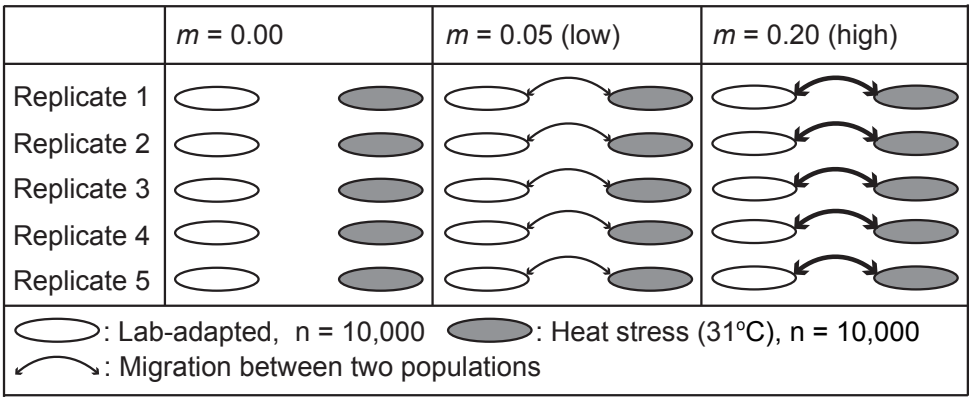


Figure 3.2: Experimental design. Experimental evolution was carried out for 40 generations. Five independent replicate population pair (one pair = a lab-adapted population and a heat stress adapted population) were created for each migration treatment

migration–selection treatment combination was independently replicated five times (Figure 3.2). A replicate consisted of one population in the ancestral environment and one population in the chronic heat stress environment. Throughout this paper 'population' will refer to one replicate of a particular migration-selection treatment combination, so there were 30 experimental populations total. There were approximately 10,000 individuals per population (Figure 3.2).

Worms were transferred to agar plates every four days, for a total of 40 cycles. During each transfer a 20 μm Nitrix filter (purchased from Dynamic Aqua-Supply, item #NTX20) was used to ensure that only juvenile worms would be transferred to the subsequent generation. Worm concentration in S basel solution, a common salt buffer used in *C. elegans* research (Stiernagle 2006), was estimated to calculate the appropriate number of individuals to transfer into each environment. Reciprocal migration was also done at this stage. There were 2000 worms per plate; if there was a migration treatment the appropriate concentration of worms was plated such that the number of native plus

migrant individuals equaled 2000. For example, for the $m = 0.05$ treatment populations, 100 migrant and 1900 native worms were plated each generation.

Fecundity Assays

We measured lifetime female fecundity as a signal of adaptation to the heat stress environment. We choose not to use survival to adulthood as an alternative fitness measure because the proportion of individuals surviving to adulthood was highly variable between pilot trials and generally very close to the number of survivors in the control environment (Figure S3.2).

Fecundity assays were conducted in the 31°C heat stress environment for both control and selected populations after 20, 30 and 40 generations of experimental evolution. For these assays, populations that were evolved in the ancestral control environment were moved to the 31°C environment one generation prior to the assessment of fecundity to ensure that assays were conducted on acclimated individuals. Populations were age synchronized via bleaching and hatched in S basel for 16-20 hours before being placed on agar plates.

Between 20–30 female worms were assayed from each population, yielding between 100–150 female worms per migration-selection combination across the five replicates. *C. remanei* is an obligate outcrossing/gonochoristic species (Diaz et al. 2008) (as opposed to the self-fertile *C. elegans*), and so for the assays themselves, one L4/young adult female and two L4/young adult males were picked to a 4 mL agar plate. All three worms were transferred to new plates every 22–26 hours, for 5 days total. If a male worm was not found on the Day 1 transfers then an additional male worm was added to the fecundity-assay plate from the population plate. After the transfer to the Day

2 plates, if a male worm was either dead or missing then they were not replaced. As long as there was at least one male and one female found on a plate those individuals were transferred to the next day's plate. If the female worm was found without any live males she was transferred to the next day's plate but not to new plates for any subsequent days remaining in the assay. If the female worm was either dead or missing, the assay was ended for that female. Individuals did not have to survive for the full 5 day of the assay to be included in overall analysis, as there is natural variation in lifespan within this species, especially under mated conditions (Palopoli et al. 2015). Fecundity assays were also performed for all generation 40 evolved populations in 20°C. The protocol was the same as described for the 31°C assay, except that the assay was carried out for 7 days instead of 5, as at 20^{o}C 90% of reproductive output occurs by day 6 (Diaz, Lindström, and Hayden 2008). The fecundity assay was only carried out for 5 days at 31^{o}C because the majority of worms had died by day 5 at that temperature.

Statistical Analysis

We tested for evolved differences in lifetime female fecundity using a generalized linear mixed model (GLMM) with a square root link and Poisson distribution in the *lme4* package (Bates et al. 2015) in R (R Core Team 2017). Migration, selection and the interaction between migration and selection were included as fixed effects. Replicate population and researcher collecting the data were included as random effects. In addition, individual was included as a random effect to correct for overdispersion.

We also tested for differences in proportions of sterile versus fecund females, using lifetime fecundity data collected during fecundity assays. A female worm was defined as sterile if she had 0 or 1 offspring over her lifetime. As *C. remanei* is a sexually

reproducing species we considered having less than the replacement rate number of offspring to be effectively sterile. We fit a logistic regression model with a binomial distribution in the *lme4* package (Bates et al. 2015) in R (R Core Team 2017). Input data were the number of sterile and fecund worms for each replicate of a migration–selection treatment replicate (Figure 3.2). Migration, selection and the interaction between migration and selection were included as fixed effects. Tukey contrasts were used in a post-hoc analysis to determine which group means were significantly different from each other using the 'glht' function in the R package *multcomp* (Hothorn, Bretz, and Westfall 2008).

Whole genome sequencing

We used a pooled population sequencing approach to determine the genetic response to selection, as described in Chapter II of this dissertation (Schlötterer et al. 2014). Pooled population whole genome shotgun sequencing data was obtained for the lab adapted ancestor population and all 30 evolved populations at generation 40 of experimental evolution (that is at the end of the experiment). Samples were prepared using Illumina's Nextera library preparation kit. Multiple sequencing runs were done as necessary to obtain sufficient mean coverage. Samples were sequenced on an Illumina HiSeq 2500 and Illumina HiSeq 4000.

Whole genome sequencing data analysis

Raw sequence reads were quality filtered using the 'process_shortreads' component of *Stacks* (Catchen et al. 2013). Reads that passed quality filtering were aligned to the *C. remanei* reference genome assembled from the PX506 *C. remanei*

reference strain (A. Teterina and P.C. Phillips in prep) using *GSNAP* (Wu and Nacu 2010).

Allele frequency differences were analyzed on an individual single-nucleotide polymorphism (SNP) basis between the ancestor and each evolved population. We used the Cochran-Mantel-Haenszel test to identify consistent changes in allele frequency between the ancestor population and each of the five replicates of a migration–selection treatment. To be analyzed, each SNP had to have a minimum of 10x coverage, a maximum coverage of 98% of total sequence coverage and the minor allele had to have at least two copies at a site. We used the Bonferroni corrected *p*-value and filtered all SNPs for a minimum allele frequency difference of 0.25 between the ancestor and each evolved population. The allele frequency filter was added to eliminate sites called as significant due to high coverage or random changes, such as genetic drift. Fisher's exact test was used to identify significantly divergent loci between the ancestor population and each evolved population separately. Both the Conchran-Mantel-Haenszel test and Fisher's exact test were run using scripts from *PoPoolation2* (Kofler, Pandey, and Schlötterer 2011).

The significance of allele frequency changes were also analyzed over 1000 base pair (bp) windows. We used a permutation test to determine whether or not a sliding window was significantly diverged from the ancestor population, as described in Chapter II. Briefly, we randomly sampled SNPs from the output of single locus Fisher's exact test results and calculated the means to create a genome-location agnostic window. The mean number of SNPs in empirical 1000 bp sliding windows was used to create each random sliding window; 2.3 million random windows were created for each evolved population.

An empirical (genome location based) window was divergent if it had a mean value greater than random sliding windows in the $5x10^{-7}$ upper percentile. To measure overlap in response to selection we calculated the proportion of divergent sliding windows in one migration-selection treatment replicate population that overlapped with all five populations in that migration–selection treatment combination. To test if either the migration or selection treatment fit the overlap in response to selection between replicates, we fit a logistic regression model with a binomial distribution in the *lme4* package (Bates et al. 2015) in R (R Core Team 2017). Input data were the number of divergent windows that overlapped between all divergent windows in a migration–selection treatment and the total number of divergent windows for each population. Migration, selection and the interaction between migration and selection were included as fixed effects. Tukey contrasts were used in a post-hoc analysis to determine which group means were significantly different from each other using the 'glht' function in the R package *multcomp* (Hothorn, Bretz, and Westfall 2008).

Nucleotide diversity was calculated for the ancestor and all evolved populations using a perl program available in the software *PoPoolation* (Kofler et al. 2011).We also calculated F_{ST} values on an individual SNP basis between each evolved population pair shown in Figure 3.2 to test if migration affected the amount of structure between two populations. We used the F_{ST} calculation found in Hartl and Clark (1997) implemented in *PoPoolation2* (Kofler, Pandey, and Schlötterer 2011). We calculated the genome wide mean F_{ST} for each evolved replicate pair (15 samples total) and fit an analysis of variance to determine if migration had a significant effect on genome wide mean F_{ST} using R 'ANOVA' and 'lm' functions (R Core Team 2017).

We used the R (R Core Team 2017) package *PoolSeq* (Jonas et al. 2016) (https://github.com/ThomasTaus/poolSeq) to estimate effective population size (N_e) of each experimental replicate using approaches for temporal allele frequency changes to estimate N_e, using established equations (R. K. Waples, Larson, and Waples 2016; Jorde and Ryman 2007; Krimbas and Tsakas 1971) modified for samples taken from pooled population data. The input data, allele frequency and coverage data for each SNP from the ancestor population and generation 40 (each evolved population) was obtained from the 'snp-frequency-diff.pl' program from *Popoolation2* (Kofler, Pandey, and Schlötterer 2011).

RESULTS

Response to selection for chronic heat stress survival and migration

Adaptation to the chronic heat stress environment was rapid and varied with migration treatment (Figure 3.3). Female fecundity increased in all chronic heat stress adapted populations ($F_{(9,1507)} = 114.98$, $p < 2x10^{-16}$), but only lab-adapted female worms that experienced high migration ($m = 0.20$) had an increased number of offspring ($z = 6.78$, $p = 1.19x10\text{-}11$); the low migration rate ($m = 0.05$) treatment did not affect lab-adapted female fecundity ($z = 1.471$, $p = 0.142$). Fecundity changed between the three time-points measured ($F_{(9,1507)} = 10.51$, $p < 2x10^{-16}$), although not in the same way for different migration and selection treatment combinations. In particular, the fecundity of chronic heat stress populations without migration of non-locally adapted individuals increased rapidly and plateaued at generation 30, while fecundity in chronically heat stress selected populations undergoing low migration increased rapidly at generation 20 but then decreased from generation 20 to generation 30.

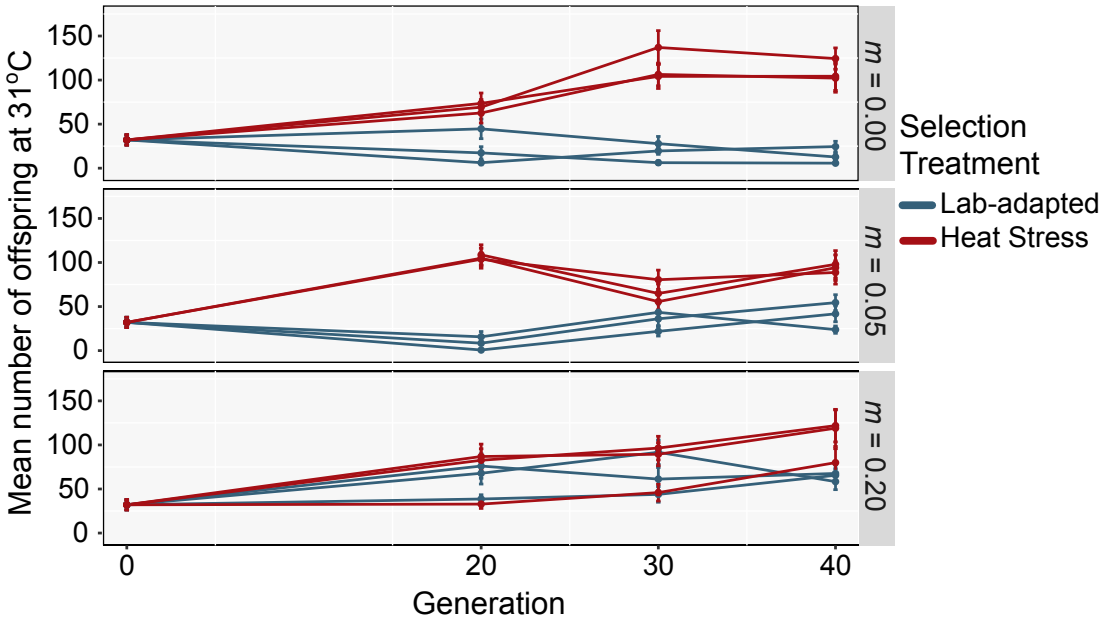


Figure 3.3: Mean female fecundity though time for three evolved replicate pairs. Each row is a different migration treatment. Mean plus and minus standard error shown. Generation 0 is the 75 generation lab adapted base population. All heat stress evolved populations show an increase in female fecundity relative to the ancestor population, although patterns of adaptation vary with migration treatment. Female fecundity also increases in the low (m = 0.05) and high (m = 0.20) migration lab-adapted populations.

In order to analyze the overall evolutionary response in more detail, we assayed female fecundity from all evolved populations at generation 40. As was seen in the time series sample of evolved populations, female fecundity increased significantly in all heat stress adapted lines (Figure 3.4). Migration of heat stress evolved individuals into the lab adapted populations increased female fecundity in the heat stress environment ($F_{(7,860)}$ = 19.76, $p < 2\text{x}10^{-16}$). The positive effect of migration on female fecundity demonstrates that there was successful gene flow, not just migration. While migration of heat stress adapted individuals into the lab environment had a positive effect on female fecundity in the heat stress environment, there was no cost of gene flow at generation 40 in the heat stress environment (Figure 3.4), which was unexpected. Finally, there was also a

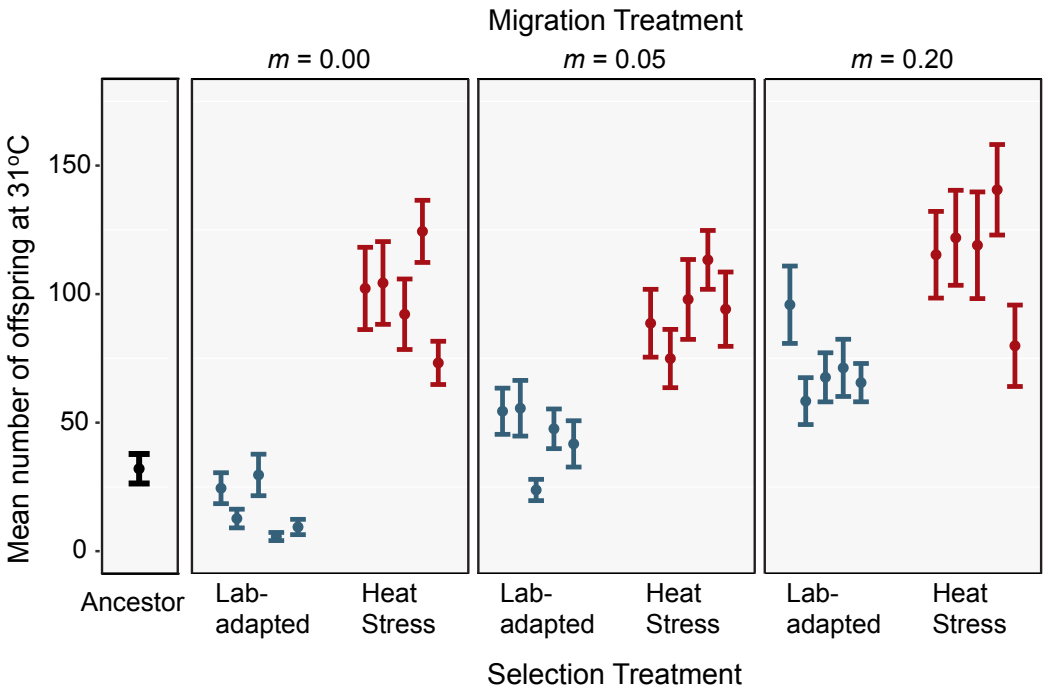


Figure 3.4: Female fecundity of 75 generation lab-adapted ancestor population and all evolved populations at generation 40 in 31°C. Mean plus and minus standard error shown. All heat stress adapted populations show an increase in female fecundity. Female fecundity also increased in the low (m = 0.05) and high (m = 0.20) migration lab-adapted populations, and there is a significant interaction between migration and selection.

significant interaction between migration and selection ($F_{(7,860)} = 24.67$, $p < 2\text{x}10^{-16}$) shown by the increase in fecundity in lab-adapted females who experienced migration of heat stress adapted individuals, compared to the lab-adapted no migration populations (Figure 3.4 and S3.3). Migration of heat-stress adapted individuals was also sufficient to increase the proportion of fecund females in the low and high migration lab-adapted populations to the same proportion as in all heat stress adapted populations. Only the no-migration, lab-adapted individuals had a significantly lower proportion of fecund females (Figure 3.5 and S3.4).

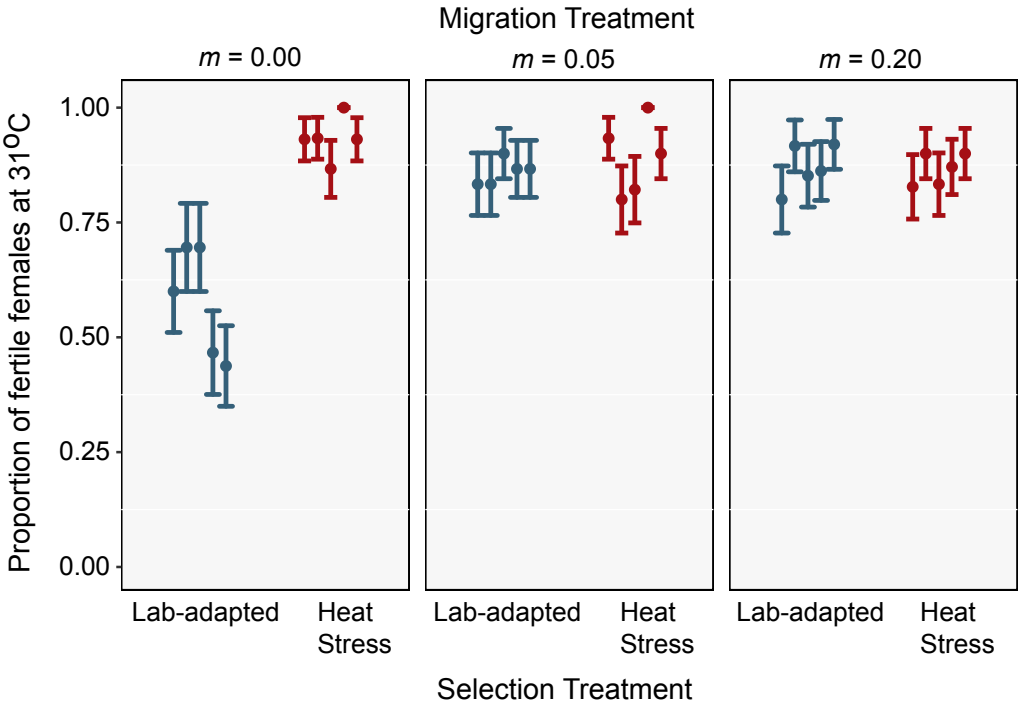


Figure 3.5: Proportion of female worms that are fecund or sterile at 31°C at generation 40. Females were counted as fecund if they at least two offspring over their lifetime or as sterile if they had zero or one offspring. Only the no migration ($m = 0.00$) lab-adapted populations had a significantly lower proportion of fecund females worms.

