UNIVERSITY OF CALIFORNIA,
IRVINE

Genomics of adaptation in Drosophila experimental evolution

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILISOPHY

in Biological Sciences

by

Mark Anthony Phillips

Dissertation Committee:
Professor Michael R. Rose, Chair
Professor Laurence D. Mueller
Assistant Professor J.J. Emerson

2018

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

AWKNOWLEDGEMENTS

First and foremost, I would like to thank my adviser Dr. Michael R. Rose for all of the guidance and support he provided over the course of my graduate studies. His scientific insights were an invaluable asset, and the confidence he showed in my ability to succeed was a constant motivator.

I would also like to acknowledge Dr. Laurence D. Mueller who in many ways served as a co-adviser along with Michael. Specifically, I would like to thank Larry for always making time when I was in need of his advice or professional expertise despite his many duties as a department chair.

I would like to thank Dr. J.J. Emerson who served as the third member of my dissertation committee. His expertise was a welcome addition during my committee meetings.

I am also grateful to two other faculty members at UC, Irvine, Dr. Anthony D. Long and Dr. Steve Frank. Tony was instrumental in the completion of my first chapter, and I have benefited greatly from all of my interactions with him. Steve was a member of my pre-advancement committee, and his unique perspective contributed greatly to my development as a scientist.

I was also fortunate to have Dr. Molly K. Burke as a collaborator of two of my dissertation projects. I have benefited greatly from her feedback, mentorship, and guidance at all stages of my graduate career.

Next, I would like to thank all of the members of the Rose and Mueller lab where I completed my dissertation research. In particular, I would like to acknowledge Thomas Barter, James Kezos, Grant Rutledge, Larry Cabral, and Zachary Greenspan. Working with these individuals was always a pleasure, and they have all contributed to my thesis in some fashion.

I would like to thank all of the friends I had made since moving to California to pursue my graduate studies. Their companionship and support has made graduate school one of the most enjoyable periods of my life thus far. In particular, I would like to acknowledge Aide Macias-Muñoz, Caitlin Looby, Jenifer Briner, Kevin Estipular, Thelma Castro, and Ramin Tasbhichi.

Lastly, I would like to acknowledge my closest friend, Georgios Lazos. His focus and work ethic been a constant source of inspiration.

CURRICULUM VITAE

Mark A. Phillips
Department of Ecology and Evolutionary Biology
321 Steinhaus Hall
University of California, Irvine
Irvine, CA 92697

Phone: 727-637-6797
Email: phillim1@uci.edu

Education

2018	Ph.D. Biological Sciences, University of California, Irvine
2015	M.S. Biological Sciences, University of California, Irvine
2011	B.S. Biological Sciences, University of Florida, Gainesville

Appointment History

2012-present	Graduate Student, Department of Ecology and Evolutionary Biology, Laboratory of Dr. Michael Rose, University of California, Irvine
2010-2011	Undergraduate Researcher, Department of Biology, Laboratory of Stuart McDaniel, University of Florida
2010	Undergraduate Researcher, Department of Biology, Laboratory of Doug Levey, University of Florida

Mentoring and Volunteer Experience

2012- present	Mentoring Undergraduate Researchers, Laboratories of Dr. Laurence Mueller and Dr. Michael Rose, University of California Irvine
2013 – present	Resident Scientist at Santa Ana High School
2014 - 2015	Irvine Unified School District Science Fair Judge
2015	Irvine Unedified School District Ask a Scientist Night

Teaching Experience

2012-present	**Graduate Teaching Assistant at UC Irvine** Bio E127 Physiology and Plant Ecology Bio 1A Life Sciences Bio E179L Field Freshwater Ecology Lab Bio E106 Ecology and Evolutionary Biology Bio 94 Organisms to Ecosystems Bio E112L Physiology Lab Bio E182 Mediterranean Ecosystems Bio 97 Genetics Bio 100 Scientific Writing

Funding Sources

2014-2018	Graduate Assistance in Areas of National Need Fellowship
2016	NSF Doctoral Dissertation Improvement Grant
2015	Carl Storm Underrepresented Minority Fellowship
2012	Competitive-Edge Summer Research Program, UC Irvine

Publications

Phillips M.A., Rutledge G.A., Kezos J.N., Talbott A., Matty S., Arian H., Mueller L.D., Rose M.R, Sharestani P. Effects of evolutionary history on genome wide and phenotypic convergence in Drosophila populations. In review *BMC Genomics*.

Mueller L.D., **Phillips M.A**., Barter T.T., Greenspan Z.G., Rose M.R. Genome-wide mapping of gene-phenotype relationship in experimentally evolved populations. Accepted *Mol. Biol. Evol.*

Phillips M.A. and Rose M.R. Experimental evolution. (2018) Oxford Bibliographies in Evolutionary Biology. New York: Oxford University Press. DOI: 10.1093/OBO/9780199941728-0107

Graves J.L, Hertweck K.L, **Phillips M.A.***, Han M.V., Cabral L.G., Barter T.T., Greer L.F., Burke M.K., Mueller L.D., and Rose M.R. (2017) Genomics of parallel experimental evolution in Drosophila. *Mol. Biol. Evol.* 34:831-842.

Rose M.R., Greer L.F., Phung K.H., Rutledge G.A., **Phillips M.A.**, Anderson C.N.K., and Mueller L.D. (2017). A Hamiltonian Demography of Life History. In R. Shefferson, O. Jones, & R. Salguero-Gómez (Eds.), *The Evolution of Senescence in the Tree of Life* (pp. 40-55). Cambridge: Cambridge University Press.

Phillips M.A., Long A.D., Greenspan Z.S., Greer L.F., Burke M.K., Bryant V., Matsagas K.C., Rizza C.L., Mueller L.D., and Rose M.R. (2016) Genome-wide analysis of long-term evolutionary domestication in Drosophila melanogaster. *Sci. Rep.* 6:39281; doi:10.1038/srep39281

Burke M.K., Barter T.B., Cabral L.G., Kezos J.N., **Phillips M.A**., Rutledge G.A., Phung K.H., Chen R.H., Nguyen H.D., Mueller L.D., and Rose M.R. (2016) Rapid convergence and divergence of life-history in experimentally evolved Drosophila melanogaster. *Evolution* 70:2085-2098

Rose M.R., Cabral L.G., Kezos J.N, **Phillips M.A**., Smith B.L., and Burnham T.C. (2015). Four steps towards the control of aging: Following the example of infectious disease. *Biogerontology* 17:21-31

Rose M.R., Rutledge, G.A., Phung K.H., **Phillips M.A**., Greer L.F., and Mueller L.D. (2014) An evolutionary and genomic approach to the challenges and opportunities for eliminating aging. *Current Aging Science* 7:54-49

Rose M. R, Cabral L.G., **Philips M.A**., Rutledge G.A., Phung K.H., Mueller L.D., and Greer L.F. (2014). The Great Evolutionary Divide: Two Genomic Systems Biologies of Aging. *Interdiscip. Top. Gerontol.* 40:63-73