To investigate possible costs of adaptation, all generation 40 evolved populations were also assayed in the ancestral lab environment of 20°C (Figure 3.6). While three populations had lower fecundity at 20°C as compared to the ancestor population and all other replicates, there was no significant effect of selection ($F_{7,823} = 0.005$, $p = 0.9999$) or migration ($F_{7,823} = 1.06$, $p = 0.3876$), nor of their interaction ($F_{7,823} = 0.11$, $p = 0.9974$). These results indicate that there was no cost of adaptation to the chronic heat stress environment, as least as measured by female fecundity.

Response to selection for chronic heat stress resistance and lab-adaptation

In keeping with the rapid phenotypic response, the genomic response to selection was strong and widespread throughout the genome. The ancestor population had a high

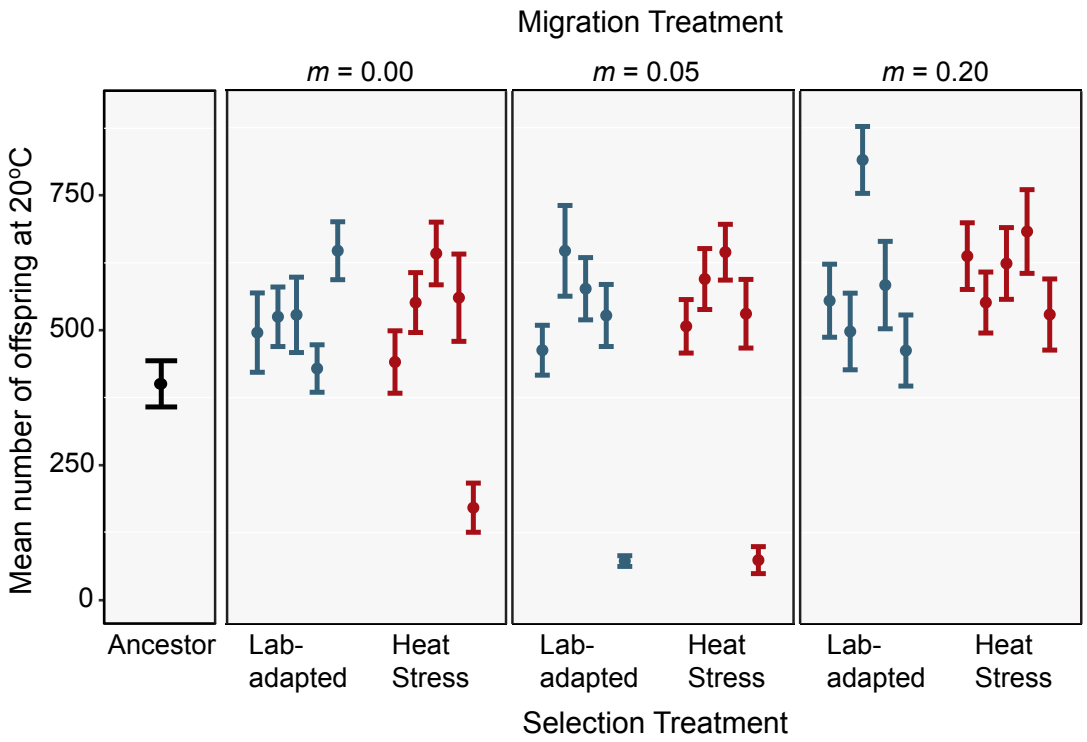


Figure 3.6: Female fecundity at 20°C for all evolved populations at generation 40. Mean plus and minus standard error is shown for all replicates and the ancestor population shown. There was no significant effect of selection or migration on female fecundity at 20°C.

level of nucleotide diversity, and patterns of nucleotide diversity did not change during the course of the experiment (Figure 3.7, Table S3.1). We first investigated the overall response to selection using the Cochran-Mantel-Haenszel test. We found that even after accounting for replicate, tens of thousands of sites responded to selection (Figures 3.8, S3.5 and Table 3.1). The number of significantly divergent sites in the heat stress adapted populations decreased as migration rate increased, and the number of significantly divergent sites in the lab adapted populations increased as migration rate increased (Figure 3.8 and S3.5, Table 3.1).

To see if this overall effect of the migration–selection interaction was seen in each replicate separately, we also tested for significant changes in allele frequency between the ancestor population and each evolved population using Fisher's exact test. We saw clear effects of migration of non-locally adapted individuals in our results; migration of non-adapted individuals greatly reduced the number of divergent loci in particular in the heat stress adapted populations experiencing high migration of non-locally adapted individuals (Figures 3.8 and 3.9). The number of significant divergent sites decreased with increasing migration rate, and this trend is particularly pronounced in the high migration heat stress evolved populations (Figures 3.8 and 3.9).

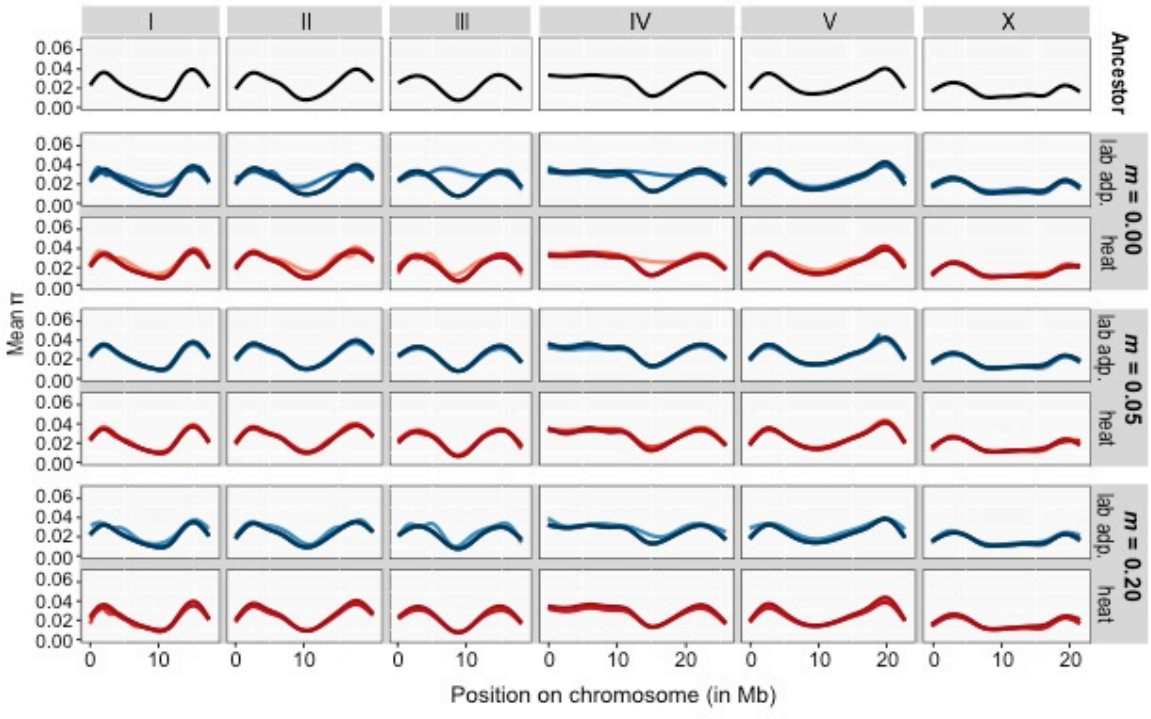


Figure 3.7: Mean nucleotide diversity (π) for ancestor and all evolved populations. Mean π smoothed over 10 kb windows shown. Different colors within a row are for each replicate. Nucleotide diversity does not change with experimental evolution.

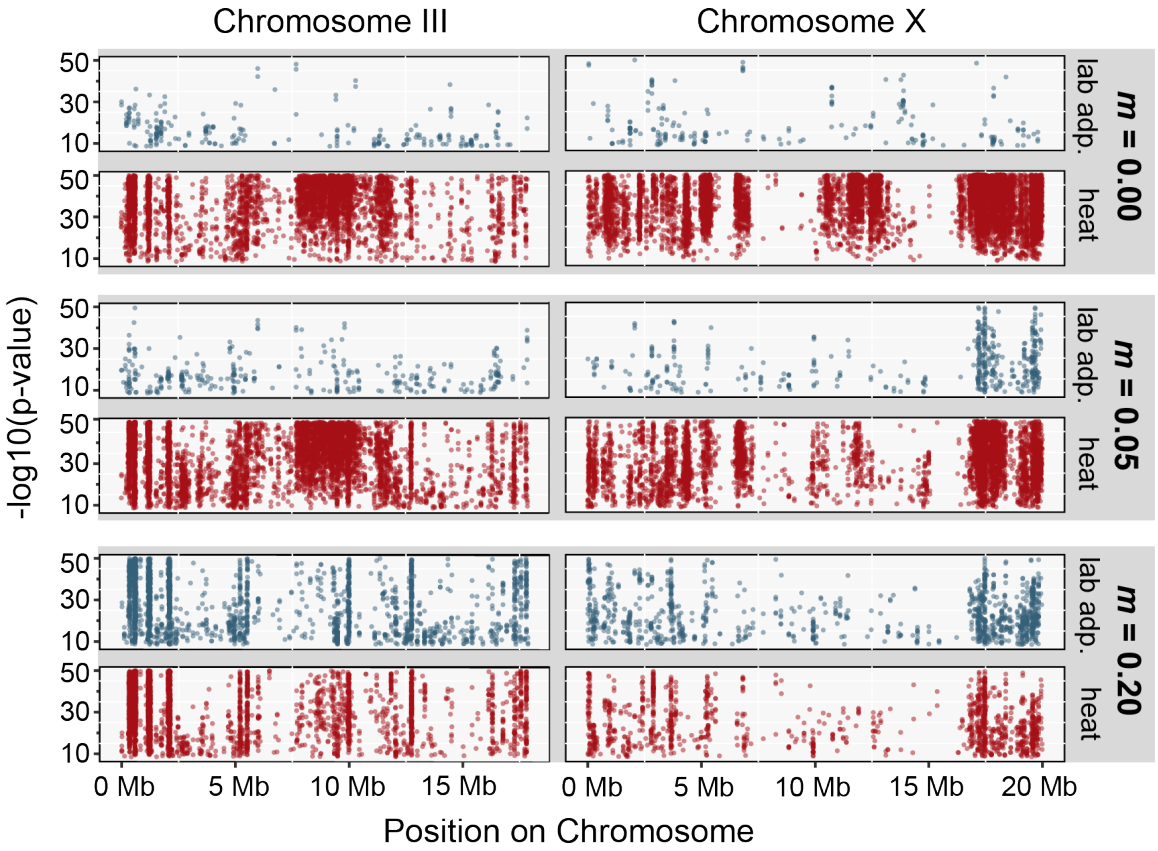


Figure 3.8: Number of divergent SNPs, compared to the ancestor population, for each migration-selection treatment combination. Cochran–Mantel–Haenszel test used to identify consistent changes in allele frequency across replicates. Panels separated by migration rate. Number of SNPs identified in the heat stress adapted populations decreases with increasing migration rate.

Migration Treatment	Selection Treatment	# divergent SNPs
m = 0.00	lab-adapted	4,603
	heat stress	95,479
m = 0.05	lab-adapted	5,722
	heat stress	84,693
m = 0.20	lab-adapted	43,667
	heat stress	54,922

Table 3.1: Number of divergent sites identified via the Cochran–Mantel–Haenszel test. Number of divergent sites in the heat stress adapted populations decreases with increasing migration rate.

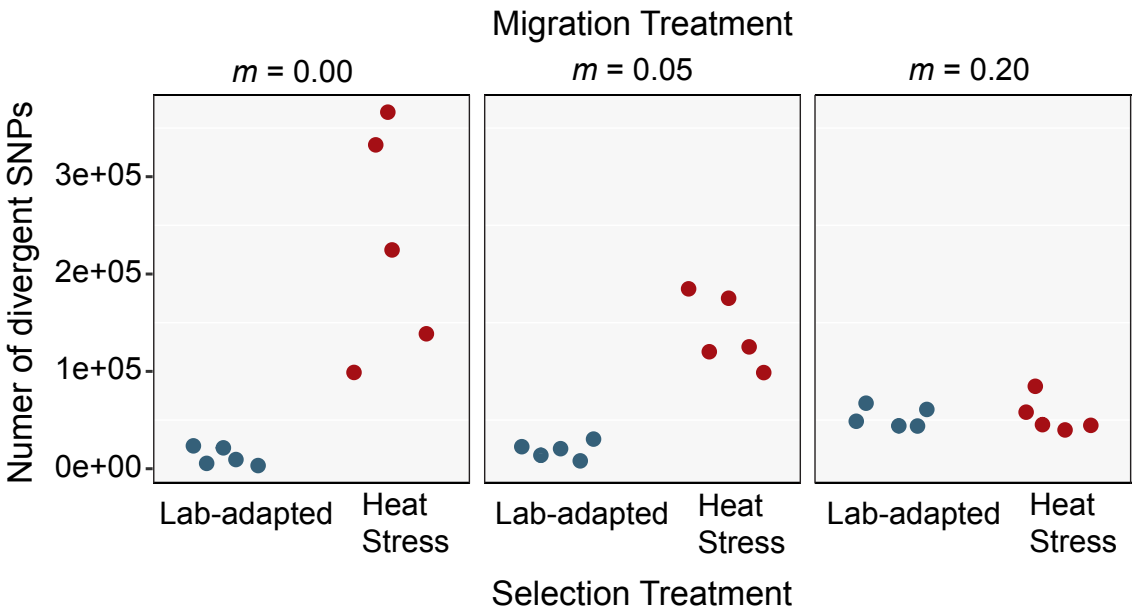


Figure 3.9: Number of divergent SNPs for each evolved replicate identified via Fisher's exact test. Number of divergent SNPs heat stress adapted populations decreases with migration of non-locally adapted individuals. Variation in total number of SNPs within a migration-selection treatment also decreases with migration.

We also found that many sites were significantly divergent from the ancestor population in the no-migration lab-adapted populations (Figures 3.8 and 3.9). These divergent loci are likely divergent due to a combination of random effects, such as genetic drift, and response to off-target selection, i.e. adaptation to the lab. In addition, these patterns of genomic change are similar to the ones seen in the control populations described in Chapter II, indicating that 75 generations of adaptation to the lab environment is not long enough to achieve mutation–selection balance.

Migration affects repeatability of evolution and population structure

In addition to gene flow affecting the number of divergent loci we also expected that gene flow would affect the locations of divergent regions. If there was successful gene flow, and if there was no cost of adaptation to the chronic heat stress environment, we would expect to see similar signals of divergence within the five replicate populations

that experienced the same migration–selection treatment. To test this we looked at overlap between divergent 1000 bp sliding windows. The proportion of divergent windows in common within a migration–selection treatment increased significantly with migration (Figure 3.10 and S3.6). Migration also affected the amount of population structure between replicate population pairs (i.e. the populations that exchanged individuals during the experiment in the low and high migration treatment populations). Genome wide mean F_{ST} decreased significantly with migration ($F_{(2,12)} = 8.39$, $p = 0.0053$) (Figures 3.11, 3.12 and S3.7).

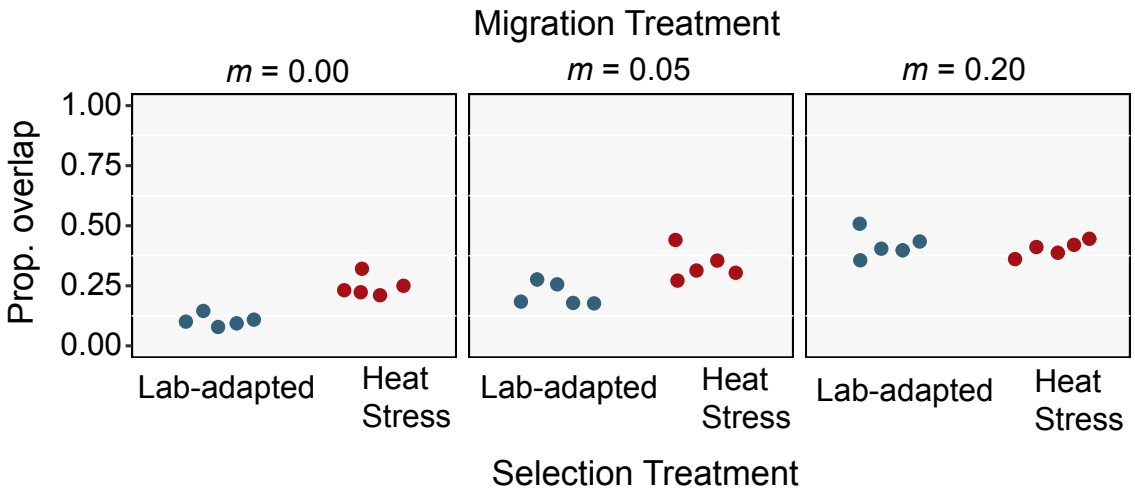


Figure 3.10: Proportion of significantly divergent 1000 bp windows for each evolved population that overlap with all windows in common within a migration–selection treatment combination. Proportion of windows that overlap within a migration-selection treatment increases with migration and in the heat stress adapted populations.