*Equal contribution with first author

ABSTRACT OF DISSERTATION

Genomics of Adaptation in Drosophila Experimental Evolution

By

Mark A. Phillips

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2018

Professor Michael R. Rose, Chair

The combination of experimental evolution and next-generation sequencing, termed E&R, has emerged as a powerful tool for parsing the genetic foundations of adaptation. Rapid progress has been made in the genomic analysis of adaptation in asexual microbial populations. Such adaptation chiefly features "selective sweeps" in which new advantageous mutations arise at low frequency and proceed to fixation. If this is representative of adaptation in outbreeding sexual populations, then we can understand the genetic basis of long-term adaptation using microbial findings alone. However, results from early Drosophila E&R studies have failed to support this notion; instead adaptation appears to be primarily fueled by standing genetic variation. But since these studies were limited in duration, little is known about the long-term dynamics of experimental evolution in Drosophila. My work addresses this issue by comparing patterns of genomic variation and differentiation in the dozens of experimentally evolved populations *D. melanogaster* maintained in the Rose Lab at UC, Irvine. This experimental radiation dates back to the 1970s, and features groups of replicate populations that have been subjected to various selection regimes for dozens to hundreds of generations. Using a group of populations

maintained on a laboratory domestication regime for ~1000 generations, I find that adaptation in sexual E&R is indeed characterized by a lack of fixation and populations actually harbor more genetic variation than conventional population genetic theory predicts (Chapter 1). Work comparing newly-derived and long-standing populations subjected to the same selections pressures led me to conclude that adaptation can be fast and highly repeatable at the level of genotypes and phenotypes due to standing genetic variation, and it is not dependent on the appearance beneficial *de novo* mutation (Chapter 2). Studying populations that were previously subjected to intense selection for desiccation resistance and their controls led me to conclude that evolutionary history does not have major long-lasting impacts in Drosophila E&R studies (Chapter 3). Lastly, using findings from two starvation-selection experiments performed at different population sizes, I show that population size is an important experimental parameter to maximize in studies aimed at deciphering the genetic architecture of complex phenotypes (Chapter 4).

INTRODUCTION

Broadly, the field of evolutionary biology is primarily concerned with studying and characterizing the evolutionary processes responsible for the patterns of diversity found in nature. Experimental evolution is an approach that seeks to characterize the microevolutionary dynamics underlying these processes in real time by studying populations across multiple generations as they evolve in response to deliberately imposed conditions (Rose and Garland 2009). The earliest well-documented study in the field dates back to 1878 when William Dallinger cultured unicellular organisms in an incubator for several years while gradually increasing the temperature from 60 ^{o}F to 158 ^{o}F (Dallinger 1887). His findings provided direct evidence of Darwinian adaptation by showing that while early cultures could not survive at 158 ^{o}F, cultures present at the end of the experiment were largely unaffected by such high temperature.

In the decades since Dallinger's pioneering work, experimental evolution has been chiefly employed to further our understanding of evolutionary processes associated with microbial ecology (Chao et al. 1977; Chao and Levin 1981; Crill et al. 2000; Kaltz and Bell 2002; Turner and Chao 1999), life history evolution (Mueller and Ayala 1981; Luckinbill et al. 1984; Rose 1984; Service et al. 1988; Chippendale et al. 1997), and evolutionary physiology (Graves et al. 1992; Rose et al. 1992; Gibbs et al. 1997; Djawdan et al. 1998; Swallow et al. 1998; Swallow et al. 1999; Roff et al. 1999; Roff and Fairbairn 2002). However, with the advent of next-generation sequencing, experimental evolution has emerged as a powerful tool for parsing the genetic foundations of adaptation (Long et al. 2015; Schlötterer et al. 2015). The use of this approach, termed evolve and resequence

("E&R") by Turner et al. (2011), to generate insights into the genomics of adaptation is the primary focus of my thesis.

The goal of my thesis is to generate insights into the genomics of adaptation in sexually reproducing populations. Unlike early E&R studies featuring strictly asexual populations where adaptation was found to be characterized by the fixation of beneficial mutations and hard selective sweeps (Barrick et al. 2009; Tenaillon et al. 2012), early E&R studies using sexual populations, typically *D. melanogaster*, concluded that adaptation was primarily driven by standing genetic variation and characterized by a lack of genetic fixation (Teotónio et al. 2009; Burke et al. 2010; Turner et al. 2011; Orozco-terWengel et al. 2012, Tobler et al. 2014). However, these early E&R studies in sexual populations lacked the replication and generations under selection seen in their asexual counterparts. As such, there were certain limits to what researchers could argue based on the strength of their experiments. For instance, is adaptation in sexual populations really characterized by soft sweeps and a lack of genetic fixation (Burke 2012)? Or do alleles simply not have time to fix given the duration of most E&R studies with sexual populations? Here I seek to address such concerns using the dozens of experimentally evolved populations *Drosophila melanogaster* maintained in the Rose Lab at the University of California, Irvine

The Rose Lab's experimental radiation dates back to the 1970s, and features groups of replicate populations that have been subjected to various selection regimes for dozens to hundreds of generations. Using a group of population subjected to the same laboratory domestication regime for nearly ~1000 generations, I characterize the long-response to selection (Chapter 1). Following this, I use newly-derived (dozens of generations under selection) and long-standing (hundreds of generations under selection) populations to

compare and contrast long and short-term responses to selection (Chapter 2). Next, I examine the impact of evolutionary history in sexual E&R studies using populations that were previously subjected to intense selection for desiccation resistance (Chapter 3). Lastly, I empirically test how population size impacts the power to detect causal variants in sexual E&R studies using findings from two starvation selection experiments: one with moderately outbred populations, and the other with populations where population size was deliberately compressed (Chapter 4).

References

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, et al. Genome evolution and adaptation in long-term experiment with *Escherichia coli*. Nature. 2009;461:1243-1274.

Burke MK. How does adaptation sweep through the genome? Insights from long-term selection experiments. Proc R Soc. 2012;279:5029-5038.

Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. Genome-wide analysis of a long-term evolution experiment with Drosophila. Nature. 2010;467:587–590.

Chao L, Levin BR, Stewart FM. A complex community in a simple habitat: an experimental study with bacteria and phage. Ecology. 1997;58:369-378.

Chao L, Levin BR. Structured habitats and the evolution of anticompetitor toxins in bacteria. Proc Natl Acad SCI U.S.A. 1981;78:6324-6328.

Chippindale AK, Alipaz JA, Chen H-W, Rose MR. Experimental evolution of accelerated development in Drosophila. 1. Development spped and larval survivial. Evolution. 1997;51:1536-1551.

Crill WD, Wichman HA, Bull JJ. Evolutionary reverasals during viral adaptation to alternatings hosts. Genetics. 200;154:27-37.

Dallinger W. The president's address. Journal of the Royal Microscopical Society. 1887;10:184-199.

Djawdan M, Chippindale AK, Rose MR, Bradley TJ. Metabolic reserves and evolved stress resistance in Drosophila melanogaster. Physiol Zool. 1998;71:584-594.

Graves JL, Toolson EC, Jeong C, Vu Ln, Rose MR. Desiccation, flight, glycogen, and posted senescence in *Drosophila melanogaster*. Physiol Zool. 1992;65:268-286.

Gibbs, A.G., Chippindale, A.K. & Rose, M.R. Physiological mechanisms of evolved desiccation resistance in Drosophila melanogaster. J Exp Biol. 1997;200:1821-1832.

Kaltz O, Bell G. The ecology and genetics of fitness in Chlamydomonas. XII. Repeated sexual episodes increase rates of adaptation to novel environments. Evolution. 2002;56:1743-1753.

Long AD, Liti G, Luptak A, Tenaillon O. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. Nat. Rev. Genet. 2015;16:567–582.

Luckinbill LS, Arking R, Clare MJ, Cirocco WC, Buck SA. Selection for delayed senescence in *Drosophila melanogaster*. Evolution. 1984;38:996-1003.

Orozco-ter Wengel P, Kapun M, Nolte V, Kofler R, Flatt T, Schlötterer C. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected traits. Mol Ecol. 2012;21:4931–4941.

Roff DA, Tucker J, Stirling G, Fairbairn DJ. The evolution of threshold traits: effects of selection on fecundity and correlated response in wing dimorphism in the sand cricket. J Evol Biol. 1999;12:535-546

Roff DA, Fairbairn DJ. Laboratory evolution of the migratory polymorphism in the sand cricket: combining physiology with quantitative genetics. Physiol Biochem Zool. 2007;80:358-369.

Rose MR. Laboratory evolution of postponed senescence in Drosophila melanogaster. Evolution. 1984;38:1004–1010.

Rose MR, Vu LN, Park, SU, Graves JL. Selection for stress resistance increases longevity in Drosophila melanogaster. Exp Gerontol. 1992;27:241-250.

Rose MR, Garland Jr. T. Experimental evolution. Berkley: University of California Press; 2009.

Schlötterer C, Kofler R, Versace E, Tobler R, Franseen SU. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. Heredity 2015;114:331–440.

Service, P.M., E.W. Hutchinson, and M.R. Rose. 1988. Multiple genetic mechanisms for the evolution of senescence in Drosophila melanogaster. Evolution 42:708-716.

Swallow JG, Garland Jr. T, Carter PA, Zhan W, Sieck GC. Effects of voluntary activity and genetic selection on aerobic capacity in house mice (*Mus domesticus*). J Appl Physiol. 1998;84:69-76

Swallow JG, Koteja P, Carter PA, and Garland Jr. T. Artificial selection for increased wheel-running activity in house mice results in decreased body mass at maturity. J Exp Biol. 1999;202:2513-2520.

Tenaillon O, Rodriguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, et al. The molecular diversity of adaptive convergence. Science. 2012;335:457-461.

Teotónio H, Chelo IM, Bradiá M, Rose MR, Long AD. Experimental evolution reveals natural selection on standing genetic variation. Nat Genet. 2009;41:251–257.

Tobler R, Franssen SU, Kofler R, Orozco-ter Wengel P, Nolte V, Hermisson J, Schlötterer C. Massive habitat-specific genomic response in D. melanogaster populations during experimental evolution in hot and cold environments. Mol Biol Evol. 2014;31:364–375.

Turner PE, Chao L. Prisoner's Dilemma in an RNA Virus. Nature. 1999;398:441-443

Mueller LD, Ayala FJ. Trade-off between r-selection and k-selection in Drosophila populations. Proc Natl Acad SCI U.S.A. 1981;78:1303-1305.

Turner TL, Steward AD, Fields AT, Rice WR, Tarone AM. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in Drosophila melanogaster. PLoS Genet.2011; 7:e10001336.

CHAPTER 1

Genome-wide Analysis of Long-term Evolutionary Domestication in *Drosophila melanogaster*

ABSTRACT

Experimental evolutionary genomics now allows biologists to test fundamental theories concerning the genetic basis of adaptation. We have conducted one of the longest laboratory evolution experiments with any sexually-reproducing metazoan, *Drosophila melanogaster*. We used next-generation resequencing data from this experiment to examine genome-wide patterns of genetic variation over an evolutionary time-scale that approaches 1,000 generations. We also compared measures of variation within and differentiation between our populations to simulations based on a variety of evolutionary scenarios. Our analysis yielded no clear evidence of hard selective sweeps, whereby natural selection acts to increase the frequency of a newly-arising mutation in a population until it becomes fixed. We do find evidence for selection acting on standing genetic variation, as independent replicate populations exhibit similar population-genetic dynamics, without obvious fixation of candidate alleles under selection. A hidden-Markov model test for selection also found widespread evidence for selection. We found more genetic variation genome-wide, and less differentiation between replicate populations genome-wide, than arose in any of our simulated evolutionary scenarios.

INTRODUCTION

The genetic basis of adaptation has historically been a major point of contention among evolutionary biologists. In recent years, combining genome-wide sequencing and experimental evolution has emerged as a powerful method for parsing the genetic underpinnings of adaptation (Burke et al. 2010; Turner et al. 2011; Tenaillon et al. 2012; Schlötterer et al. 2014). Termed the "evolve and resequence" (E&R) approach, these experiments involve sequencing laboratory populations that have been exposed to clearly defined selective pressures in the hopes of making direct connections between patterns of genotypic and phenotypic change (Long et al. 2015). In the case of largely or wholly asexual populations, genome-wide sequencing has been performed on clones derived from single individuals after many generations of adaption to novel conditions (Tenaillon et al. 2012; Barrick et al. 2009). Since such asexual populations are expected to undergo successive rounds of selective sweeps that purge genetic variation genome-wide (Burke 2012), this is a reasonable approach to the characterization of chiefly clonal evolutionary processes.

In the case of outbreeding sexual species, such as *Drosophila melanogaster,* the more common sequencing strategy in E&R experiments has been to pool multiple individuals within or across evolving replicated laboratory populations(Burke et al. 2010; Orozco-terWengel et al. 2012). This is often referred to as the "pool-seq" approach (Schlötterer et al. 2014). Results from E&R experiments in outbreeding sexual species using this pool-seq approach have revealed abundant genetic variation genome-wide, and suggest that adaptation is primarily due to selection on standing genetic variation (Schlötterer et al. 2014). However, as most of those studies typically feature populations with relatively small effective population sizes that have been subjected to only a few dozen generations of

selection, there is a limit to the conclusions that can be drawn from them regarding the relative importance of selective sweeps versus selection acting on standing genetic variation (Burke 2012; Hermisson and Pennings 2005; Pennings and Hermisson 2006; Przeworski et al. 2005), particularly as selective sweeps are likely to take much longer than the few dozens of generations commonly used in selection experiments with metazoa, as opposed to experimental evolution with microbes (Barrick et al. 2009).

Genome-wide sequencing of genetic variation present in experimentally evolving sexual populations after many generations of selection remains of interest as a method for addressing the relative importance of selective sweeps, particularly from the standpoint of alleles being driven to fixation. Hitchhiking effects arising from successive selective sweeps are not expected to purge genetic variation genome-wide in sexual populations immediately (Burke 2012; Maynard Smith and Haigh 1974). But sufficiently many such selective sweeps acting in conjunction with reductions in heterozygosity resulting from background selection and genetic drift conceivably could progressively purge genetic variation, given the moderately small population sizes commonly used in experimental evolution with sexual species (Mueller et al. 2013). The data analyzed by Burke et al. (2010) do not show a widespread purging of genetic variation in populations that had evolved in the lab for some decades. However, save for a single replicate sequenced individually, this study featured data generated from pooling across replicates, which could have potentially masked genetic fixation in individual replicate populations. This raises the question whether genome-wide sequencing of independently evolving, replicate, sexual populations that have been maintained in the lab for hundreds of generations will indeed

show a pattern of generally purged genetic variation when these populations are resequenced separately.