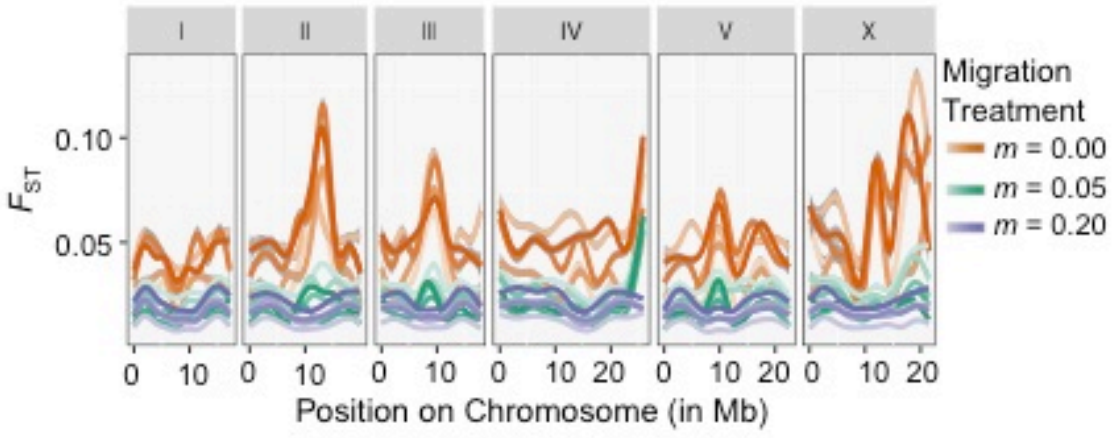


Figure 3.11: Mean F_{ST} between evolved replicate pairs over the whole genome. Mean F_{ST} smoothed over 10 kb non-overlapping sliding windows. F_{ST} between evolved replicate pairs is increased and more variable in the absence of migration. A replicate pair refers to the lab-adapted and heat stress adapted populations that reciprocally exchanged individuals in the low- and high- migration treatments.

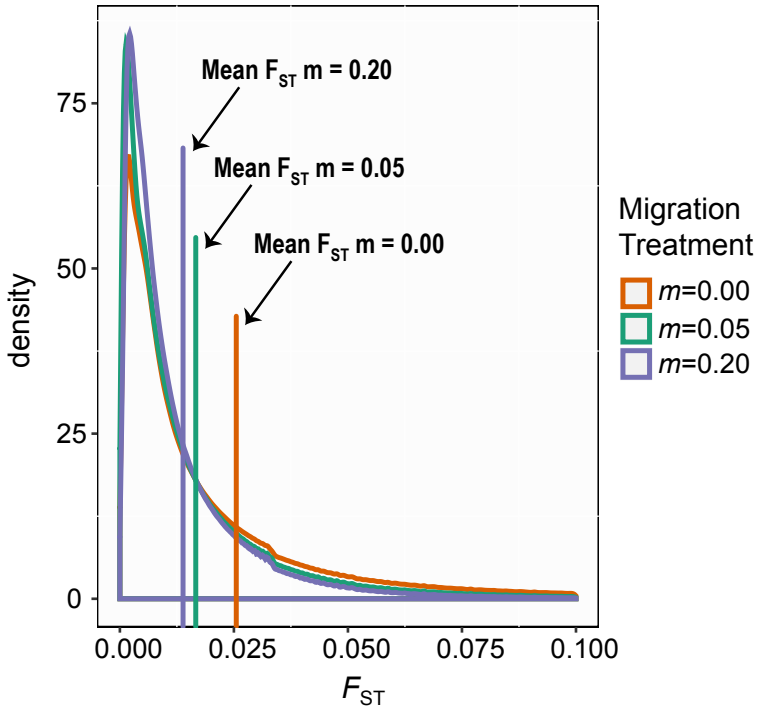


Figure 3.12: Distribution of F_{ST} values between replicate pairs for each migration treatment shown. F_{ST} calculated on a single SNP basis between the lab-adapted and heat stress selected evolved populations for each replicate pair. Mean of genomic F_{ST} means identified. Mean F_{ST} increases significantly in the absence of migration.

Effective population size

Effective population size (N_e) was greatly reduced compared to census population size in all populations (Figure 3.13). This follows a pattern seen in previous experimental evolution projects using this strain of *C. remanei* (Chapter II). *C. remanei* has an XX/X0 sex determination system (Thomas, Woodruff, and Haag 2012; Hodgkin 1987) so we would expect that the estimated effective population size using only SNPs found on the X chromosome would be around 75% of the estimated effective population size using SNPs autosomes. However, we instead found that the estimated N_e on the X chromosome ranged from 55% to 122% of the estimated N_e from just the autosomes (Figure 3.13).

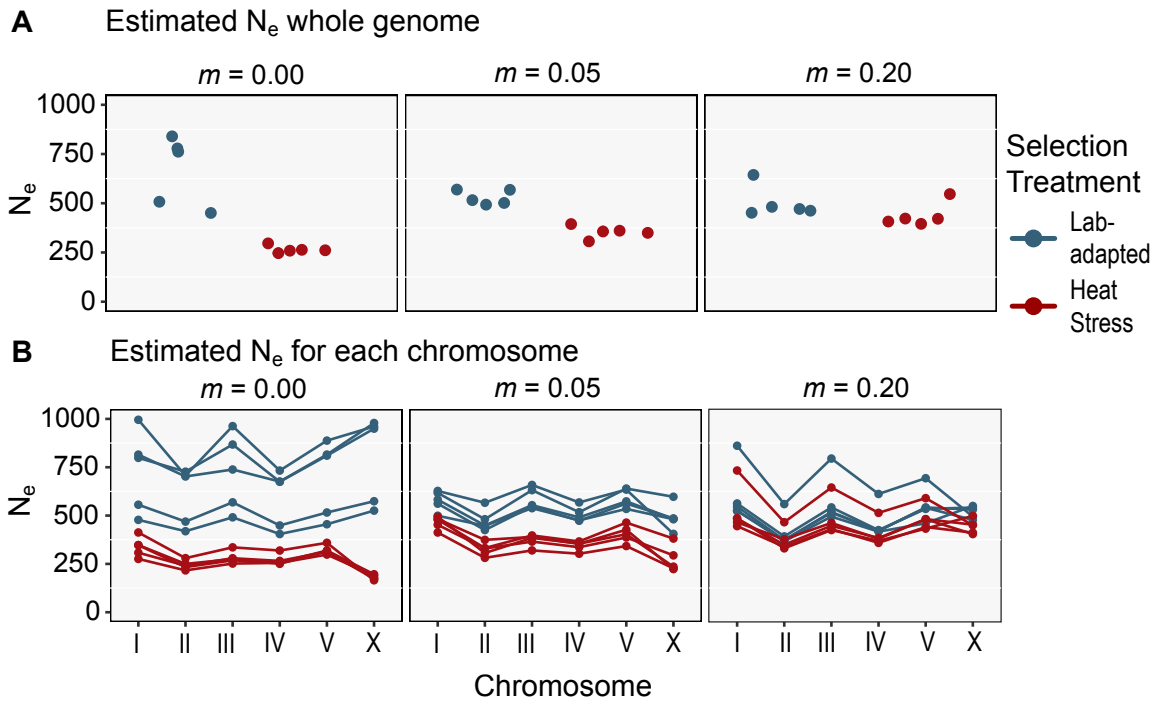


Figure 3.13: Estimated effective population size (N_e) for all evolved populations. (A) Estimated N_e for the whole genome. (B) Estimated N_e for each chromosome separately. N_e is greatly reduced compared to census population sizes.

DISCUSSION

Selection and migration interact to affect phenotypic and genomic divergence

The study of divergence and speciation has traditionally focused on how genetic barriers to gene flow, that is reproductive isolation, evolve in the context of ecological or environmental pressures and important genes that cause hybrid incompatibility (Nosil and Feder 2012; Coyne and Orr 2004). Additionally, studies have tended to focus on the roles of one or two genes of major effect without consideration of the possible influences of genomic context. However, as whole-genome sequencing has become more readily available, researchers have begun to incorporate next generation sequencing to study not only one or two genes thought to be important for the evolution of reproductive isolation between two populations but also patterns of divergence across the whole genome. This effort has driven the realization that, while there was a good understanding of specific genes of major effect thought to be important for speciation, we still have a poor understanding of how those genes fit into the genome overall (Nosil and Feder 2012).

In particular, there has been a focus on the potential interaction between "speciation genes" and "genomic islands of divergence:" locations in the genome that show elevated genetic divergence between two populations compared to expectations under neutrality (Nosil, Funk, and Ortiz-Barrientos 2009; Nosil and Feder 2012). These islands are predicted to be regions of the genome that contain genes important for the process of divergence between two populations and that are thus under strong selection and resistant to introgression even in the face of continued gene flow (Feder, Egan, and Nosil 2012; Nosil and Feder 2012; Turner, Hahn, and Nuzhdin 2005; Turner and Hahn 2007; Nosil, Egan, and Funk 2008). Genomic islands of divergence have been

hypothesized to be important for catalyzing the process of reproductive isolation with gene flow: as two populations diverge, regions in the genome with a reduced effective migration rate form because they contain a gene or genes important for local adaptation. Hitchhiking and linkage cause these genomic islands of divergence to grow in size, becoming 'continents' of divergence that eventually lead to reproductive isolation among populations (Feder et al. 2014; Feder et al. 2013; Via 2012).

This theory is controversial, however. A 2014 reanalysis of published genomic data concluded that genetic diversity was reduced in previously identified genomic islands of divergence, instead of absolute measures of divergence being higher, as predicted by the islands of divergence hypothesis (Cruickshank and Hahn 2014). Reductions in nucleotide diversity can bias relative measures of divergence such as F_{ST}, so if putative genomic islands of divergence are only found in regions of reduced genetic diversity that could mean that such regions do not contain genes important for divergence, and are only sites of reduced genetic diversity, not unexpectedly high divergence (Noor and Bennett 2009; Cruickshank and Hahn 2014). Additionally, negative correlations between gene flow and adaptive divergence do not imply causation one way or the other (Räsänen and Hendry 2008).

Indeed, a number of recent studies have found that heterogeneous genomic divergence correlates more closely with genomic architecture, in particular regions of low recombination or low genetic diversity, as opposed to patterns gene flow. A recent study of benthic and limnetic and marine and freshwater stickleback found that divergent regions were concentrated in regions of low recombination (Samuk et al. 2017). A 2012 study of divergent between two naturally hybridizing flycatcher species found a highly

heterogeneous signal of divergence, and originally attributed that heterogeneity to genomic islands of divergence (Ellegren et al. 2012). However, a 2015 reanalysis of the same species plus two more closely related flycatchers found that islands of divergence did not correlate with patterns of gene flow but instead with variation in recombination rate (Burri et al. 2015). This result matches a study done of 12 species pairs of Darwin's finches (Han et al. 2017), in which the authors found that the number of 'genomic islands' identified was similar for both allopatric and sympatric species pairs.

The above issues are difficult to disentangle using historical data and inference, although attempts have been made (Nosil and Crespi 2004). The experiment described here allowed us to explicitly test theory concerning patterns of genomic divergence between two populations experiencing the opposing effects of selection and migration. Specifically we can test theory concerning patterns of genomic divergence between two populations in primary contact, such as what has likely occurred during the formation hawthorn- and apple-infesting host races of the apple maggot fly *Rhagoletis pomonella* (Michel et al. 2010), and the role that divergent selection plays in divergence with gene flow (Feder et al. 2013). Both the phenotypic and genetic results clearly show that there was successful gene flow between populations and that that gene flow greatly affected genomic divergence in all populations experiencing migration of non-locally adapted individuals. We have shown that increasing the rate of migration of non-locally adapted individuals into a novel environment drastically alters patterns of genomic divergence, not only patterns of phenotypic divergence (Figures 3.3, 3.4, 3.8 and 3.9). In addition to changes in the number of divergent loci, we also saw a decrease in the level of population

structure between replicate population pairs with increasing migration rate and an increase in the reproducibility of the response to selection (Figures 3.10–3.12).

However, we do not find strong evidence that genomic islands divergence formed. While we see large divergent genomic regions in the no migration heat stress evolved populations, those regions break down with migration of lab-adapted individuals (Figures 3.8 and S3.5). This breakdown is mixed in the low migration heat stress evolved populations, where the number of divergent SNPs identified in each replicate is less than the number of SNPs identified in some of the no migration heat stress evolved populations, but ranges overlap (Figure 3.9). In addition, large regions of divergence in the no-migration and low-migration heat stress evolved populations often overlap with regions of low nucleotide diversity (Figures 3.7 and 3.8). Low nucleotide diversity can upwardly bias relative measures of population divergence (B. Charlesworth, Charlesworth, and Nordborg 1997; Noor and Bennett 2009). As such, some of these large divergent regions seen in the no-migration and low-migration heat stress evolved populations may be due to low recombination.

The divergent SNPs identified in Figures 3.8 and S3.5 measure divergence in respect to the ancestor population, so those patterns of divergence do not give us information about divergence between evolved replicate pairs. For that, we need to look at F_{ST} between evolved replicate pairs. We find that genome wide mean F_{ST} between evolved replicate pairs is significantly decreased with the addition of migration, and a spike in F_{ST} at one end of Chromosome IV overlaps with a region of reduced nucleotide diversity (Figures 3.7 and 3.11). Thus, in spite of strong selection, we do not see extended regions of divergence between low- and high-migration evolved replicate pairs,

compared to regions of divergence between no-migration replicate pairs. This is in contrast to patterns of genomic divergence seen in putative cases of genomic islands of divergence (Michel et al. 2010).

Migration affects patterns of phenotypic divergence

The increase in female fecundity in the lab-adapted populations experiencing migration of heat stress adapted individuals demonstrates that there was successful gene flow, not just migration, between populations. Of the possible scenarios illustrated in Figure 3.1B, out results most closely reflect the scenario in which positively selected alleles in the heat stress environment swamp out alleles present in the lab-adapted environment and there is no or little cost to maintaining those alleles, so even individuals who have themselves never experienced the heat stress environment have increased fecundity compared to the no migration lab-adapted individuals. Unexpectedly, there were no apparent trade-offs between adaptation to the heat stress and lab environments: the migration treatment did not dampen the response to selection at 31°C by the end of the experiment and there was no effect of migration or selection on fecundity in the ancestral environment of 20°C (Figures 3.4 and 3.6). Another interesting phenotypic result was the rapid increase in female fecundity seen at generation 20 in the low migration heat stress adapted populations and the subsequent decline seen at generation 30. This could have been due to a rapid sweep in adaptive alleles of smaller effect that were subsequently lost to further migration of the lab-adapted populations.

Migration alters the genomic signature of selection

Continuous migration of non-locally adapted individuals affected the signature of divergence and adaptation seen at the genomic level. Specifically, migration reduced the

number of divergent sites identified (Figures 3.8 and 3.9), as predicted by models and simulations of the effects of gene flow on divergence (Haldane 1930; Yeaman and Whitlock 2011). Migration affected the total number of divergent sites, the overlap in divergent sites between evolved replicates and amount of population structure between evolved replicates (Figured 3.8–3.12). The large degree of overlap between divergent regions in the lab-adapted and heat stress adapted population pairs shows that not only was there successful gene flow between replicate population pairs, but that heat stress adaptive alleles were not disadvantaged in the ancestral, standard laboratory environment (Figure 3.10).

Chronic heat stress is complex and polygenic

The response to chronic heat stress was widespread throughout the genome in all heat stress selected populations, regardless of selection treatment, similar to what was seen for selection for the acute heat stress selected population described in Chapter II. We originally predicted that the signal of chronic heat stress response would be weaker than the response to acute heat stress described in Chapter II, given the different effects on survival (Figure 2.1, Figure S3.2). However, we found the opposite. Through all forms of genomic analysis, we found tens of thousands or hundreds of thousands of sites that responded to selection. This result, coupled with the very small estimated effective population sizes compared to census population sizes (Figure 3.13) indicates very strong selection.

Migration of non-locally adapted individuals does dampen response to selection at the genomic level, but the response is still very complex (Figures 3.8 – 3.10, S3.5 and S3.6). Given the convergence in the number of divergent SNPs identified in the high

migration lab-adapted and heat stress evolved populations (Figures 3.8 and 3.9, Table 3.1) and the strong phenotypic response seen in the lab-adapted high migration populations (Figure 3.4), we might find the clearest signal of heat stress adaptation by identifying overlap in divergent sites between all heat stress evolved populations and the high migration lab-adapted populations. Additionally, if this experiment were to be continued for hundreds or thousands of generations, we might then find a simple signal of heat stress adaptation, similar to how a few thousand years of migration and selection eroded most genetic divergence between deer mice on and off Sand Hills in Nebraska, except at loci important for coat coloration (Pfeifer et al. 2018).

This highly polygenic response is not without precedent. Studies of selected traits in plant and animal domestication have found that many phenotypes under selection during domestication have are complex and polygenic (Pettersson et al. 2013; Sheng et al. 2015; Kemper et al. 2014; Hirsch et al. 2014). A highly polygenic response to artificial selection has been seen in other experimental evolution projects, often after only a few generations of selection. Orozco-ter Wengel et al. (2012) identified 5000 divergent SNPs after only 15 generations of selection to a high temperature, laboratory environment and Behrman et al. (2018) identified functional immune response differences and variable SNPs between populations of *Drosophilia melanogaster* collected during different seasons.

We found that N_e estimates are more than an order of magnitude smaller that the census population sizes, similar to what we found in the experimental evolution project described in Chapter II (Tables 2.1 and 2.2, Figures 3.2 and 3.13). Unexpectedly, migration did not increase effective population size above the effective population sizes

estimated for the lab-adapted no migration populations. The effective population size estimates were also very consistent between replicates. As was described previously, variance in mating success and fluctuating population size will reduce N_e (Kimura 1983) and here strong selection may have also reduced genetic variation at linked sites via the Hill-Robertson effect (Hill and Robertson 1966; B. Charlesworth 1996; B. Charlesworth 2012; Comeron, Williford, and Kliman 2007).

CONCLUSION

Gene flow and selection are powerful and opposing forces and gene flow of non-locally adapted individuals is expected to affect patterns and signatures of both phenotypic and genetic divergence. Here, we have shown how reciprocal migration of nematode worms between two very different environments affected both phenotypic and genomic patterns of divergence. Gene flow of non-locally adapted individuals did not dampen the phenotypic response by the end of the experiment, but did reduce the number of divergent SNPs identified and significantly reduced population structure between replicate population pairs.

With the strong selection imposed on the heat stress adapted populations, we would have predicted to see genomic islands of divergence form in the heat stress adapted populations that experienced migration of non-locally adapted individuals, if proximity to selected alleles was sufficient to locally reduce the effective migration rate. However, we did not see any evidence for large genomic islands of divergence in either of the heat stress evolved populations that experienced migration of non-locally adapted individuals. Our results indicate that any genomic islands that may form from standing

genetic variation will be easily swamped out migration, even in the presence of strong selection.

CHAPTER IV

CONCLUSION

Organisms are complex, integrated systems and the manner in which they respond to selection reflects that. While evolutionary quantitative genetics has provided a powerful statistical framework for analyzing complex traits, even in the absence of important pieces of information such as the number of genes that underlie variation in a trait or the distribution of their effects (Barton and Turelli 1989), taking the next step to identify the actual genes that underlie this variation has proved difficult (Rockman 2012). Here, I have shown that by integrating phenotyping, whole genome sequencing and an evolve and re-sequence approach, I can dissect the basis of abiotic stresses, as model complex traits, in the nematode *Caenorhabditis remanei* and investigate pressing questions in evolutionary genetics.

In Chapter II, I demonstrated how acute heat stress and acute oxidative stress response are independent traits, contrary to what is predicted by the molecular biology literature (Rodriguez et al. 2013). Although some level of pleiotropy throughout the genome may be ubiquitous (Paaby and Rockman 2013; Wagner and Zhang 2011), highly pleiotropic genes are not likely the most common responders to selective pressures. Additionally, alleles, not genes, are the appropriate unit of variation, as different alleles can have different pleiotropic effects (Wagner and Zhang 2011; Phillips and McGuigan 2006; Wright 1968; Lande 1984). As such, seeking to understand the genetic architecture of traits solely using gene knockouts is likely to leave an incomplete picture of the overall

structure of the pleiotropic network, especially in respect to segregating variation within natural populations.