Here we show that such long-evolved moderate-sized sexual populations do not exhibit the general lack of genetic variation that is characteristic of long-evolved clonal populations. Instead, after sequencing five independent replicate populations sharing a common selection regime, and more than 900 generations of sustained directional selection, we find widespread maintenance of genetic variation genome-wide. While it could be argued we still do not have sufficient generations of selection or population sizes large enough, our study features the best data collected from laboratory evolution to date with respect to the long-term evolution of patterns of genetic variation in sexually reproducing populations of multicellular eukaryotes. Furthermore, we compare measures of variation within and differentiation between our populations to simulated data from a number of evolutionary scenarios. We consistently find that there is more variation maintained in our populations, and less differentiation between replicate populations, than is found in any of the evolutionary scenarios we simulated. Lastly, we look at patterns of genetic variation and the frequency distribution of genetic variation to test for selective sweeps.

MATERIALS AND METHODS

Experimental material

The novel experimental material analyzed here is pooled genomic DNA obtained from each of the five "B" populations maintained in the Rose laboratory (Rose 1984, Rose et al. 2003). These five B populations were founded in February 1980 from a single generation of the "IV" stock studied by Rose and Charlesworth (1980 and 1981), which was

in turn founded in August 1975 from a sample of 200 wild-caught *Drosophila melanogaster* females obtained by Phillip Ives from his long-studied South Amherst, Massachusetts endemic population (Ives 1970). From their first founding, IV and B populations have been cultured using 14-day discrete generations, with mixing of flies across all culture vessels *within* replicate populations, at temperatures of 23-25 degrees Celsius. Effective populations size (N_e) are around ~1000 based on calculations from demographic data (Mueller et al. 2013).

The B populations were pooled and sampled for sequencing at generation 785 from their founding in 1980, in March 2010, though this was after a total of 915 generations of lab domestication since Ives supplied the wild-caught females for founding the laboratory IV stock. From August of 1975 to June of 1981, the IV and B populations were cultured using corn meal based medium in 16 glass milk pint bottles per population, with 12L:12D light exposure. From June 1981 to March 2010, the B populations sequenced here were cultured in 40 shell vials at densities of 60-80 eggs per vial, yielding 50-75 adults per vial. This equates to around a minimum census size of 2000 each generation. During this period, these B populations were cultured using banana-molasses medium with 24L:0D light exposure (Rose 1984).

The pooled genomic DNA was obtained by isolating 250 female flies from each of the five population replicates of the B flies, with harvesting ten days after the pupal stage, using the Gentra Puregene Cell Kit (Qiagen, Valencia, CA) according to standard protocol, after maceration of the fly tissues using Dounce Tissue Grinders (Daigger, Vernon Hill, IL). Genomic DNA concentrations and purities of the samples were assayed by DNA spectrophotometer. Size distributions were visualized by low agarose gel electrophoresis

with DNA size markers. Genomic DNAs were stored at –20° C before shipment to Beckman Coulter (Brea, CA) for Illumina paired end sequencing. Reads were 76 bp in length. The fastq files used in our analyses are available through NCBI SRA (BioProject ID: PRJNA350701).

Mapping of reads

We mapped reads with BWA (version 0.7.8) (Li and Durbin 2009) against the *D. melanogaster* reference genome (version 5.55) using bwa mem (Li 2013) with default settings. We filtered and sorted the resulting SAM files for reads mapped in proper pairs with a minimum mapping quality of 20 using and converted them to the BAM using the view and sort commands in SAMtools (Li et al. 2009). These files were then converted to mpileup format once again using SAMtools. Using the PoPoolation2 (Kofler et al. 2011a) software package, these files were converted to "synchronized" files, which is a format that allele counts for all bases in the reference genome and for all populations being analyzed. Lastly, we used RepeatMasker 4.0.3 (http://www.repeatmasker.org) (Smith et al. 2013) to create a gff file to mask simple sequence repeats and transposable elements of the *D. melanogaster* genome version 5.55.

A table with major and minor allele counts for each SNP in each population was then generated from this synchronized file. SNPs where discarded if coverage in any of the populations was less than 20X or greater than 150X. We also required a minimum minor allele frequency of 2% across all five populations. Based on these parameters, ~1.2 million SNPs were identified across the major chromosome arms. The average coverage at each called SNP was 62X, 65X, 57X, 66X, and 69X in B_1 through B_5 respectively.

Characterizing genetic variation

Local depressions in genetic variation are considered one of the primary means of detecting selective sweeps from population level data (Oleksyk et al. 2010). We calculated and plotted heterozygosity across the five major chromosome arms to see if we could find any such evidence for such depressions in our real and simulated data sets. Heterozygosities were calculated over 100kb non-overlapping windows directly from the major and minor counts in our SNP table. Watterson theta (Θ) was also calculated using *PoPoolation* (Kofler et al. 2011b), where the details of these calculations can be found. Mpileups were first made for each population using the bam files mentioned above. We then subsampled (without replacement) to a uniform coverage level of 30X across all populations, as these calculations can be sensitive to coverage variation. Estimates of genetic parameters were then calculated over 100kb non-overlapping windows across the major chromosome arms. For a SNP to be called at a given position, we required a minimum minor allele count of 2. Lastly, sufficient coverage (30X) was required for at least 60% of our 100kb windows for estimates to be generated.

F_{ST} estimates

F_{ST} estimates were obtained using the formula: $F_{ST} = \frac{Ht-Hs}{Ht}$ where H_T is heterozygosity based on total population allele frequencies, and H_S is the average subpopulation heterozygosity in each of the B populations (Hedrick 2009). F_{ST} estimates were made at every polymorphic site in the data set. This was done to quantify the levels of differentiation between our five B populations, as well as between replicate populations in our simulated scenarios. F_{ST} estimates were calculated for the B populations along

chromosome 3R at every polymorphic site. This was also done for the simulated data, and once again all polymorphic positions along 3R not present in the SNP table created from our real data set were discarded.

We also used the formula: $F_{ST} = 1 - (1 - \frac{1}{2N})^t$ to generate a predicted F_{ST} value where N is the effective population size of each subpopulation and t is the time since divergence from their ancestral population (Hartl and Clark 1997). This model assumes that population diverge randomly over time and that there is no migration. In the event that there is some level of migration between our populations, we used the formula: $F_{ST} = \frac{1}{4Nm+1}$ where m is equal to the migration rate and the quantity Nm is equal to the number of migrants per generation (Dobzhansky and Wright 1941; Hartl and Clark 1997). This model assumes no mutation and that the migration rate is small. As with the previous model, it corresponds to a scenario where a single population is split into subpopulations at some point and diverges randomly over time. But in this scenario, migration has placed a limit on how much these subpopulations can diverge and assumes that the populations have reached this limit and are at equilibrium. It is worth noting that this assumption of equilibrium might not be met in our system.

Simulations

To perform our first set of simulations we used MimicrEE (https://sourceforge.net/projects/mimicree/) (Kofler 2015), a forward simulation specifically designed to mimic experimental evolution. It simulates populations of diploid individuals where genomes are provided as haplotypes with two haplotypes constituting a

diploid genome. There are no changes in the demography once the initial population file is submitted and a list of selected loci may be provided.