Indeed, contrary to predictions, I have shown that acute heat stress and acute oxidative stress response were independent at all levels investigated – survival phenotype, constitutive changes in gene expression and locations of significant changes in allele frequency across the genome. These results demonstrate that the space of possible interactions, i.e. the number of pleiotropic interactions a gene has as shown by a knock out study, is not necessarily the same thing as the space of realized interactions, i.e. the number of pleiotropic interactions two allelic variations of a gene have.

Gene flow and selection are powerful and opposing forces, and gene flow of non-locally adapted individuals is expected to affect patterns and signatures of both phenotypic and genetic divergence. In Chapter III I showed how reciprocal migration of nematode worms between two very different environments affected both phenotypic and genomic patterns of divergence. While gene flow did not dampen adaptation to the novel heat stress environment in the populations that experienced migration of non-locally adapted individuals, gene flow did reduce the number of divergent SNPs identified and significantly reduced population structure between replicate population pairs, as shown by the significant reduction in mean F_{ST} between evolved replicate population pairs with the addition of migration.

The controlled experimental design of the project described in Chapter III also allowed me to test hypotheses about the formation of genomic islands of divergence (Nosil and Feder 2012). Hypothesis and theory concerning the formation of genomic islands of divergence and genetic hitchhiking have been formed using examples from

natural populations (Feder et al. 2003; Nosil, Egan, and Funk 2008; Larson et al. 2016). While such natural experiments are the inspiration for novel theory, and understanding the world around us is the ultimate inspiration and goal of biology, they can be lacking when it comes to rigorous testing of hypotheses. In particular, important parameters including population history, migration rate, population size and strength of selection must be estimated. In contrast, with experimental evolution I was able to control those parameters.

With strong selection and knowledge that gene flow was occurring, we did not see evidence for large genomic islands of divergence in the high migration heat stress evolved population and mixed evidence for genomic islands of divergence in the low migration heat stress evolved populations. Our results are more concordant with the theory laid out by Haldane (1930) that as migration increases, the number of divergent sites will decrease, as they will be swamped out by migration. The results described here indicate that any genomic islands that may form will be swamped out migration, even in the presence of strong selection, and that genome architecture is an important factor determining the genetic architecture of adaptation, even in the face of gene flow, similar to what was seen in flycatchers (Burri et al. 2015). Here, strong selection was not sufficient to reduce migration at sites adjacent to those under selection.

The components of the genomic islands of divergence/genetic hitchhiking related hypotheses can be tested by the experimental set up described in Chapter III are limited. In particular, I was only able to test hypotheses related to the formation of islands of divergence during primary contact, and over a relatively short time on an evolutionary timescale (40 generations). There are a myriad of possible future directions with this

experimental system. The reciprocal migration treatments could be continued to see if over more generations of divergence and adaptation with gene flow a more precise genetic architecture of chronic heat stress adaptation emerges, as has been seen in the Nebraska Sand Hill deer mice populations (Pfeifer et al. 2018). I could also see if one-way migration (in this case only from the lab-adapted populations to the heat stress adapted populations) affects genomic patterns of divergence. Finally, some studies of divergence or speciation with gene flow investigate divergence with gene flow following a period of allopatry, such as what occurred in benthic and limnetic North American lake whitefish (Gagnaire et al. 2013). I could introduce migration of non-locally adapted individuals into the no-migration heat stress adapted populations to test if any of the large genomic regions of divergence are maintained in the face of gene flow, instead of asking if such regions of divergence can form in spite of gene flow, which is what was done in the experiment described in this dissertation. I would then be able to compare patterns of divergence and adaptation to the heat stress environment given histories of no gene flow, continuous gene flow and gene flow after a period of allopatry. Those population histories represent some scenarios for current examples of divergence with gene flow that can be difficult to distinguish between using only data from extant populations (Stankowski, Sobel, and Streisfeld 2015).

Nearly 100 years of quantitative genetics has demonstrated that dissecting the genetic basis of complex traits is, well, complex, and a multi-prong approach is needed (Wellenreuther and Hansson 2016). In Chapter II, I showed how single gene knockout studies can paint an incomplete picture of the genetic architecture of a model complex trait. In Chapter III, I showed how gene flow affects the genetic architecture of adaptation

and divergence. I was also able to experimentally test, in an animal system, hypotheses concerning the formation of genomic islands of divergence during the process of divergence with gene flow. For both projects described here, multiple analysis methods were necessary, and the project would not have worked without all methods.

Life is complex, and it is complex at all levels. Within an individual, a single gene can affect variation in more than one trait and tens or hundreds of genes may control variation in a single trait. The ability of a population, the unit of evolution, to respond to selection pressures is dependent on not only the strength of selection and the genetic architecture of the trait under selection, but also on the vagaries of historical processes: the total amount of genetic variation present in the population, population size and the rate at which novel genetic variation (be via mutation or recombination) is created. The work described in this dissertation demonstrates that in order to understand the genetic architecture of complex traits, all these levels of complexity must be taken into account.

APPENDIX A

SUPPORTING INFORAMTION FOR CHAPTER II

Population	Q1	Median	Mean	Q3	% bases > 15x
Ancestor	56	104	121	153	92.7%
Replicate 1 Control – 1	9	17	21	27	55%
Replicate 1 Control – 2	29	56	66	84	86%
Replicate 1 Heat stress	51	102	118	150	91.6%
Replicate 1 Oxidative stress	58	111	126	160	92.6%
Replicate 2 Control – 1	45	90	102	132	90.5%
Replicate 2 Control – 2	24	44	52	67	84.4
Replicate 2 Heat stress	45	84	98	123	91.0%
Replicate 2 Oxidative stress	56	102	111	140	92.5%

Table S2.1: Coverage data for all sequenced populations.

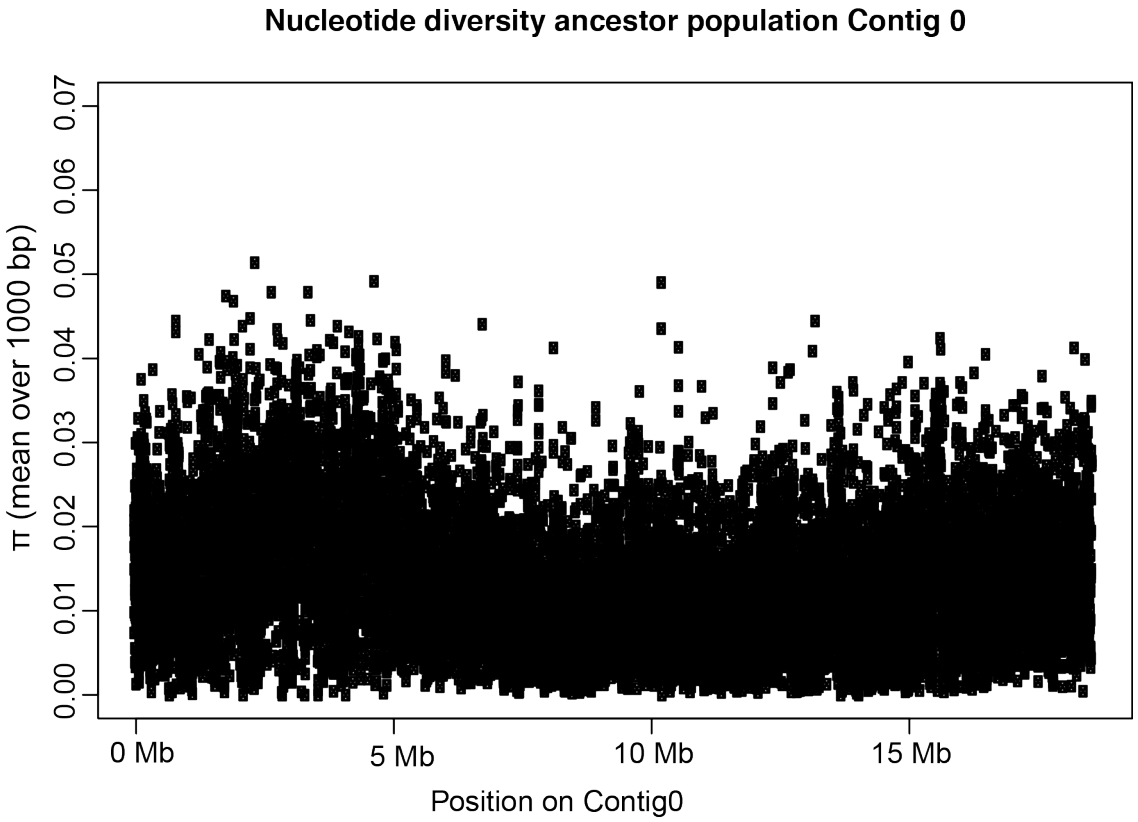


Figure S2.1: Mean nucleotide diversity (π) over 1000 bp sliding windows on Contig 0 (X Chromosome) in the ancestor population.

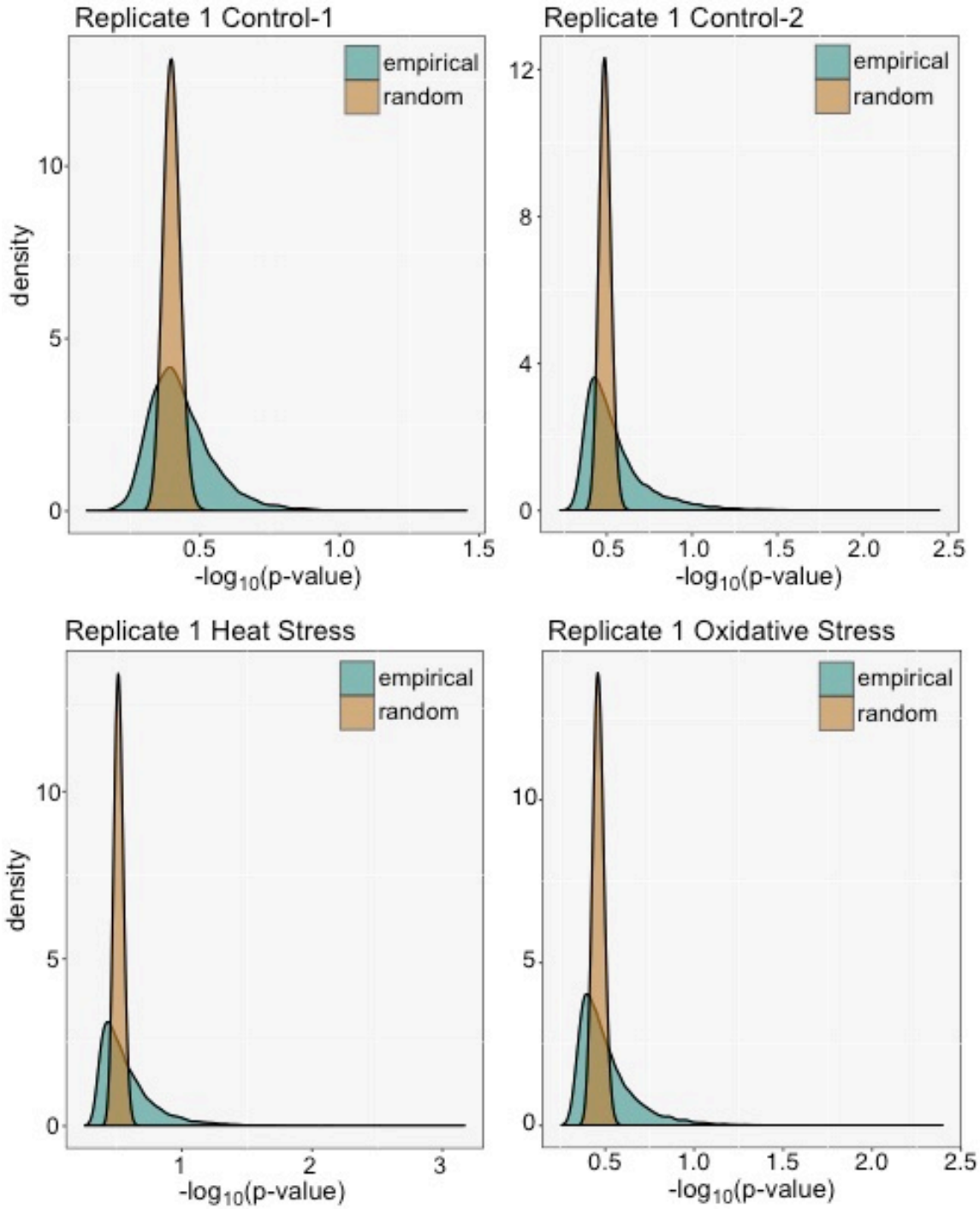


Figure S2.2: Distribution of random sliding window and gnome location based (empirical) sliding windows for all Replicate 1 populations. X-axis is the mean $-\log_{10}(p\text{-value})$ for each window.

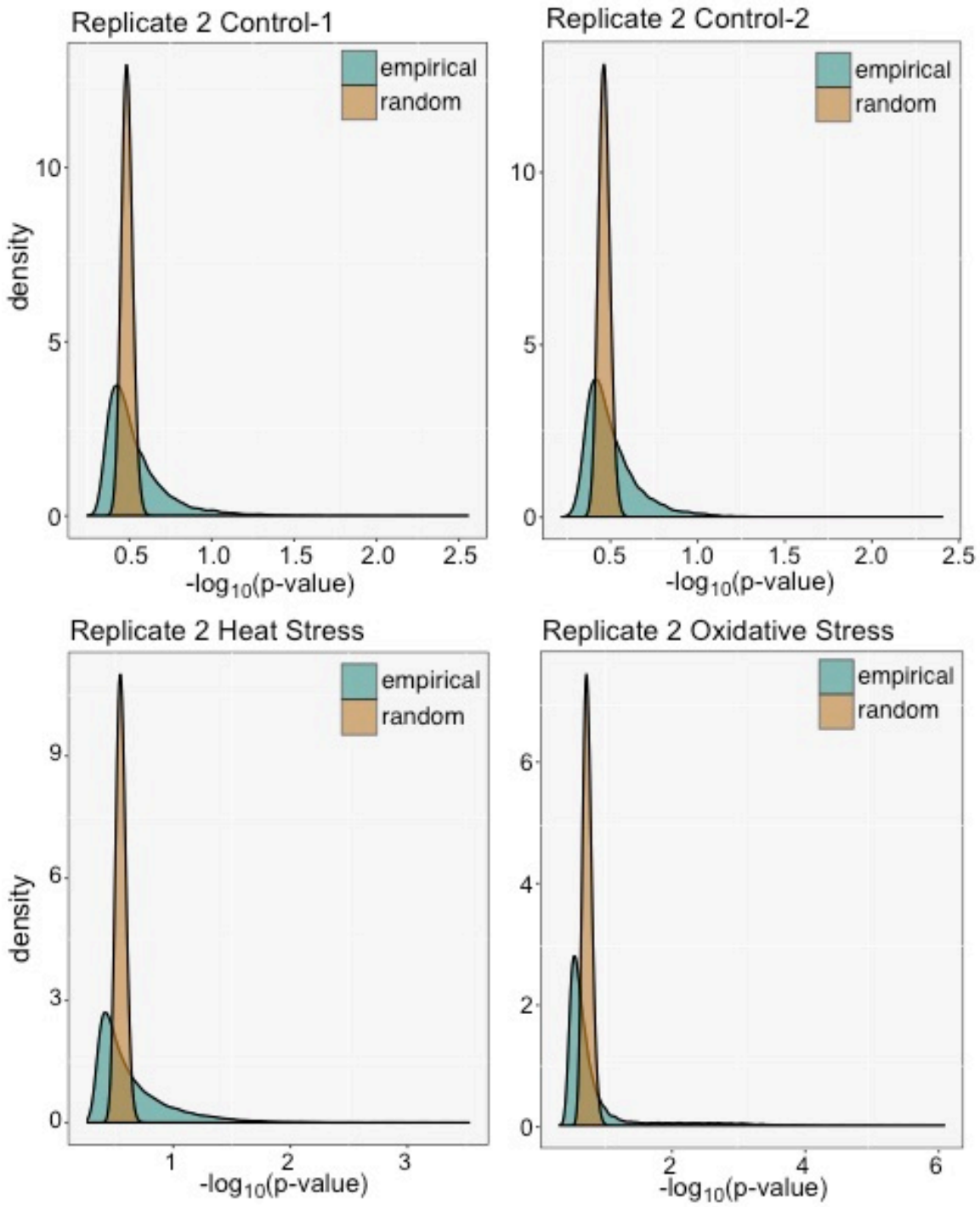


Figure S2.3: Distribution of random sliding window and gnome location based (empirical) sliding windows for all Replicate 2 populations. X-axis is the mean $-\log_{10}(p\text{-value})$ for each window.