For each selected locus, the selection coefficient (s), the dominance coefficient (h), and the nucleotide of the nonselected allele are provided (w_{11}). The fitness of the heterozygous and homozygous individuals is given by: $\omega_{11} = 1, \omega_{12} = 1 + hs$, and $\omega_{22} = 1 + s$ (Gillespie 2010). The simulation assumes multiplicative fitness when several selected loci are specified. No *de novo* mutations are considered, as its purpose is to simulate scenarios where adaptation results from selection on standing genetic variation. The simulated populations have non-overlapping generations and all individuals are hermaphrodites (though selfing is excluded). At each generation, matings are performed, where mating success (number of offspring) scales linearly with fitness, until the total number of offspring in the population equals the targeted population size (fecundity selection). Each parent contributes a single gamete to the offspring. Crossing-over events are introduced according to a user-specified recombination rate.

To generate our starting haplotypes, we started with 105 individuals from the Drosophila Genetics Reference Panel (DGRP) (Mackay et al. 2012). We only used positions along chromosome 3R and only sites that were polymorphic in the B populations. In total, there were 238,291 polymorphic sites after this filtering. From these 105 haplotypes, we randomly sampled with replacement to 1000 to achieve our desired population size. Recombination rates were specified for 100kb windows and were obtained from the *D. melanogaster* recombination rate calculator v2.2 (Fiston-Lavier et al. 2010). As recombination does not occurs in male Drosophila, the empirically estimated female recombination rate was divided by two for the simulations.

We first performed neutral simulations, featuring only drift and recombination, to establish a baseline. We then ran simulations across a variety of evolutionary scenarios that involved different numbers of selected loci. Our goal was to see which, if any, of these scenarios would produce the sorts of patterns we observer in our real data set. For each scenario, we simulated five populations for 800 generations. This was then done 100 times for each scenario. In our selection scenarios, we simulated populations with 5, 10, or 20 randomly distributed beneficial sites. For one set of simulations, the reference nucleotide (A_1) was defined as the beneficial allele in each case and was also defined as dominant (h=0). For another set, the A_2 allele was defined as beneficial and dominant (h=1). And for a final set, all selected loci were codominant (h=.5). For each set, we simulated scenarios with selection coefficients ranging from 0.03 (low) to 0.1 (high) (Table 1.1). As we increased the number of selected sites, we reduced the selection coefficients to generate scenarios with either few sites of large effect or many sites of small effect. Lastly, we simulated scenarios featuring sites with overdominance to see if extensive balancing selection could be behind the patterns we observe in our data set. In these scenarios, we simulated populations with 20 or 30 randomly distributed sites with overdominance. And once again, we simulated scenarios with a range of selection coefficients (Table 1.1).

From each set of 5 simulated populations under each scenario, we calculated average heterozygosity and average F_{ST} across all polymorphic sites to compare to values observed in the B populations. We also looked at heterozygosity and F_{ST} over 50kb windows, and calculated the variance between windows as a means of comparing spatial variation in heterozygosity and F_{ST} to what we observe in the B populations.

Simulations with migration

Given how long our populations have been maintained in the lab, it is easy to imagine that there may have been some instances of migration due to accidental cross-contamination. Thus, in addition to the selection scenarios mentioned above, we performed simulations featuring migration. These simulations also feature optimizing selection, as opposed to our other simulations featuring directional and over dominant selection. We once again simulated an evolve-and-resequence experiment for a 63 cM long *D. melanogaster* chromosome 3R. An R program was created to simulate an initial population of F founder chromosomes expanded and used to found five populations that were then evolved for G generations at a population size of N gametes per population with M migrant gametes in the meta-population per generation. The simulation was accomplished by tracking founder segments and recombination breakpoints over time. So, the N gametes used to found a subpopulation initially consist of a random sample of size N drawn from the numbers 1 through F with replacement. Then each generation to create N new gametes (n=1...N) we draw two gametes with replacement from the previous generation and create a recombination breakpoint at position r = unif(0,1), if $r < 0.63$ (to simulate a chromosome of 63 cM) and n modulus 2 equal zero (since recombination only takes place in females). Recombinant chromosomes are represented as a pair of vectors: a founder state vector, and a recombination breakpoint vector. So, for example, the n^{th} gamete in a population might be $\boldsymbol{S_n}$={3,17,31,3} and $\boldsymbol{B_n}$={0.20,0.25,0.60}, indicating that gamete has material from founder #3 from 0 to 20cM, founder 17 from 20 to 25 cM, etc. This sampling scheme models drift and recombination in a Wright-Fisher population.

Next we added selection to the simulation. We simulated Q evenly spaced quantitative trait loci (QTL) on the chromosome, with a vector of effect sizes and starting allele frequencies (***E*** and ***F***), with QTL states randomly assigned via binomial sampling given F_q for each locus. The resulting quantitative trait has a heritability due to the QTL on 3R of 12%, and an additional polygenic heritability of 38% due to the other chromosomes, and a total phenotypic variance of one. We then model a gamete's phenotype as the sum of its effect sizes plus a random Gaussian deviate representing a polygenic component, plus a second random Gaussian deviate representing an environmental deviation, plus the current polygenic mean of the population. The fitness of each gamete is a standard Gaussian fitness function proportional to w=(pheno – NewOptimum)^2/(2*VarianceFitness), normalized to total fitness. Under this selection scheme each generation gametes are resampled proportional to w, resulting in allele frequency changes at the underlying QTL, and resulting changes in the mean phenotype due to those QTL. Furthermore, each generation the polygenic mean (the mean phenotype due to chromosomes other than 3R) changes according to the Breeder's equation(Falconer and Mackay 1996).That is we partition the trait variance into the variance due to tracked loci (each having $V_{a,i}=2p_iq_ia_i^2$), a polygenic component with Gaussian variance $V_{a;poly}$, and environmental variation (V_e). We held V_e /V_t constant, but allowed the ratio of V_a to $V_{a;poly}$ to vary. Each generation an individual's phenotypic value is the sum of allelic effects due to tracked loci, a Gaussian deviation due to the polygenic component, and a Gaussian environmental deviate. Those phenotypic values determines a vector of length N consisting of each individual's average fitness based on the Gaussian fitness function, with N individual's chosen with replacement from that vector to create the next generation, with

the probability of being chosen proportional to average fitness. Between generations, the mean of the population then shifts due to changes in allele frequencies at ***both*** the tracked loci ***and*** untracked loci, as predicted by the Breeder's equation: $h^2_{poly} * S$. In this context, h^2_{poly} is $V_{a;poly}/(V_{a;poly} + V_e)$ and S is the observed selective differential (i.e., the mean phenotype of individuals chosen to reproduce minus the mean of the population). By modeling adaptation in this fashion, the population approaches the new optimum due to genetic changes at both tracked and untracked loci. This model thus accommodates adaptation at both the explicitly modeled chromosome arm and the remainder of the genome.

Under our model, the overall rate at which the mean phenotype changes in the population is controlled by the distance to the new phenotypic optimum and the variance in fitness, which are set to 15 and 12, respectively. Parameters scaling results in V_e=0.5 and V_p=1.0 at generation zero, with a new optimum that is 15 phenotype standard deviations away from the population mean, a shift in optimum that was chosen to match experimental evolution experiments in Drosophila (vid. Teotonio and Rose 2000). The simulation was set up so that a population can reach a new optimum phenotype before all underlying QTL are fixed.

Each generation we simulated migration by randomly taking M pairs of gametes in the 5*N set being tracked and replacing the first member of the pair with the second. This corresponds to a one-way island model of migration. This means that if the entire population is 5*N gametes (where 5 is the number of populations), each generation k gametes are chosen and they essentially overwrite K other gametes. That is to say, "one-way" means gametes are not exchanged between populations and "island" means migration

is equally likely between any two demes in any direction. We iterated this entire process for G generations to obtain a final set of 5*N gametes. Given the relatively short time scale of the experimental evolution experiment, and the relatively modest number of gametes, this simulation is fairly efficient in R on a desktop computer. At the end of the simulation we took a set of 100 (F) 3R chromosomes from the DGRP (Mackay et al. 2012), and for each polymorphic position in the real DGRP data we calculated an allele frequency at that site based our simulated populations. This was done because, as with our previous simulations, our starting haplotypes were based on lines from the DGRP. Performing this calculation required a function that maps physical position in bp to cM, and then simply iterates over the founder states at each of the 5*N gametes for each SNP. Despite the fact that poolseq estimates allele frequency in the population based on a finite sample of gametes, with the accuracy of that estimate a function of coverage depth and number of gametes sampled for the Illumina library, we used the exact allele frequency estimates calculated using the method described above in our downstream calculations. Since libraries are made using a large number of individuals (>200) and the per site coverage approaches 60X, this simplification is likely acceptable.

Using this framework, we simulated a number of evolutionary scenarios that involved varying the following: migration rates (M), number of selected QTL, effect sizes of selected QTL, and starting frequencies of selected QTL. Scenarios were simulated using groups of five populations to mimic our observed fly populations. Each scenario was simulated 300 times. All simulations ran for 800 generations (G) and all simulated populations consisted of 2000 gametes (N). We simulated scenarios with 0, 1 or 5 migration events per generation (M). We ran simulations with 0 (i.e., a control with only

random genetic drift), 3, 10 and 20 selected QTL. Selected QTL were evenly distributed across the chromosome arm. We also looked at the effects of the starting frequency of selected QTL (F) by running simulations where all selected alleles started at either 0.05 or 0.5. As we increased the number of QTL, we reduced their effect sizes (E) so that the sum of squared effect sizes was held constant. This was done to prevent changes in the heritability of the character. For simulations with 3 QTL, effect sizes were 1, 2, and 2. For simulations with 10 QTL, they were 1, 1, 1, 1, 1, 1, 1, 1, 0.71, and 0.71. And lastly, they were 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.71, 0.5, 0.5, 0.5, and 0.5 for simulations with 20 QTL. [See Table 1.2 for all simulated scenarios]

Once again, we calculated average heterozygosity and average F_{ST} across all polymorphic sites to compare with the values observed in the B populations. We then again looked at heterozygosity and F_{ST} over 100kb windows and calculated the variance between windows as a means of comparing spatial variation in heterozygosity and F_{ST} to what we observe in the B populations.

Selection detection

To test for footprints of selection across the genome, we relied on a hidden Markov Model developed with the intention to detect sweeps in pooled sequence data, developed by Boitard et al. (2012), implemented in the Pool-HMM software package (Boitard et al. 2013). Their method involves estimating the allele frequency spectrum (AFS) across genomic regions and detecting distortions relative to the background AFS, which are expected to occur in regions subject to selection. Though it is worth noting that while this method was developed primarily to detect selective sweeps, it could be instead identifying signatures of any process that produces similar perturbations of the AFS.

Mpileups for individual populations were used as test inputs, since Pool-HMM can only process data from one population at a time. Scans were performed along each of the major chromosome arms using the following parameters: -n 500 (number of chromosomes in each pool), -pred (predicts the hidden state, "Neutral" (far away from sweep site) , "Selected" (close to sweep site) , or "Intermediate" (between Neutral and Selected) of each SNP), -C 150 (maximum coverage allowed for sites used in this analysis), -r 5 (where 1/r is the proportion of sites that are used to estimate the background AFS), -theta Θ (average Θ for each population was approximately 0.003 based on estimates from PoPoolation, and increasing or decreasing window size did not affect this result), and -k 10^{-10}. The –k parameter is the per site transition probability *q* between neutral and selected states, which is an important tuning parameter for the hidden Markov model underlying this test. As *q* increases, less evidence is required for a transition to selection and more sweep candidates should be detected. We also ran tests under more ($q = 10^{-11}$) and less stringent ($q = 10^{-9}$) conditions, which only led to slight differences in the number of footprints detected (Table 1.3). A confidence index was calculated for each selective sweep window detected using this method as $-\log_{10}(1-p)$, where *p* is the maximum of the posterior probability of hidden state "Selection" within the window.

We applied this Pool-HMM test to results from our neutral simulations using the settings listed above, as a means of evaluating our false positive rate. For such tests of simulated data, we used 100 kb regions extracted from different runs of our neutral simulations. Essentially, we converted output from the simulations to mpileup files. Sequence and coverage variation were introduced based on what was found in the actual 3R sequences from the B populations. In addition to the 100 kb regions just mentioned, we

ran tests for selection on results for the entirety of 3R from 15 simulated populations each taken from a different evolutionary scenario. We did this for populations taken from each of following scenarios: neutral evolution with migration rates M=0, 5 and 20; 3 selected QTL's with low starting frequencies with M=0, 5 and 20; 3 selected QTL's with high starting frequencies with M=0, 5 and 20; 10 selected QTL's with low starting frequencies with M=0, 5 and 20; 10 selected QTL's with high starting frequencies with M=0, 5 and 20.

RESULTS

Genetic variation

Plotting measures of genetic variation, heterozygosity and Watterson theta (Θ), across 100 kb non-overlapping windows reveals depressions in genetic variation across all major chromosomes arms in the B populations (Figure. 1.1, see Figure S1.1 for mean heterozygosity and Θ across all 5 populations). This pattern is robust to both increased (150 kb) and decreased (30 and 50 kb) window size (Figure S1.2-S1.4). Many depressions in heterozygosity are consistent across the 5 replicate populations, which may be indicative of selection on standing variation. In general, there is a great deal of similarity in patterns of heterozygosity across replicates (Figure S1.5 shows pair-wise comparisons between all replicates). As in Burke et al. (2010), we find no regions where genetic variation has been completely expunged in an unambiguous manner. However, there are regions that show very low levels of heterozygosity (0.2) and theta ($\Theta < 0.001$) consistently across replicates (Table S1.1 and Table S1.2). In addition, the vast majority of these regions are located in chromosome X. Such regions may arise from incomplete selective sweeps or balanced selective equilibria that are close to fixation boundaries. Nonetheless, we have not found cases that conform to the pattern of heterozygosity expected with hard selective sweeps

proceeding all the way to fixation (Burke 2012), despite almost 1,000 generations of sustained selection.