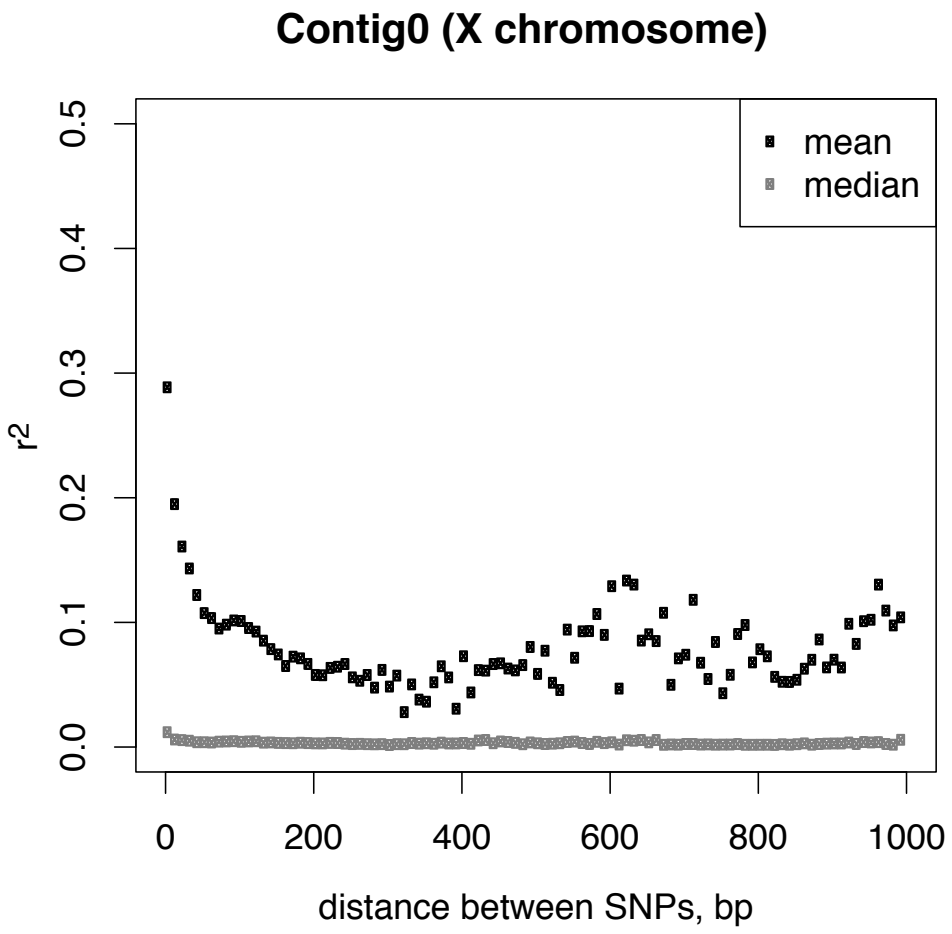
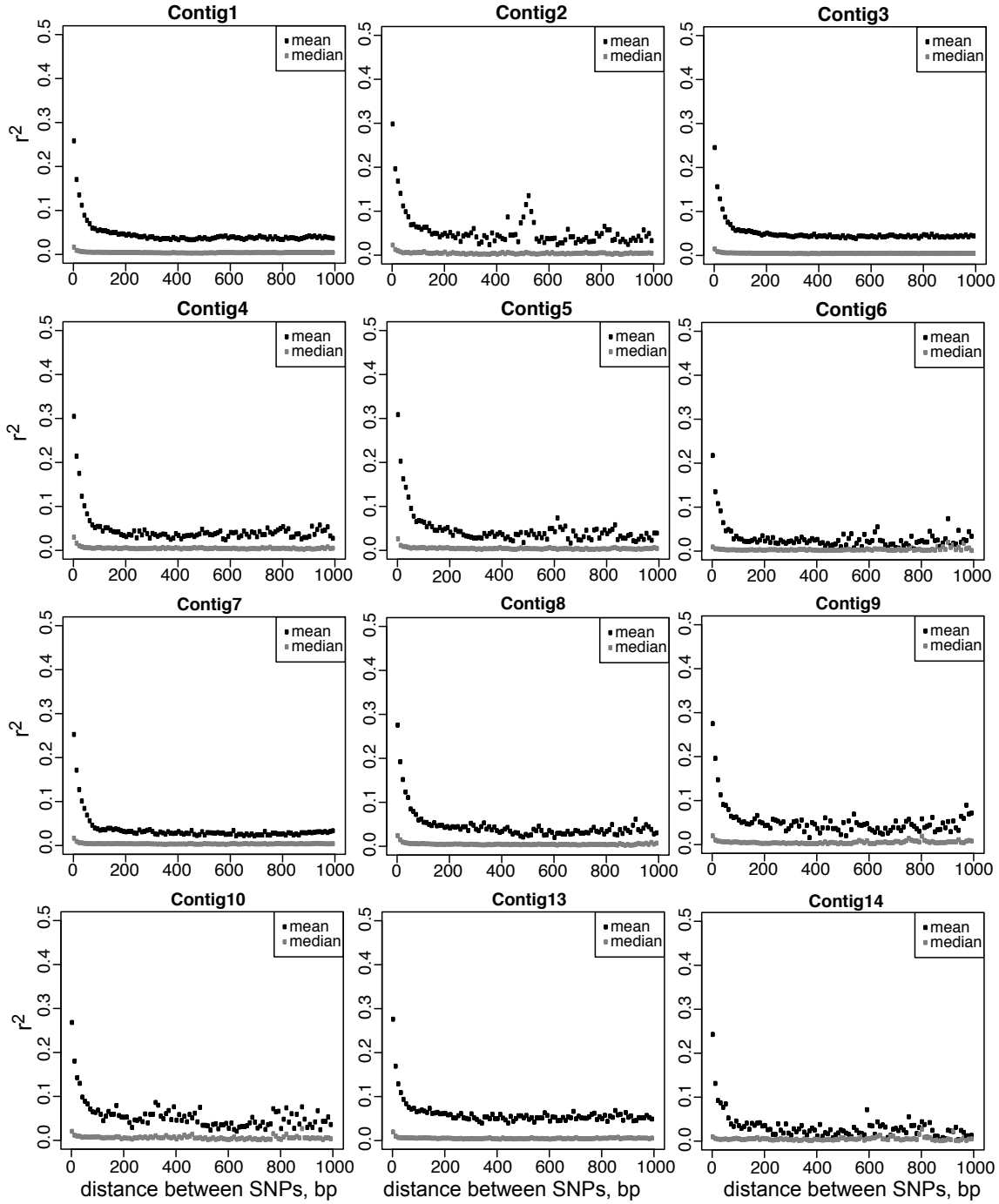


Figure S2.4: Mean and median linkage disequilibrium for all SNPs within 10 kb of a focal SNP on Contig 0 (X Chromosome). Only 1 kb shown for graphing purposes. Within 10 bp of a focal SNP, r2 averages only around 0.2 and decays to background levels of ~0.089 within 200 bp.



Figures S2.5: Linkage disequilibrium decay on the 12 largest contigs after Contig 0 (X Chromosome). See Figure S2.3 for full description.

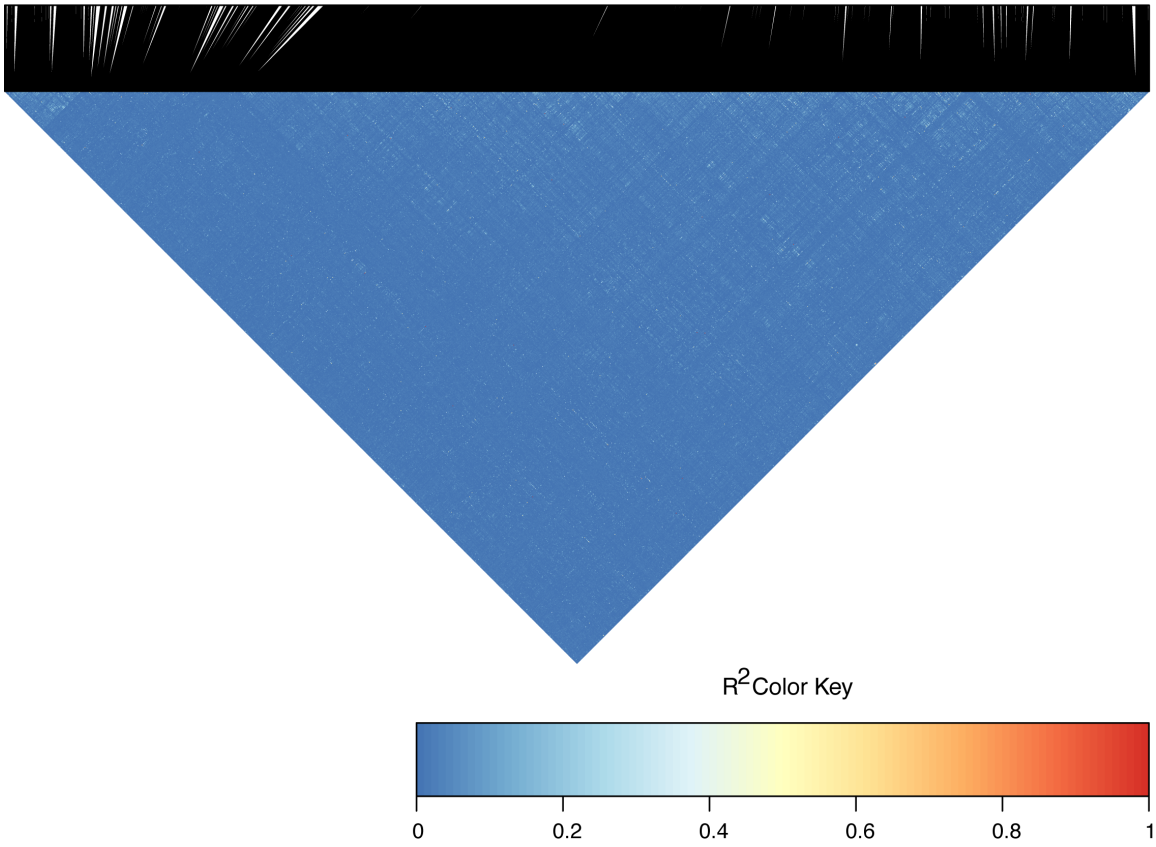


Figure S2.6: Pairwise linkage disequilibrium for 3880 high confidence SNPs on Contig 0 (X Chromosome). Sites fixed for non-reference alleles were removed. Bin size: 10 kb. Physical length: 18.5 Mb.

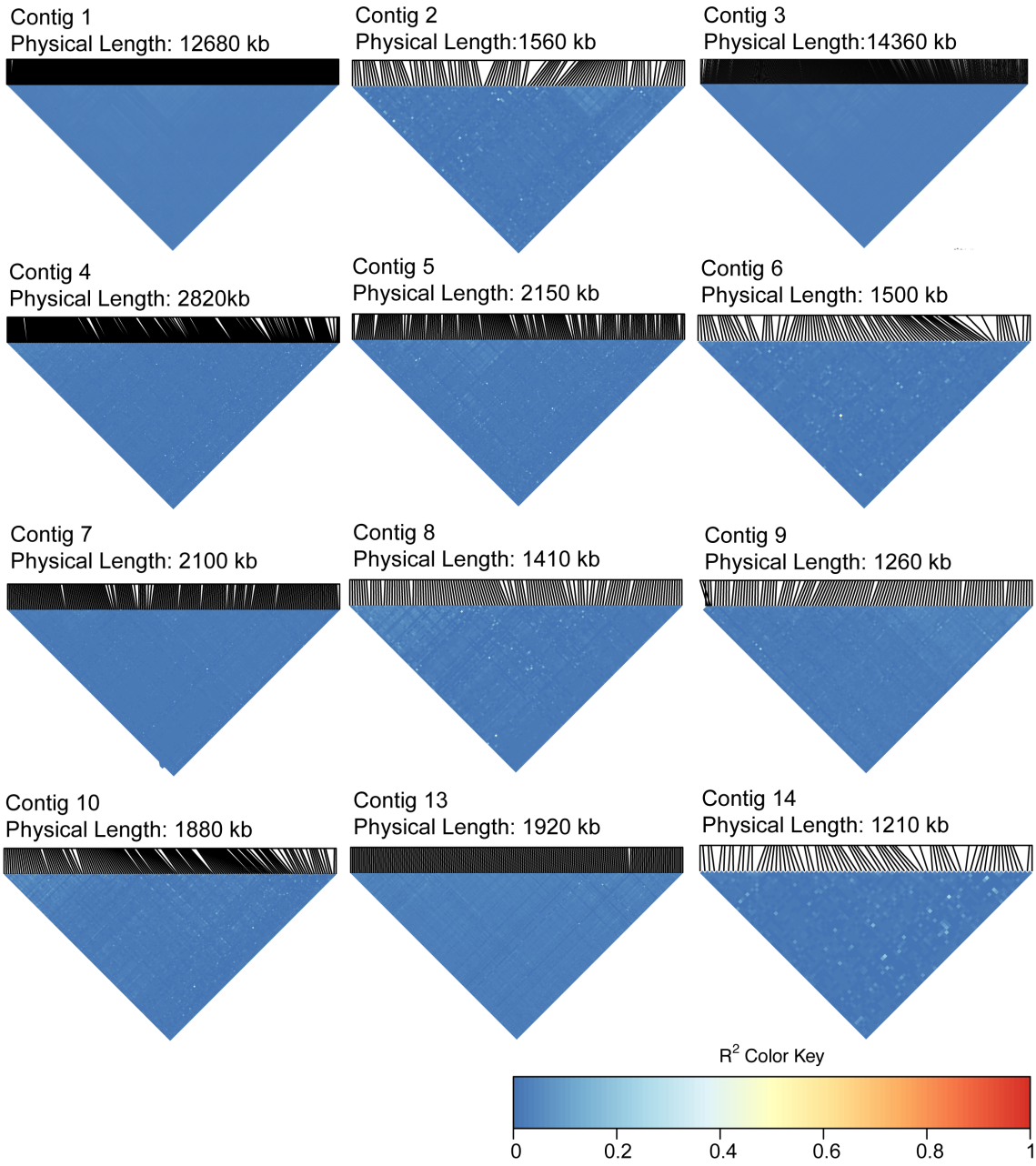


Figure S2.7: Pairwise linkage disequilibrium of 12 largest contigs after Contig 0. Results are binned into 10 kb windows. For full description see Figure S2.6.

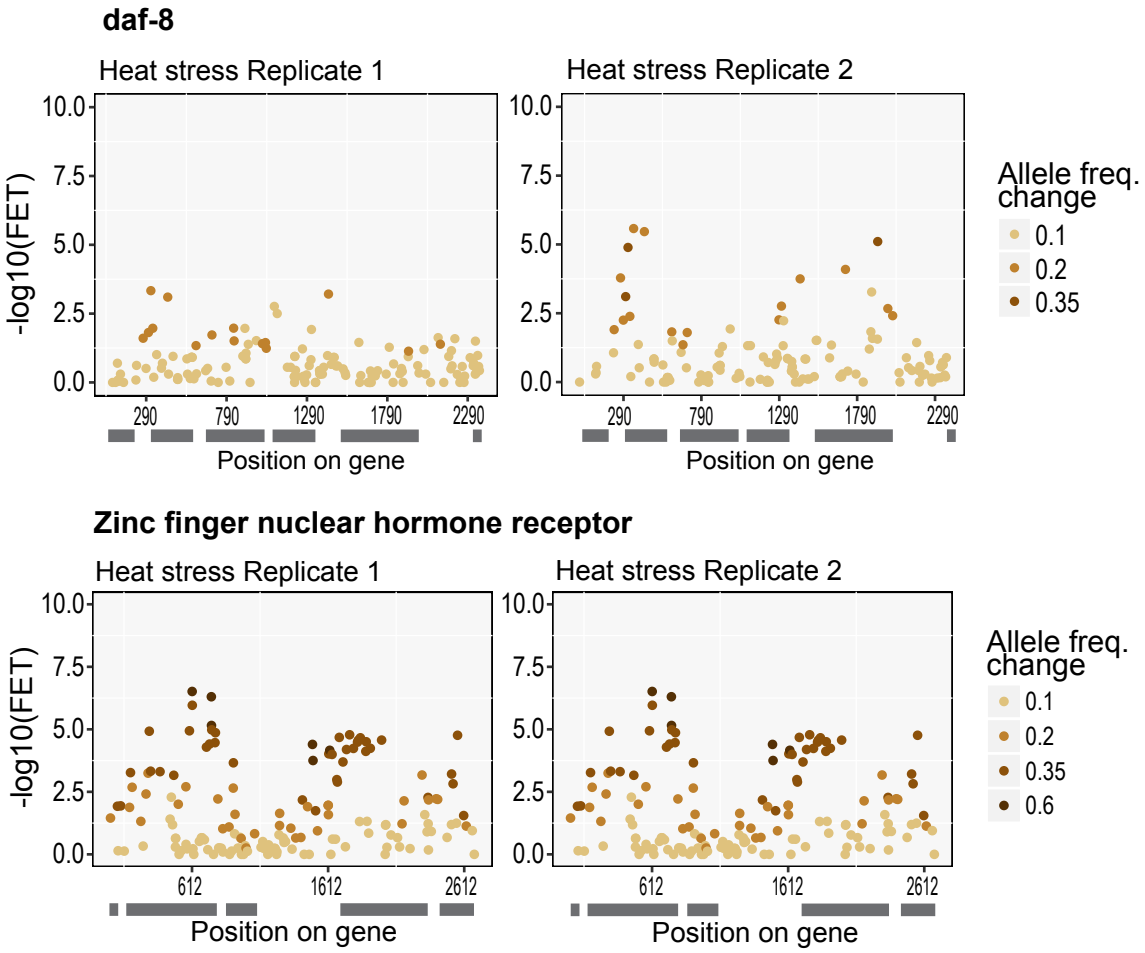


Figure S2.8: Plot of $-\log_{10}$ *p*-value of Fisher's exact test colored by allele frequency changes on a SNP by SNP basis in both acute heat stress evolved populations for two genes: daf-8 and a zinc finger nuclear hormone receptor. Magnitude of allele frequency change is similar in both replicates.

APPENDIX B

SUPPORTING INFORAMTION FOR CHAPTER III

Migration Treatment	Replicate	π lab-adapted population	π heat stress population
m = 0.00	1	0.0239	0.0262
m = 0.00	2	0.023	0.0233
m = 0.00	3	0.0251	0.0234
m = 0.00	4	0.0222	0.236
m = 0.00	5	0.0243	0.0226
m = 0.05	1	0.0229	0.0241
m = 0.05	2	0.0229	0.0243
m = 0.05	3	0.0243	0.0248
m = 0.05	4	0.0237	0.0239
m = 0.05	5	0.0245	0.0239
m = 0.20	1	0.022	0.0229
m = 0.20	2	0.0256	0.0238
m = 0.20	3	0.0226	0.0222
m = 0.20	4	0.0224	0.0224
m = 0.20	5	0.0224	0.0244

Table S3.1: Genome wide mean nucleotide diversity (π) values. Means calculated as mean of 1000 bp non-overlapping sliding windows. There is no pattern of change in mean π with either migration or selection treatment ($F_{(5,24)} = 0.90$, $p = 0.5$). Mean π in ancestor population: 0.0234.

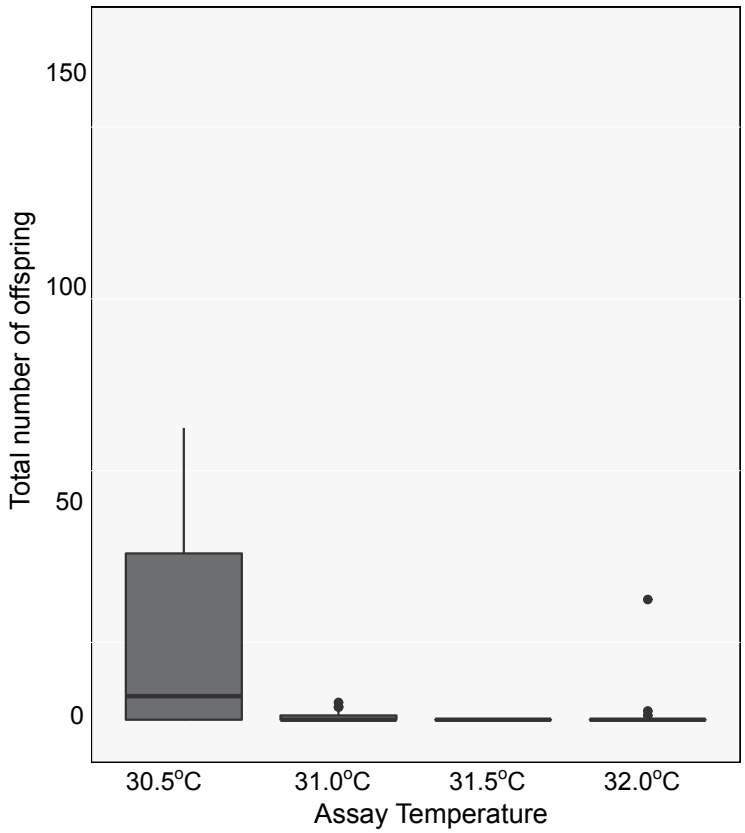


Figure S3.1: Female fecundity in chronic heat stress condition in the 75 generation lab adapted population (experiment base application). Worms are effectively sterile above 31°C.