We used average heterozygosity as our primary measure of variation when comparing our quantitative results from the actual data obtained from the B populations to the corresponding results obtained from the different evolutionary scenarios we simulated. Average heterozygosity in the *starting* base population used in our simulations (see Materials and Methods) was 0.32. Average heterozygosity across 3R was lower in the five B-population replicate data at .28, .29, .28, .27, and .27, respectively. We find lower levels of heterozygosity after 800 generations in all of our evolutionary scenarios (Table 1.1 and Table 1.2). Spatial variance in heterozygosity along the chromosome arm, based on calculations from 50kb windows, were as follows for the five B replicates: .21, .24, .24, .37, and .37.

In our simulations performed using MimicrEE, we found that the addition of selection does result in greater losses in heterozygosity than genetic drift alone, as expected (Table 1.1). In our scenarios where the A_1 allele is dominant and beneficial at each selected site, we found that increasing the number of selected sites produced greater reductions in heterozygosity (Table 1.1). This was also true when the A_2 allele was dominant or if we used a selection model featuring codominance. In terms of spatial variance in heterozygosity, we find the highest levels in scenarios featuring strong selection at 3 (0.055 to 0.060) sites or weaker selection at 20 sites (Table 1.1). The lowest levels of spatial variance, which were comparable to the higher end of what was observed in the B populations, were found in scenarios with selection at 10 sites (0.0035 to 0.0040).

In our scenarios with overdominance, we found that increasing the number of sites under selection and increasing selection coefficients both produce greater reductions in heterozygosity near locations undergoing selection. Heterozygosity was maintained at the selected sites themselves and selected alleles approach predicted equilibrium frequencies based on our settings, but heterozygosity was nonetheless reduced in surrounding regions. We also found spatial variance in heterozygosity to be higher than what we typically observed in the B populations, with the exception of scenarios featuring 30 sites with small selection coefficients (.34) (Table 1.1).

In our simulations with migration, we once again found that the addition of selection results in greater reductions in heterozygosity overall (Table 1.2). Increasing the number of selected QTLs again produced a greater decrease in heterozygosity. Increasing migrations rates resulted in more variation being maintained, as did increasing the starting frequencies of selected QTLs; however, the simulated levels of genetic variation never achieved the levels seen in the actual B populations. In terms of genome-wide variance in heterozygosity, allowing migration and increasing the starting frequencies of selected QTLs produced results closer to those observed in the B populations (Table 1.2).

F_{ST}

Mean F_{ST} across the 5 B populations was 0.08 across all chromosome arms, including 3R individually. This value is far lower than what we would predict assuming no migration using the formula $F_{ST} = 1 - (1 - \frac{1}{2N})^t$, which predicts F_{ST} should be around 0.33 assuming N = 1,000 and t ~ 800 generations. Substituting our observed F_{ST} into this equation and instead solving for N suggests that in order to produce an F_{ST} estimate of 0.08, assuming no migration and random divergence, we would have to have an effective

population size of around 4,700. To assesses how much migration would be required to produce this result, assuming random divergence and populations at equilibrium, we used the formula $F_{ST} = \frac{1}{4Nm+1}$. Solving for Nm, the number of migrants per generations, suggests our observed F_{ST} could be produced if there were 2.88 successful migrants per generation, each and every generation, if the assumptions of this model are met.

Our observed F_{ST} was also far lower than anything produced in the different evolutionary scenarios we simulated (Table 1.1 and Table 1.2). In our simulations performed using MimicrEE, we found that scenarios where A_1 allele was dominant and beneficial at each selected site all produced greater F_{ST} estimates than were produced by drift alone. Scenarios with overdominance and low selection coefficients produced modest reductions in mean F_{ST}, but never to the level observed in the B populations. This effect was lost when selection coefficients were increased, once again giving mean F_{ST} estimates greater than those produced by drift alone. All simulations also produced much greater variance in genome-wide F_{ST}, 0.0028 at the lowest, than we observe in the B populations (.0004), for the single chromosome arm.

In our simulations featuring migration, increasing the number of QTL under selection and/or altering the starting frequencies of favored genotypes both failed to have any appreciable effects on mean F_{ST} (Table 1.2). Increasing the migration rates did reduce F_{ST} as expected. However, even with migration rates as great as 5 gametes per generation, far higher than we consider likely, F_{ST} estimates failed to approach the values observed in the actual B populations. Increasing migration rates also reduced variance in F_{ST} along the chromosome arm, but once again not to the levels observed in the B populations (Table 1.2).

Footprints of selection

We used the pool-HMM method (Boitard et al. 2013) to detect selective sweeps or changes in allele frequency due to selection across all major chromosome arms in the B populations. When applying Pool-HMM to our real sequence data, we detected dozens of signatures of selection on each of the major chromosome arms for all of the B populations (Figure 1.2 and Table 1.3). Additionally, nearly twice as many regions were detected in the B_1 population compared to the other four replicates. Of the hundreds of candidate selected regions detected, there were ~35 regions that overlapped across all five replicates (Figure 1.3). However, as many of these regions were in excess of 100kb, these results do not definitively point to any specific genes as being targets of selection.

We also applied pool-HMM to results from our neutral simulations, both those including and those excluding migration. We applied the test to regions consisting of 100kb sampled from our neutral simulations using the same settings we applied to the data from the B populations. Using these settings, we found very few instances where selection was detected, suggesting a low false positive rate (Table S1.3). For instance, we only found 2 instances where selection was falsely detected after applying pool-HMM to 300 100kb regions sampled from our neutral simulations with no migration. This was also true for 300 regions sampled from neutral simulations with M=1. Lastly, when M=5 there were zero instances where selection was falsely detected.

We also ran pool-HMM on results from the entirety of 3R for several of the simulated neutral and selective scenarios we tested. We did this for one simulated population from each of the following scenarios: neutral evolution with M=0, 5 and 20; 3 selected QTL's with low starting frequencies with M=0, 5 and 20; 3 selected QTL's with high

starting frequencies with M=0, 5 and 20; 10 selected QTL's with low starting frequencies with M=0, 5 and 20; 10 selected QTL's with high starting frequencies with M=0, 5 and 20. Once again, few regions were identified as being under selection when pool-HMM was applied to results from our neutral simulations relative to what we observe in the B populations. Seven regions in total were detected when we applied it to results from a neutral simulation with M=0, 2 regions when M=5, and 0 when M=20 (Figure S1.6).

However, we found pool-HMM's ability to detect selected QTL to be highly dependent on the starting frequency of the selected QTL and on assumed migration rates. When starting frequencies are low (.05) and there is no migration, there is some correspondence between the regions identified by pool-HMM and the locations of the actual selected QTL with pool-HMM identifying regions overlapping or adjacent to selected QTL (Figure S1.7 and Figure S1.8). However, when the starting frequencies of selected QTL were high (.5), this correspondence broke down (Figure S1.9 and Figure S1.10). For instance, when we applied pool-HMM to results from a simulation with zero migration and 10 selected QTLs starting at low frequency, 11 regions were identified as being under selection by pool-HMM (Figure S1.8). Five of these regions directly overlapped with the locations of our selected QTLs, and the remaining regions were adjacent to selected QTL. However, when we applied pool-HMM to results from a simulation with zero migration and 10 selected QTLs starting at high frequency, only 3 regions were detected (Figure S1.10). Of these 3 regions, only one overlapped with the location of a selected QTL.

Migration also had a pronounced effect on pool-HMM's ability to detected selected QTL. Across all the scenarios we tested, we found that increased migration rates resulted in reductions in the number of regions identified as being under selection (Figure S1.6-S1.10).