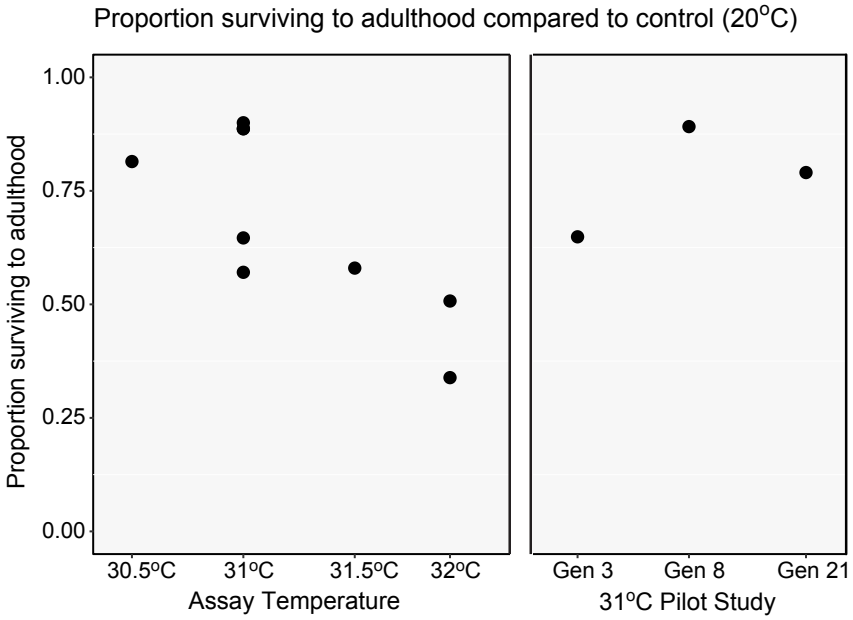


Figure S3.2: Results of pilot survival assays. Proportions shown were obtained by dividing the number of surviving adult worms in the heat stress condition over the number of worms in the control (20°C) condition. Effect of survival was not consistent or severe enough to use as signal of adaptation to the heat stress environment.

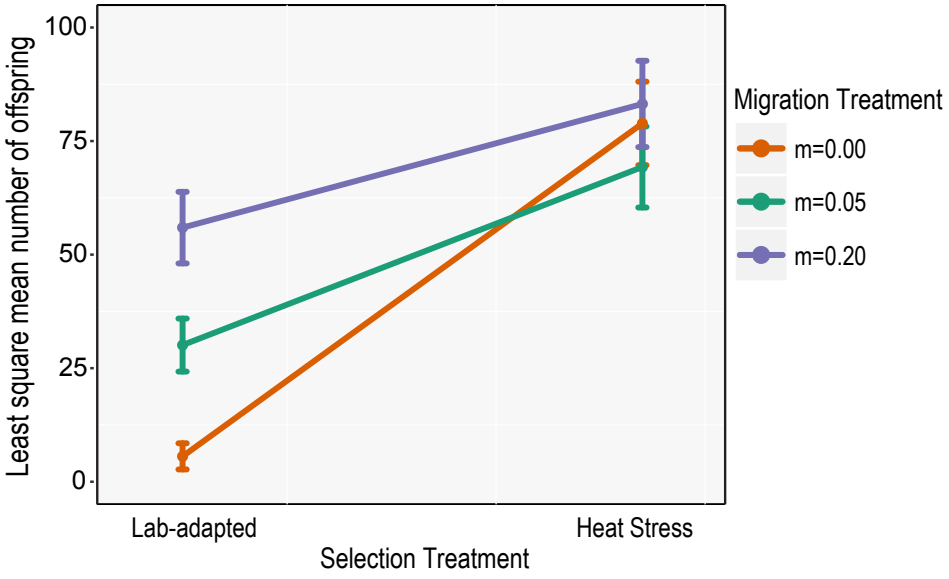


Figure S3.3: Least square means of female fecundity at 31°C (generation 40). The interaction between the migration and selection treatments is clearly seen here as a change in the slope of each line.

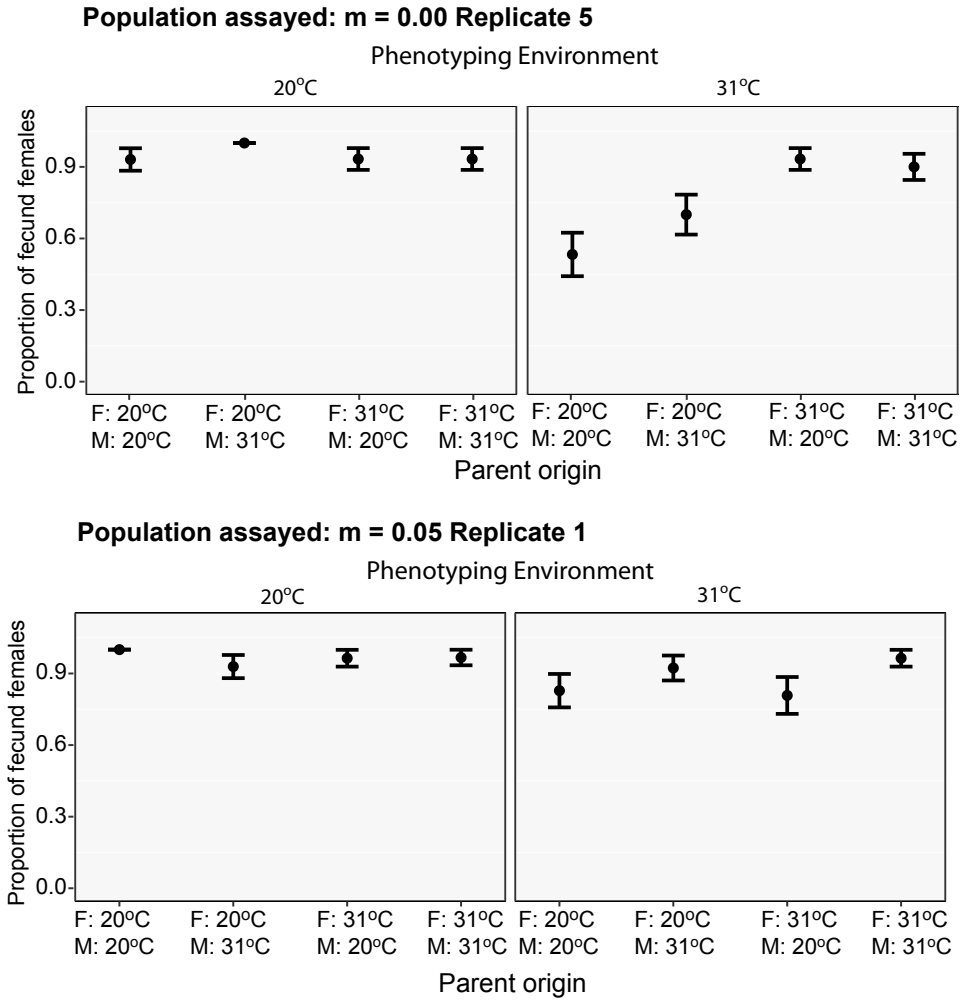


Figure S3.4: Proportion (plus and minus standard error) of fecund and sterile worms in 20°C and 31°C, separated by origin of parent. Twenty-five to thirty individuals were assayed per parent origin combination. Individuals were moved into the phenotyping environment post population age synchronization and were in the phenotyping environment for ~24h (31°C) or 48h (20°C) before being picked to mating plates. Migrants were able to successfully mate in both environments.

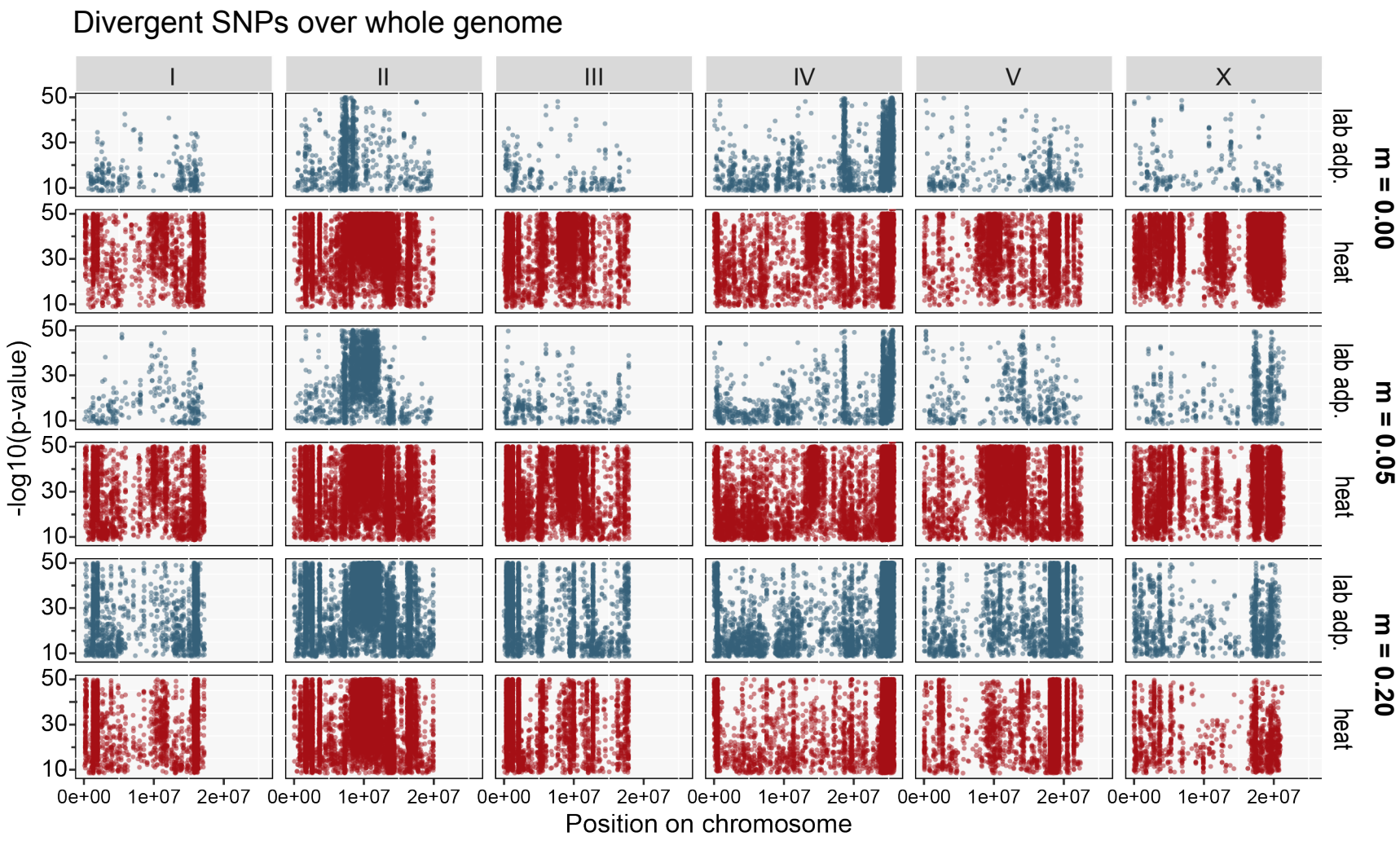


Figure S3.5: Divergent SNPs from the ancestor population in all evolved populations identified by Cochran–Mantel–Haenszel test over whole genome.

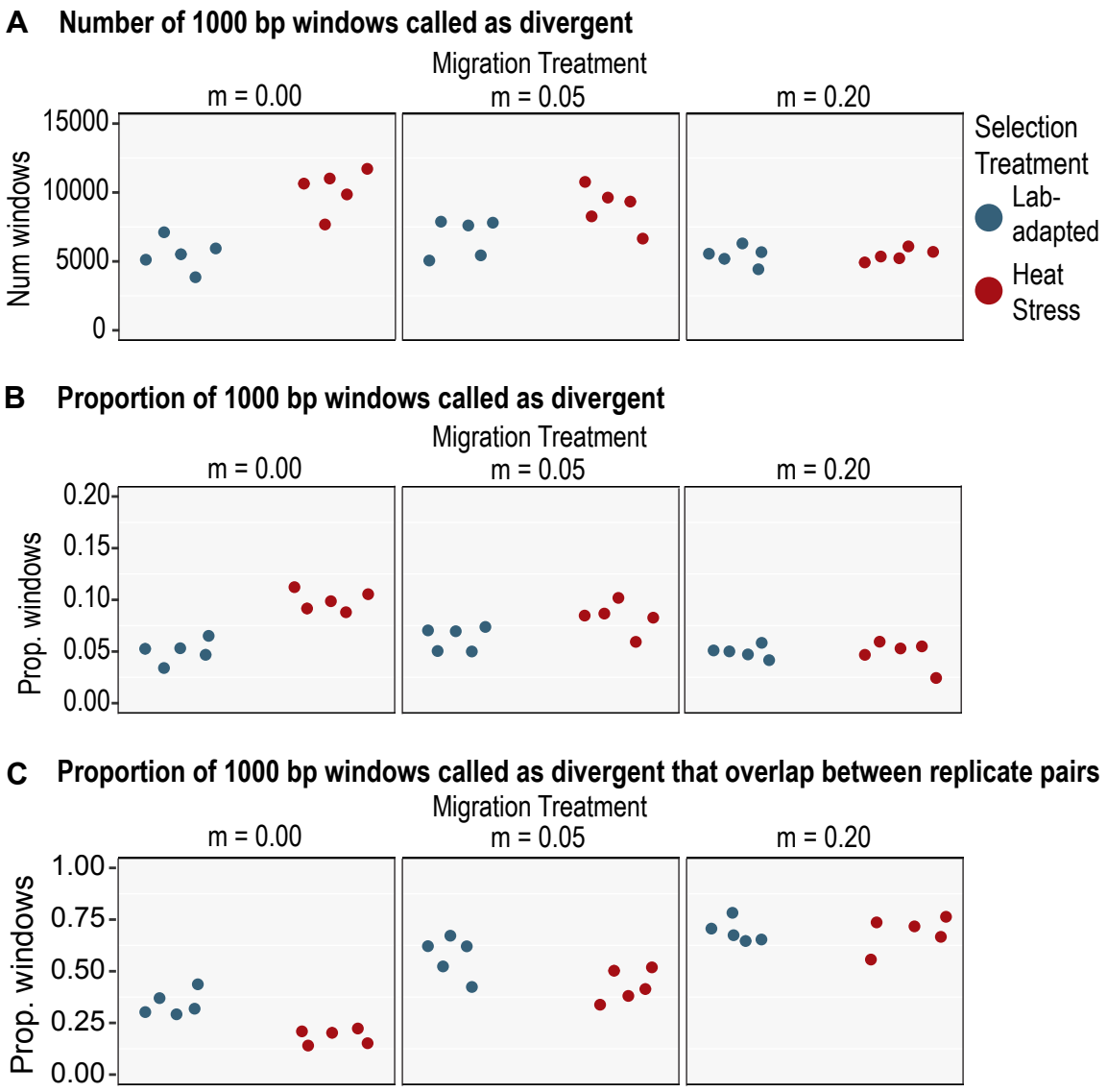


Figure S3.6: Summary statistics for significantly divergent 1000 bp sliding windows. (A) Total number of divergent windows identified in each evolved population. (B) Proportion of 1000 bp windows called as divergent – overall, a large proportion of the genome response to selection (including in the no migration lab adapted populations). (C) Proportion of divergent windows that overlap between replicate pairs. A replicate pair is made up of the two evolved populations that exchanged individuals in the low and high migration treatment populations. Increasing migration increases the similarity of the genomic response between replicate pairs.

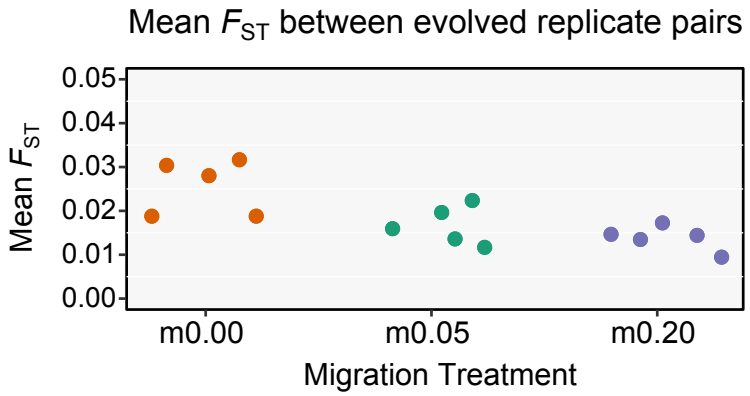


Figure S3.7: Genome wide mean F_{ST} between evolved replicate pairs (i.e. between a lab-adapted and heat stress evolved population pair) for each replicate separated by migration treatment. Mean F_{ST} increases in the absence of migration.

REFERENCES CITED

Ali, Omar A, Sean M O'Rourke, Stephen J Amish, Mariah H Meek, Gordon Luikart, Carson Jeffres, and Michael R Miller. 2016. "RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping." *Genetics* 202 (2): 389–400. doi:10.1534/genetics.115.183665.

An, Jae Hyung, and T Keith Blackwell. 2003. "SKN-1 Links *C. Elegans* Mesendodermal Specification to a Conserved Oxidative Stress Response." *Genes & Development* 17 (15). Cold Spring Harbor Lab: 1882–93. doi:10.1101/gad.1107803.

Anders, Simon, Paul Theodor Pyl, and Wolfgang Huber. 2014. "HTSeq—a Python Framework to Work with High-Throughput Sequencing Data." *Bioinformatics* 31 (2). Oxford University Press: btu638–169. doi:10.1093/bioinformatics/btu638.

Arnold, S J. 1992. "Constraints on Phenotypic Evolution." *The American Naturalist* 140 Suppl 1 (November). University of Chicago Press: S85–S107. doi:10.1086/285398.

Bailey, Susan F, and Thomas Bataillon. 2016. "Can the Experimental Evolution Programme Help Us Elucidate the Genetic Basis of Adaptation in Nature?." *Molecular Ecology* 25 (1): 203–18. doi:10.1111/mec.13378.

Barrett, Rowan D H, Sean M Rogers, and Dolph Schluter. 2008. "Natural Selection on a Major Armor Gene in Threespine Stickleback." *Science* 322: 255–57. doi:10.1029/2005GL025531.

Barton, N H, and M Turelli. 1989. "Evolutionary Quantitative Genetics: How Little Do We Know?." *Annual Review of Genetics* 23 (January): 337–70.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using Lme4." *Journal of Statistical Software* 67 (1): 1–48. doi:10.18637/jss.v067.i01.

Behrman, Emily L, Virginia M Howick, Martin Kapun, Fabian Staubach, Alan O Bergland, Dmitri A Petrov, Brian P Lazzaro, and Paul S Schmidt. 2018. "Rapid Seasonal Evolution in Innate Immunity of Wild Drosophila Melanogaster." *Proceedings. Biological Sciences / the Royal Society* 285 (1870): 20172599–10. doi:10.1098/rspb.2017.2599.

Brenner, S. 1974. "The Genetics of *Caenorhabditis Elegans*." *Genetics* 77 (1). Genetics Society of America: 71–94.

Bulmer, M G. 1972. "Multiple Niche Polymorphism." *American Naturalist*. doi:10.2307/2459931.

Burri, Reto, Alexander Nater, Takeshi Kawakami, Carina F Mugal, Pall I Olason, Linnea Smeds, Alexander Suh, et al. 2015. “Linked Selection and Recombination Rate Variation Drive the Evolution of the Genomic Landscape of Differentiation Across the Speciation Continuum of *Ficedula* Flycatchers.” *Genome Research* 25 (11): 1656–65. doi:10.1101/gr.196485.115.

Catchen, Julian, Paul A Hohenlohe, Susan Bassham, Angel Amores, and William A Cresko. 2013. “Stacks: an Analysis Tool Set for Population Genomics.” *Molecular Ecology* 22 (11): 3124–40. doi:10.1111/mec.12354.

Charlesworth, B. 1996. “Background Selection and Patterns of Genetic Diversity in Drosophila Melanogaster.” *Genetics Research* 68 (2): 131–49.

Charlesworth, B, D Charlesworth, and Magnus Nordborg. 1997. “The Effects of Local Selection, Balanced Polymorphism and Background Selection on Equilibrium Patterns of Genetic Diversity in Subdivided Populations,” November, 1–20.

Charlesworth, Brian. 1990. “Optimization Models, Quantitative Genetics, and Mutation.” *Evolution* 44 (3): 520–38.

Charlesworth, Brian. 2012. “The Effects of Deleterious Mutations on Evolution at Linked Sites.” *Genetics* 190 (1): 5–22. doi:10.1534/genetics.111.134288.

Cheverud, J M. 1984. “Quantitative Genetics and Developmental Constraints on Evolution by Selection.” *Journal of Theoretical Biology* 110 (2): 155–71.

Cheverud, James M. 2006. “Genetic Architecture of Quantitative Varaition.” In *Evolutionary Genetics*, edited by Charles W Fox and Jason B Wolf, 288–309.

Comeron, J M, A Williford, and R M Kliman. 2007. “The Hill–Robertson Effect: Evolutionary Consequences of Weak Selection and Linkage in Finite Populations.” *Heredity* 100 (1): 19–31. doi:10.1038/sj.hdy.6801059.

Consortium, Schizophrenia Working Group of the Psychiatric Genomics, Stephan Ripke, Benjamin M Neale, Aiden Corvin, James T R Walters, Kai-How Farh, Peter A Holmans, et al. 2014. “Biological Insights From 108 Schizophrenia-Associated Genetic Loci.” *Nature* 511 (7510). Nature Publishing Group: 421–17. doi:10.1038/nature13595.

Coyne, Jerry A, and Allen H Orr. 2004. *Speciation*. Sunderland.

Cruickshank, Tami E, and Matthew W Hahn. 2014. “Reanalysis Suggests That Genomic Islands of Speciation Are Due to Reduced Diversity, Not Reduced Gene Flow.” *Molecular Ecology* 23 (13): 3133–57. doi:10.1111/mec.12796.

Cutter, A D, S E Baird, and D Charlesworth. 2006. “High Nucleotide Polymorphism and Rapid Decay of Linkage Disequilibrium in Wild Populations of *Caenorhabditis Remanei*.” *Genetics* 174 (2): 901–13. doi:10.1534/genetics.106.061879.

Danecek, Petr, Adam Auton, Goncalo Abecasis, Cornelis A Albers, Eric Banks, Mark A DePristo, Robert E Handsaker, et al. 2011. “The Variant Call Format and VCFtools.” *Bioinformatics* 27 (15): 2156–58. doi:10.1093/bioinformatics/btr330.

Darling, Nicola J, and Simon J Cook. 2014. “The Role of MAPK Signalling Pathways in the Response to Endoplasmic Reticulum Stress.” *BBA - Molecular Cell Research* 1843 (10). Elsevier B.V.: 2150–63. doi:10.1016/j.bbamcr.2014.01.009.

Diaz, S Anaid, Jan Lindström, and Daniel T Hayden. 2008. “Basic Demography of *Caenorhabditis Remanei* Cultured Under Standard Laboratory Conditions.” *Journal of Nematology* 40 (3): 167–78.

East, Edward M. 1910. “A Mendelian Interpretation of Variation That Is Apparently Continuous.” *The American Naturalist* 44 (518): 65–82.

Ehrlich, Paul R, and Peter H Raven. 1969. “Differentiation of Populations” 165 (3899): 1228–32.

Ellegren, Hans, Linnea Smeds, Reto Burri, Pall I Olason, Niclas Backström, Takeshi Kawakami, Axel Künstner, et al. 2012. “The Genomic Landscape of Species Divergence in *Ficedula* Flycatchers.” *Nature* 491 (7426). Nature Publishing Group: 756–60. doi:10.1038/nature11584.

Emms, David M, and Steven Kelly. 2015. “OrthoFinder: Solving Fundamental Biases in Whole Genome Comparisons Dramatically Improves Orthogroup Inference Accuracy.” *Genome Biology* 16 (1). BioMed Central: 157. doi:10.1186/s13059-015-0721-2.

Ewbank, Jonathan. 2006. “Signaling in the Immune Response.” *WormBook*, 1–12. doi:10.1895/wormbook.1.83.1.

Falconer, D S, and Trudy F C Mackay. 1996. *Introduction to Quantitative Genetics*. Harlow.

Feder, J L, P Nosil, A C Wacholder, S P Egan, S H Berlocher, and S M Flaxman. 2014. “Genome-Wide Congealing and Rapid Transitions Across the Speciation Continuum During Speciation with Gene Flow.” *Journal of Heredity* 105 (S1): 810–20. doi:10.1093/jhered/esu038.

Feder, Jeffrey L, Joseph B Roethele, Kenneth Filchak, Julie Niedbalski, and Jeanne Romero-Severson. 2003. "Evidence for Inversion Polymorphism Related to Sympatric Host Race Formation in the Apple Maggot Fly, Rhagoletis Pomonella" *Genetics* 163 (3). Genetics Society of America: 939–53.

Feder, Jeffrey L, Samuel M Flaxman, Scott P Egan, Aaron A Comeault, and Patrik Nosil. 2013. "Geographic Mode of Speciation and Genomic Divergence." *Annual Review of Ecology, Evolution, and Systematics* 44 (1): 73–97. doi:10.1146/annurev-ecolsys-110512-135825.

Feder, Jeffrey L, Scott P Egan, and Patrik Nosil. 2012. "The Genomics of Speciation-with-Gene-Flow." *Trends in Genetics : TIG* 28 (7): 342–50. doi:10.1016/j.tig.2012.03.009.

Felsenstein, Joseph. 1976. "The Theoretical Population Genetics of Variable Selection and Migration." *Annual Review of Genetics* 10 (January): 253–80.

Felsenstein, Joseph. 1981. "Skepticism Towards Santa Rosalia, or Why Are There So Few Kinds of Animals?." *Evolution* 35 (1): 124–38. doi:10.2307/2407946.

Fierst, Janna L, John H Willis, Cristel G Thomas, Wei Wang, Rose M Reynolds, Timothy E Ahearne, Asher D Cutter, and Patrick C Phillips. 2015. "Reproductive Mode and the Evolution of Genome Size and Structure in *Caenorhabditis* Nematodes." Edited by Mark Blaxter. *PLoS Genetics* 11 (6). Public Library of Science: e1005323. doi:10.1371/journal.pgen.1005323.

Fisher, R A. 1918. "The Correlation Between Relative on the Supposition of Mendelian Inheritance." *Transactions of the Royal Society of Edinburgh* 52 (October): 399–433.

Flaxman, Samuel M, Aaron C Wacholder, Jeffrey L Feder, and Patrik Nosil. 2014. "Theoretical Models of the Influence of Genomic Architecture on the Dynamics of Speciation." *Molecular Ecology* 23 (16): 4074–88. doi:10.1111/mec.12750.

Foote, Andrew D. 2018. "Sympatric Speciation in the Genomic Era." *Trends in Ecology & Evolution* 33 (2). Elsevier Ltd: 85–95. doi:10.1016/j.tree.2017.11.003.

Franssen, Susanne U, Viola Nolte, Ray Tobler, and Christian Schlötterer. 2015. "Patterns of Linkage Disequilibrium and Long Range Hitchhiking in Evolving Experimental Drosophila Melanogaster Populations." *Molecular Biology and Evolution* 32 (2): 495–509. doi:10.1093/molbev/msu320.

Gagnaire, Pierre-Alexandre, Scott A Pavey, Eric Normandeau, and Louis Bernatchez. 2013. "The Genetic Architecture of Reproductive Isolation During Speciation-with-Gene-Flow in Lake Whitefish Species Pairs Assessed by RAD Sequencing." *Evolution* 67 (9): 2483–97. doi:10.1111/evo.12075.

Gibson, Greg. 2018. “Population Genetics and GWAS: a Primer.” *PLoS Biology* 16 (3): e2005485–86. doi:10.1371/journal.pbio.2005485.

Graham, Allie M, Philip Lavretsky, Violeta Muñoz-Fuentes, Andy J Green, Robert E Wilson, and Kevin G McCracken. 2017. “Migration-Selection Balance Drives Genetic Differentiation in Genes Associated with High-Altitude Function in the Speckled Teal (Anas Flavirostris) in the Andes.” *Genome Biology and Evolution* 10 (1): 14–32. doi:10.1093/gbe/evx253.

Griswold, C K. 2006. “Gene Flow's Effect on the Genetic Architecture of a Local Adaptation and Its Consequences for QTL Analyses.” *Heredity* 96 (6). Nature Publishing Group: 445–53. doi:10.1038/sj.hdy.6800822.

Haldane, J B S. 1930. “A Mathematical Theory of Natural and Artificial Selection. (Part VI, Isolation.).” *Mathematical Proceedings of the Cambridge Philosophical Society* 26 (2). University of Oregon Library: 220–30. doi:10.1017/S0305004100015450.

Han, Fan, Sangeet Lamichhaney, B Rosemary Grant, Peter R Grant, Leif Andersson, and Matthew T Webster. 2017. “Gene Flow, Ancient Polymorphism, and Ecological Adaptation Shape the Genomic Landscape of Divergence Among Darwin's Finches.” *Genome Research* 27 (6): 1004–15. doi:10.1101/gr.212522.116.

Hartl, Daniel L, and Andrew G Clark. 1997. *Principles of Population Genetics*. Sunderland.

Hill, W G. 2009. “Understanding and Using Quantitative Genetic Variation.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 365 (1537): 73–85. doi:10.1098/rstb.2009.0203.

Hill, W G, and Alan Robertson. 1966. “The Effect of Linkage on Limits to Artificial Selection.” *Genetics Research* 8: 269–94. doi:10.1017/S001667230800949X.

Hill, William G. 2014. “Applications of Population Genetics to Animal Breeding, From Wright, Fisher and Lush to Genomic Prediction.” *Genetics* 196 (1): 1–16. doi:10.1534/genetics.112.147850.

Hirsch, Candice N, Sherry A Flint-Garcia, Timothy M Beissinger, Steven R Eichten, Shweta Deshpande, Kerrie Barry, Michael D McMullen, et al. 2014. “Insights Into the Effects of Long-Term Artificial Selection on Seed Size in Maize.” *Genetics* 198 (1): 409–21. doi:10.1534/genetics.114.167155.

Hoban, Sean, Joanna L Kelley, Katie E Lotterhos, Michael F Antolin, Gideon Bradburd, David B Lowry, Mary L Poss, Laura K Reed, Andrew Storfer, and Michael C Whitlock. 2016. “Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions.” *The American Naturalist* 188 (4). University of Chicago PressChicago, IL: 379–97. doi:10.1086/688018.

Hodgkin, Jonathon. 1987. “Primary Sex Determination in the Nematode.” *Development (101), Suppl.*, January, 5–15.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. “Simultaneous Inference in General Parametric Models.” *Biometrical Journal. Biometrische Zeitschrift* 50 (3). Wiley-Blackwell: 346–63. doi:10.1002/bimj.200810425.

Johansson, Anna M, Mats E Pettersson, Paul B Siegel, and Örjan Carlborg. 2010. “Genome-Wide Effects of Long-Term Divergent Selection.” Edited by Bruce Walsh. *PLoS Genetics* 6 (11). Public Library of Science: e1001188. doi:10.1371/journal.pgen.1001188.

Jonas, Agnes, Thomas Taus, Carolin Kosiol, Christian Schlötterer, and Andreas Futschik. 2016. “Estimating the Effective Population Size From Temporal Allele Frequency Changes in Experimental Evolution.” *Genetics* 204 (2): 723–35. doi:10.1534/genetics.116.191197.

Jorde, P E, and N Ryman. 2007. “Unbiased Estimator for Genetic Drift and Effective Population Size.” *Genetics* 177 (2): 927–35. doi:10.1534/genetics.107.075481.

Jovelin, R, B C Ajie, and P C Phillips. 2003. “Molecular Evolution and Quantitative Variation for Chemosensory Behaviour in the Nematode Genus *Caenorhabditis*.” *Molecular Ecology* 12 (5): 1325–37.

Kemper, Kathryn E, Sarah J Saxton, Sunduimijid Bolormaa, Benjamin J Hayes, and Michael E Goddard. 2014. “Selection for Complex Traits Leaves Little or No Classic Signatures of Selection.” *BMC Genomics* 15 (1). BMC Genomics: 1–14. doi:10.1186/1471-2164-15-246.

Kenyon, Cynthia. 2005. “The Plasticity of Aging: Insights From Long-Lived Mutants.” *Cell* 120 (4): 449–60. doi:10.1016/j.cell.2005.02.002.

Kenyon, Cynthia J. 2010. “The Genetics of Ageing.” *Nature* 464 (7288): 504–12. doi:10.1038/nature08980.

Kimura, Motoo. 1983. *The Neutral Theory of Molecular Evolution*. New York City.

Kirkpatrick, Mark, and Nick Barton. 2006. “Chromosome Inversions, Local Adaptation and Speciation.” *Genetics* 173 (1). Genetics Society of America: 419–34. doi:10.1534/genetics.105.047985.

Knight, Christopher G, Nicole Zitzmann, Sripadi Prabhakar, Robin Antrobus, Raymond Dwek, Holger Hebestreit, and Paul B Rainey. 2006. “Unraveling Adaptive Evolution: How a Single Point Mutation Affects the Protein Coregulation Network.” *Nature Genetics* 38 (9): 1015–22. doi:10.1038/ng1867.

Kofler, R, E Versace, R Tobler, S U Franssen, and Christian Schlötterer. 2015. "Combining Experimental Evolution with Next-Generation Sequencing: a Powerful Tool to Study Adaptation From Standing Genetic Variation." *Heredity* 114 (5). Nature Publishing Group: 431–40. doi:10.1038/hdy.2014.86.

Kofler, Robert, and Christian Schlötterer. 2013. "A Guide for the Design of Evolve and Resequencing Studies." *Molecular Biology and Evolution* 31 (2). Oxford University Press: 474–83. doi:10.1093/molbev/mst221.

Kofler, Robert, Pablo Orozco-terWengel, Nicola De Maio, Ram Vinay Pandey, Viola Nolte, Andreas Futschik, Carolin Kosiol, and Christian Schlötterer. 2011. "PoPoolation: a Toolbox for Population Genetic Analysis of Next Generation Sequencing Data From Pooled Individuals." Edited by Manfred Kayser. *PLoS ONE* 6 (1). Public Library of Science: e15925. doi:10.1371/journal.pone.0015925.

Kofler, Robert, Ram Vinay Pandey, and Christian Schlötterer. 2011. "PoPoolation2: Identifying Differentiation Between Populations Using Sequencing of Pooled DNA Samples (Pool-Seq)." *Bioinformatics* 27 (24). Oxford University Press: 3435–36. doi:10.1093/bioinformatics/btr589.

Krimbas, Costas B, and Spyros Tsakas. 1971. "The Genetics of Dacus Oleae. v. Changes of Esterase Polymorphism in a Natural Population Following Insecticide Control-Selection or Drift." *Evolution* 25: 454–60.

Lande, R. 1979. "Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry." *Evolution* 33 (1): 402–16.

Lande, Russell. 1984. "The Genetic Correlation Between Characters Maintained by Selection, Linkage and Inbreeding." *Genetical Research* 44 (3): 309–20.

Lango Allen, Hana, Karol Estrada, Guillaume Lettre, Sonja I Berndt, Michael N Weedon, Fernando Rivadeneira, Cristen J Willer, et al. 2010. "Hundreds of Variants Clustered in Genomic Loci and Biological Pathways Affect Human Height." *Nature* 467 (7317): 832–38. doi:10.1038/nature09410.

Larson, Wesley A, Morten T Limborg, Garrett J McKinney, Daniel E Schindler, James E Seeb, and Lisa W Seeb. 2016. "Genomic Islands of Divergence Linked to Ecotypic Variation in Sockeye Salmon." *Molecular Ecology* 26 (2): 554–70. doi:10.1111/mec.13933.

Lee, Raymond Y N, Kevin L Howe, Todd W Harris, Valerio Arnaboldi, Scott Cain, Juancarlos Chan, Wen J Chen, et al. 2018. "WormBase 2017: Molting Into a New Stage." *Nucleic Acids Research* 46 (D1): D869–74. doi:10.1093/nar/gkx998.

Legendre, Pierre. 2018. "Lmodel2: Model II Regression."

Levene, Howard. 1953. “Genetic Equilibrium When More Than One Ecological Niche Is Available.” *American Naturalist*, September, 331–33.

Lewontin, R C, and Jesse Krakauer. 1973. “Distribution of Gene Freqeuncy as a Test of the Theory of the Selective Neutrality of Polymorphisms.” *Genetics* 74 (May): 175–95.

Lewontin, Richard C. 1974. *The Genetic Basis of Evolutionary Change*. New York City.

Linnen, Catherine R, Yu-Ping Poh, Brant K Peterson, Rowan D H Barrett, Joanna G Larson, Jeffrey D Jensen, and Hopi E Hoekstra. 2013. “Adaptive Evolution of Multiple Traits Through Multiple Mutations at a Single Gene.” *Science* 339 (6125). American Association for the Advancement of Science: 1312–16. doi:10.1126/science.1233213.

Lotterhos, Katie E, and Michael C Whitlock. 2015. “The Relative Power of Genome Scans to Detect Local Adaptation Depends on Sampling Design and Statistical Method.” *Molecular Ecology* 24 (5): 1031–46. doi:10.1111/mec.13100.

Love, Michael I, Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12). BioMed Central Ltd. doi:10.1186/s13059-014-0550-8.

Luikart, Gordon, Phillip R England, David Tallmon, Steve Jordan, and Pierre Taberlet. 2003. “The Power and Promise of Population Genomics: From Genotyping to Genome Typing.” *Nature Reviews Genetics* 4 (12): 981–94. doi:10.1038/nrg1226.

Lynch, Michael, and Bruce Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Sunderland: Sinauer Associates.

Mackay, Trudy F. C. 2004. “The Genetic Architecture of Quantitative Traits: Lessons From *Drosophila*.” *Current Opinion in Genetics & Development* 14: 253–57.

Mackay, Trudy F. C., E A Stone, and J F Ayroles. 2009. “The Genetics of Quantitative Traits: Challenges and Prospects.” *Nature Reviews Genetics*.

Makvandi-Nejad, Shokouh, Gabriel E Hoffman, Jeremy J Allen, Erin Chu, Esther Gu, Alyssa M Chandler, Ariel I Loredo, et al. 2012. “Four Loci Explain 83% of Size Variation in the Horse.” Edited by Michael Hofreiter. *PLoS ONE* 7 (7). Public Library of Science: e39929. doi:10.1371/journal.pone.0039929.

Maynard Smith, J. 1966. “Sympatric Speciation.” *The American Naturalist* 100 (916): 637–50.

Maynard Smith, J. 1970. “Genetic Polymorphism in a Varied Environment.” *American Naturalist* 104 (939): 487–90.

Michel, Andrew P, Sheina Sim, Patrik Nosil, Jeffrey L Feder, Michael S Taylor, and Thomas H Q Powell. 2010. "Widespread Genomic Divergence During Sympatric Speciation." *Proceedings of the National Academy of Sciences*, May, 1–6. doi:10.1073/pnas.1000939107/-/DCSupplemental.

Murphy, Coleen T, and Patrick T Hu. 2013. "Insulin/Insulin-Like Growth Factor Signaling in *C. Elegans*." *WormBook*, December, 1–43. doi:10.1895/wormbook.1.164.1.

Natarajan, C, N Inoguchi, R E Weber, A Fago, H Moriyama, and J F Storz. 2013. "Epistasis Among Adaptive Mutations in Deer Mouse Hemoglobin." *Science* 340 (6138): 1324–27. doi:10.1126/science.1236862.

Nelson, Thomas C, and William A Cresko. 2018. "Ancient Genomic Variation Underlies Repeated Ecological Adaptation in Young Stickleback Populations." *Evolution Letters* 2 (1): 9–21. doi:10.1002/evl3.37.

Nolte, Ilja M, Peter J van der Most, Behrooz Z Alizadeh, Paul IW de Bakker, H Marike Boezen, Marcel Bruinenberg, Lude Franke, et al. 2017. "Missing Heritability: Is the Gap Closing? an Analysis of 32 Complex Traits in the Lifelines Cohort Study." *European Journal of Human Genetics : EJHG* 25 (7). Nature Publishing Group: 877–85. doi:10.1038/ejhg.2017.50.

Noor, M A F, and S M Bennett. 2009. "Islands of Speciation or Mirages in the Desert? Examining the Role of Restricted Recombination in Maintaining Species." *Heredity* 103 (6). Nature Publishing Group: 439–44. doi:10.1038/hdy.2009.151.

Nosil, P, and B J Crespi. 2004. "Does Gene Flow Constrain Adaptive Divergence or Vice Versa? a Test Using Ecomorpohlogy and Sexual Isolation in *Timema Cristinae* Walking-Sticks." *Evolution* 58 (1): 102–12.

Nosil, P, and J L Feder. 2012. "Genomic Divergence During Speciation: Causes and Consequences." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 367 (1587): 332–42. doi:10.1098/rstb.2011.0263.

Nosil, Patrik, Daniel J Funk, and Daniel Ortiz-Barrientos. 2009. "Divergent Selection and Heterogeneous Genomic Divergence." *Molecular Ecology* 18 (3). Blackwell Publishing Ltd: 375–402. doi:10.1111/j.1365-294X.2008.03946.x.

Nosil, Patrik, Scott P Egan, and Daniel J Funk. 2008. "Heterogeneous Genomic Differentiation Between Walking-Stick Ecotypes: 'Isolation by Adaptation' and Multiple Roles for Divergent Selection." *Evolution* 62 (2). Blackwell Publishing Inc: 316–36. doi:10.1111/j.1558-5646.2007.00299.x.

Orozco-terWengel, Pablo, Martin Kapun, Viola Nolte, Robert Kofler, Thomas Flatt, and Christian Schlötterer. 2012. "Adaptation of Drosophila to a Novel Laboratory Environment Reveals Temporally Heterogeneous Trajectories of Selected Alleles" *Molecular Ecology* 21 (20). Blackwell Publishing Ltd: 4931–41. doi:10.1111/j.1365-294X.2012.05673.x.

Orr, H A, and J A Coyne. 1992. "The Genetics of Adaptation: a Reassessment." *The American Naturalist* 140 (5). University of Chicago Press: 725–42. doi:10.1086/285437.

Orr, H Allen. 2009. "Fitness and Its Role in Evolutionary Genetics." *Nature Reviews Genetics* 10 (8): 531–39.

Paaby, Annalise B, and Matthew V Rockman. 2013. "The Many Faces of Pleiotropy." *Trends in Genetics* 29 (2). Elsevier Ltd: 66–73. doi:10.1016/j.tig.2012.10.010.

Palopoli, Michael F, Colin Peden, Caitlin Woo, Ken Akiha, Megan Ary, Lori Cruze, Jennifer L Anderson, and Patrick C Phillips. 2015. "Natural and Experimental Evolution of Sexual Conflict Within *Caenorhabditis* Nematodes." *BMC Evolutionary Biology*: 1–13. doi:10.1186/s12862-015-0377-2.

Pettersson, Mats E, Anna M Johansson, Paul B Siegel, and Örjan Carlborg. 2013. "Dynamics of Adaptive Alleles in Divergently Selected Body Weight Lines of Chickens." *G3: Genes Genomics Genetics* 3 (12). Genetics Society of America: 2305–12. doi:10.1534/g3.113.008375.

Pfeifer, Susanne P, Stefan Laurent, Vitor C Sousa, Catherine R Linnen, Matthieu Foll, Laurent Excoffier, Hopi E Hoekstra, and Jeffrey D Jensen. 2018. "The Evolutionary History of Nebraska Deer Mice: Local Adaptation in the Face of Strong Gene Flow." *Molecular Biology and Evolution* 35 (4): 792–806. doi:10.1093/molbev/msy004.

Phillips, Patrick C, and Katrina L McGuigan. 2006. "Evolution of Genetic Variance-Covariance Structure." In *Evolutionary Genetics*, edited by Charles W Fox and Jason B Wolf, 310–25.

Pritchard, Jonathan K, Joseph K Pickrell, and Graham Coop. 2010. "The Genetics of Human Adaptation: Hard Review Sweeps, Soft Sweeps, and Polygenic Adaptation." *Current Biology* 20 (4). Elsevier Ltd: R208–15. doi:10.1016/j.cub.2009.11.055.

Przeworski, Molly, Graham Coop, and Jeffrey D Wall. 2005. "The Signature of Positive Selection on Standing Genetic Variation." *Evolution* 59 (11). Society for the Study of Evolution: 2312–23. doi:10.2307/3449139?ref=no-x-route:9ac3acca65f7a492513341ef8b6e57a9.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A R Ferreira, David Bender, Julian Maller, et al. 2007. "PLINK: a Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *The American Journal of Human Genetics* 81 (3): 559–75. doi:10.1086/519795.

Quinlan, Aaron R, and Ira M Hall. 2010. "BEDTools: a Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics* 26 (6): 841–42. doi:10.1093/bioinformatics/btq033.

R Core Team. 2017. "R: a Language and Environment for Statistical Computing." Vienna. https://www.R-project.org/.

Räsänen, Katja, and Andrew P Hendry. 2008. "Disentangling Interactions Between Adaptive Divergence and Gene Flow When Ecology Drives Diversification." *Ecology Letters* 11 (6). Blackwell Publishing Ltd: 624–36. doi:10.1111/j.1461-0248.2008.01176.x.

Renaut, S, C J Grassa, S Yeaman, B T Moyers, Z Lai, N C Kane, J E Bowers, J M Burke, and L H Rieseberg. 2013. "Genomic Islands of Divergence Are Not Affected by Geography of Speciation in Sunflowers." *Nature Communications* 4 (May): 1827–28. doi:10.1038/ncomms2833.

Reynolds, Rose M, and Patrick C Phillips. 2013. "Natural Variation for Lifespan and Stress Response in the Nematode *Caenorhabditis Remanei*." Edited by Matt Kaeberlein. *PLoS ONE* 8 (4): e58212. doi:10.1371/journal.pone.0058212.

Rockman, Matthew V. 2012. "The QTN Program and the Alleles That Matter for Evolution: All That's Gold Does Not Glitter." *Evolution* 66 (1): 1–17. doi:10.1111/j.1558-5646.2011.01486.x.

Rodriguez, Miriam, L Basten Snoek, Mario De Bono, and Jan E Kammenga. 2013. "Worms Under Stress: *C. Elegans* Stress Response and Its Relevance to Complex Human Disease and Aging." *Trends in Genetics* 29 (6). Elsevier Ltd: 367–74. doi:10.1016/j.tig.2013.01.010.

Rosenblum, Erica Bree. 2006. "Convergent Evolution and Divergent Selection: Lizards at the White Sands Ecotone" *The American Naturalist* 167 (1). The University of Chicago Press: 1–15. doi:10.1086/498397.

Ross-Ibarra, J, M Tenaillon, and B S Gaut. 2009. "Historical Divergence and Gene Flow in the Genus Zea." *Genetics* 181 (4): 1399–1413. doi:10.1534/genetics.108.097238.

Samuk, Kieran, Gregory L Owens, Kira E Delmore, Sara E Miller, Diana J Rennison, and Dolph Schluter. 2017. "Gene Flow and Selection Interact to Promote Adaptive Divergence in Regions of Low Recombination." *Molecular Ecology* 26 (17): 4378–90. doi:10.1111/mec.14226.

Savage-Dunn, Cathy. 2005. “TGF-B Signaling.” *WormBook*, 1–12. doi:10.1895/wormbook.1.22.1.

Schlötterer, Christian, Raymond Tobler, Robert Kofler, and Viola Nolte. 2014. “Sequencing Pools of Individuals — Mining Genome-Wide Polymorphism Data Without Big Funding.” *Nature Reviews Genetics* 15 (11). Nature Publishing Group: 749–63. doi:10.1038/nrg3803.

Sheng, Zheya, Mats E Pettersson, Christa F Honaker, Paul B Siegel, and Örjan Carlborg. 2015. “Standing Genetic Variation as a Major Contributor to Adaptation in the Virginia Chicken Lines Selection Experiment.” *Genome Biology* 16 (1). BioMed Central: 219. doi:10.1186/s13059-015-0785-z.

Sikkink, Kristin L, Catherine M Ituarte, Rose M Reynolds, William A Cresko, and Patrick C Phillips. 2014. “The Transgenerational Effects of Heat Stress in the Nematode *Caenorhabditis Remanei* Are Negative and Rapidly Eliminated Under Direct Selection for Increased Stress Resistance in Larvae.” *Genomics* 104 (6 Pt A): 438–46. doi:10.1016/j.ygeno.2014.09.014.

Sikkink, Kristin L, Rose M Reynolds, Catherine M Ituarte, William A Cresko, and Patrick C Phillips. 2014. “Rapid Evolution of Phenotypic Plasticity and Shifting Thresholds of Genetic Assimilation in the Nematode *Caenorhabditis Remanei*.” *G3: Genes Genomics Genetics* 4 (6). Genetics Society of America: 1103–12. doi:10.1534/g3.114.010553.

Sikkink, Kristin L, Rose M Reynolds, William A Cresko, and Patrick C Phillips. 2015. “Environmentally Induced Changes in Correlated Responses to Selection Reveal Variable Pleiotropy Across a Complex Genetic Network.” *Evolution* 69 (5): 1128–42. doi:10.1111/evo.12651.

Slate, Jon. 2005. “Invited Review: Quantitative Trait Locus Mapping in Natural Populations: Progress, Caveats and Future Directions.” *Molecular Ecology* 14 (2). Blackwell Science Ltd: 363–79. doi:10.1111/j.1365-294X.2004.02378.x.

Song, Hyun-Ok, Wonhae Lee, Kiyoung An, Hye-suk Lee, Jeong Hoon Cho, Zee-Yong Park, and Joohong Ahnn. 2009. “*C. Elegans* STI-1, the Homolog of Sti1/Hop, Is Involved in Aging and Stress Response.” *Journal of Molecular Biology* 390 (4). Elsevier Ltd: 604–17. doi:10.1016/j.jmb.2009.05.035.

Stankowski, Sean, James M Sobel, and Matthew A Streisfeld. 2015. “The Geography of Divergence with Gene Flow Facilitates Multitrait Adaptation and the Evolution of Pollinator Isolation in Mimulus Aurantiacus.” *Evolution* 69 (12): 3054–68. doi:10.1111/evo.12807.

Steiner, Cynthia C, Jesse N Weber, and Hopi E Hoekstra. 2007. “Adaptive Variation in Beach Mice Produced by Two Interacting Pigmentation Genes.” Edited by Mohamed A F Noor and Mohamed A F Noor. *PLoS Biology* 5 (9). Public Library of Science: e219. doi:10.1371/journal.pbio.0050219.

Stiernagle, Theresa. 2006. “Maintenance of C. Elegans.” *WormBook*, 1–11. doi:10.1895/wormbook.1.101.1.

Taus, Thomas, Andreas Futschik, and Christian Schlötterer. 2017. “Quantifying Selection with Pool-Seq Time Series Data.” *Molecular Biology and Evolution* 34 (11): 3023–34. doi:10.1093/molbev/msx225.

Teshima, Kosuke M, Graham Coop, and Molly Przeworski. 2006. “How Reliable Are Empirical Genomic Scans for Selective Sweeps?.” *Genome Research* 16 (6). Cold Spring Harbor Lab: 702–12. doi:10.1101/gr.5105206.

Thomas, Cristel G, Gavin C Woodruff, and Eric S Haag. 2012. “Causes and Consequences of the Evolution of Reproductive Mode in *Caenorhabditis* Nematodes.” *Trends in Genetics : TIG* 28 (5). Elsevier Ltd: 213–20. doi:10.1016/j.tig.2012.02.007.

Turner, Thomas L, and Matthew W Hahn. 2007. “Locus- and Population-Specific Selection and Differentiation Between Incipient Species of Anopheles Gambiae.” *Molecular Biology and Evolution* 24 (9): 2132–38. doi:10.1093/molbev/msm143.

Turner, Thomas L, Andrew D Stewart, Andrew T Fields, William R Rice, and Aaron M Tarone. 2011. “Population-Based Resequencing of Experimentally Evolved Populations Reveals the Genetic Basis of Body Size Variation in Drosophila Melanogaster.” Edited by Greg Gibson. *PLoS Genetics* 7 (3). Public Library of Science: e1001336. doi:10.1371/journal.pgen.1001336.

Turner, Thomas L, Matthew W Hahn, and Sergey V Nuzhdin. 2005. “Genomic Islands of Speciation in Anopheles Gambiae.” Edited by Nick Barton. *PLoS Biology* 3 (9). Public Library of Science: e285. doi:10.1371/journal.pbio.0030285.

Van der Auwera, Geraldine A, Mauricio O Carneiro, Chris Hartl, Ryan Poplin, Guillermo Del Angel, Ami Levy-Moonshine, Tadeusz Jordan, et al. 2013. “From FastQ Data to High Confidence Variant Calls: the Genome Analysis Toolkit Best Practices Pipeline.” *Current Protocols in Bioinformatics* 43. Hoboken, NJ, USA: John Wiley & Sons, Inc.: 11.10.1–.10.33. doi:10.1002/0471250953.bi1110s43.

Via, Sara. 1991. “The Genetic Structure of Host Plant Adaptation in a Spatial Patchwork: Demographic Variability Among Reciprocally Transplanted Pea Aphid Clones.” *Evolution* 45 (4): 827. doi:10.2307/2409692.

Via, Sara. 2012. “Divergence Hitchhiking and the Spread of Genomic Isolation During Ecological Speciation-with-Gene-Flow.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1587): 451–60. doi:10.1098/rstb.2011.0260.

Visscher, Peter M, Naomi R Wray, Qian Zhang, Pamela Sklar, Mark I McCarthy, Matthew A Brown, and Jian Yang. 2017a. “10 Years of GWAS Discovery: Biology, Function, and Translation.” *American Journal of Human Genetics* 101 (1): 5–22. doi:10.1016/j.ajhg.2017.06.005.

Visscher, Peter M, Naomi R Wray, Qian Zhang, Pamela Sklar, Mark I McCarthy, Matthew A Brown, and Jian Yang. 2017b. “10 Years of GWAS Discovery: Biology, Function, and Translation.” *The American Journal of Human Genetics* 101 (1). Cell Press: 5–22. doi:10.1016/j.ajhg.2017.06.005.

Vitti, Joseph J, Sharon R Grossman, and Pardis C Sabeti. 2013. “Detecting Natural Selection in Genomic Data.” *Annual Review of Genetics* 47 (1): 97–120. doi:10.1146/annurev-genet-111212-133526.

Wagner, Günter P, and Jianzhi Zhang. 2011. “The Pleiotropic Structure of the Genotype–Phenotype Map: the Evolvability of Complex Organisms.” *Nature Publishing Group* 12 (3). Nature Publishing Group: 204–13. doi:10.1038/nrg2949.

Waples, R K, W A Larson, and R S Waples. 2016. “Estimating Contemporary Effective Population Size in Non-Model Species Using Linkage Disequilibrium Across Thousands of Loci,” August. Nature Publishing Group, 1–8. doi:10.1038/hdy.2016.60.

Wellenreuther, Maren, and Bengt Hansson. 2016. “Detecting Polygenic Evolution: Problems, Pitfalls, and Promises.” *Trends in Genetics : TIG* 32 (3): 155–64. doi:10.1016/j.tig.2015.12.004.

Wood, Andrew R, Tõnu Esko, Jian Yang, Sailaja Vedantam, Tune H Pers, Stefan Gustafsson, Audrey Y Chu, et al. 2014. “Defining the Role of Common Variation in the Genomic and Biological Architecture of Adult Human Height.” *Nature Genetics* 46 (11). Nature Publishing Group: 1173–86. doi:10.1038/ng.3097.

Wright, Sewall. 1931. “Evolution in Medelian Populations.” *Genetics* 16 (March): 97–159.

Wright, Sewall. 1951. “The Genetical Structure of Populations.” *Annals of Eugenics* 15: 323–54.

Wright, Sewall. 1968. *Evolution and the Genetics of Populations Volume I: Genetic and Biometric Foundations*. Chicago.

Wu, T D, and S Nacu. 2010. “Fast and SNP-Tolerant Detection of Complex Variants and Splicing in Short Reads.” *Bioinformatics* 26 (7): 873–81. doi:10.1093/bioinformatics/btq057.

Yeaman, Sam. 2015. “Local Adaptation by Alleles of Small Effect*.” *The American Naturalist* 186 (S1): S74–S89. doi:10.1086/682405.

Yeaman, Sam, and Michael C Whitlock. 2011. “The Genetic Architecture of Adaptation Under Migration-Selection Balance.” *Evolution* 65 (7). Blackwell Publishing Inc: 1897–1911. doi:10.1111/j.1558-5646.2011.01269.x.

Yeaman, Sam, and Sarah P Otto. 2011. “Establishment and Maintenance of Adaptive Genetic Divergence Under Migration, Selection, and Drift.” *Evolution* 65 (7). Blackwell Publishing Inc: 2123–29. doi:10.1111/j.1558-5646.2011.01277.x.