As mentioned previously, in scenarios with 10 selected QTL's starting at low frequencies, 11 regions were detected when M=0. However, when M=20 only 4 regions were detected and only one of those overlapped with the location of a selected QTL (Figure S1.8). Combining high migration rates and high starting frequencies further impaired pool-HMM ability to detect selected QTL. For instance, when M=5 for simulations with 10 QTL starting at high frequencies, pool-HMM did not identify any regions as being under selection.

In summary, we found that both migration and the starting frequency of selected alleles affect the rate and accuracy at which pool-HMM identifies regions as being under selection. Consequently, the overall correspondence between regions identified by pool-HMM and the location of selected QTL in our simulated results was generally poor. Pool-HMM also detected far fewer regions in our simulated results than in any of the scans of our real data from the B population. This suggests there may be some other evolutionary factor(s) behind the allele frequency distributions in the evolved populations other than those we simulated (selection at modest number of sites, migration, and drift). It is not entirely clear what this factor might be from our results. For instance, this discrepancy could be the result of some demographic factor acting on the B populations, complex epistatic interactions, or some selective scenario not tested. Or perhaps a combination of the three.

DISCUSSION

Applying all our measures of genetic variation to the five observed *Drosophila* populations, we found some depressions indicating reduced genetic polymorphism. But there are no regions where it was completely expunged. When comparing these results to the combined DGRP lines we used as base populations for our simulations, we found levels

of variation in our populations to be lower on average. While there are clearly other factors at play, this disparity could also be due in large part to the nearly 1000 generations of evolutionary domestication that the experimental B populations have been subjected to, domestication that has featured both reduced effective population sizes as well as long-sustained stable patterns of selection. This hypothesis is supported by the localized reductions in polymorphism found within our populations, reductions which are consistent with adaptation involving allele frequencies moving part-way toward fixation (Burke 2012). Given that many of these reductions are consistent across our replicates and genetic variation is never entirely depleted, it also seems reasonable to infer that they result from selection on standing genetic variation.

Our tests for selection using pool-HMM are also suggestive of a widespread response to selection across the genome in our populations. However, it is unclear how many of these regions are indicative of a recent response to selection, or selection in the wild ancestral population sampled by Ives in 1975. We find a number of regions that are consistently implicated across all replicates, which is perhaps indicative of a parallel response to selection. However, further complicating matters, our tests on data from simulated populations suggest demographic factors and the starting frequencies of selected variants can have pronounced effects on pool-HMM's ability to detect regions under selection. The role of the former in particular warrants further investigation. We found that migration produced large reductions in the number of regions identified as being under selection by pool-HMM across all scenarios we tested. Given the number of regions detected when pool-HMM is applied to the B population sequence data, it seems unlikely that migration between B populations is a major confounding factor. However, it is entirely

possible that these results could be due to some other demographic factor or combinations of factors not explored in our simulations.

We find that, after almost 1,000 generations of laboratory cultivation, the five replicate B populations studied here are not generally genetically depauperate. This is somewhat surprising, given the moderate N_e estimates of Mueller et al. (2013) for these populations: generally a bit less than 1,000. If the only evolutionary processes acting on these populations were selective sweeps, background selection, and genetic drift, then it seems odd that such extensive genetic variation is maintained. The results of our simulation add to this mystery, as we consistently find greater simulated reductions in average heterozygosity than that shown by the B populations, across a range of evolutionary scenarios featuring drift and selection. Note that for more than 900 generations, these populations were maintained under stable conditions with respect to life cycle, illumination, density, and handling vessel. This provided an excellent opportunity for a selective sweep to occur, since the B populations were maintained for a long time in a consistent selection regime, much longer than would be likely to arise in nature.

That being said, our simulations are far from perfect. Given the age of our system, we have no record of the starting genetic make-up of our populations. The populations featured in this study were derived from a single population that had been maintained for 130 generations under laboratory conditions. This population was in turn created from 200 gravid females collected in the wild brought into the lab. In contrast, the base population used in our simulations was created by essentially combining a hundred inbred lines from the DGRP. This difference alone represents a major confounding factor. Additionally, while we feel we are justified in excluding the potential for *de novo* beneficial mutations in our

simulations, given the values of our actual N_e and the number of generations under selection, there no doubt remains value in exploring a wider range of evolutionary scenarios. However, addressing these issues satisfactorily would constitute a considerable undertaking, well beyond the scope of this project.

Laboratory selection experiments with *Drosophila* have provided a variety of results and interpretations concerning the underlying mechanisms of adaptation. For instance, both Turner et al. (2011) and Zhou et al. (2011) report patterns of locally-purged genetic variation in evolved populations consistent with the classic signature of complete selective sweeps. But we do not find any regions where genetic variation is locally purged in the manner associated with a complete selective sweep, as heterozygosity across the genome in our evolved populations never unambiguously achieves zero values in well-defined local regions of the genome. Our findings are more consistent with those of Burke et al. (2010), Orozco-terWengel et al. (2012) and Tobler et al. (2014); the patterns of adaptation that they found were attributed to selection on standing genetic variation, without complete fixation of favored alleles.

This discrepancy may be due to differences in experimental methods. For instance, the study of Turner et al. (2011) featured an artificial selection experiment in which flies that met specific body size criteria were selected and allowed to reproduce. Their breeding population sizes were substantially smaller than ours, at 160 females and 160 males. This is in contrast to those experimental evolution studies where there is no direct choice of individuals who will contribute to the next generation, which might have led to very different patterns of evolution. In Zhou et al. (2011), the study populations were founded from 27 isogenic lines. Our populations were not created by crossing of inbred lines. The

populations studied by Orozco-terWengel et al. (2012) and Tobler et al. (2014) were founded using113 isofemale lines, and thus should have had far more genetic variation to begin with than the populations studied by Zhou et al. (2011), perhaps even more than our founding "Ives" population, which was started with about 200 fertilized females sampled from the wild (Rose 1984).

The only published study that is closely comparable to this one is that of Burke et al. (2010), also from our laboratory, although that study was somewhat impaired by the use of a single unpooled replicate population alongside two sets of pools of five replicate populations. Nevertheless, it too featured long sustained selection, founding populations that had never been systematically inbred, and five-fold replication of the selected and ancestral treatments. A failure to detect completely depressed heterozygosity in the five-replicate pools of that study could be attributed to differentiation between replicate populations with respect to selective sweeps. However, the single unpooled population (ACO_1) from the Burke et al. study also did not show clear signatures of completed selective sweeps, despite just over 600 generations of sustained selection.

The high degree of similarity in patterns of variation between our five replicate populations is another surprising aspect of our results. We found that our observed level of F_{ST} was in fact much lower than what would be predicted by classical theory assuming no migration and random divergence between subpopulation. To produce the level of F_{ST} we observe using this model would require an effective population size nearly five times greater than estimates made from empirical data by Mueller et al. (2013), and as such we do not believe this discrepancy can be reasonably explained away by issues with our population size estimates. Our attempts to predict how much migration would be required

to produce our observed $F_{ST,}$ assuming random divergence and populations at equilibrium suggest that ~3 migrants per generation would be sufficient. However, given the nature of our system and its maintenance protocols, migration rates that high *every* generation seem unlikely. It is also worth noting we have no guarantee that the assumption of equilibrium conditions has been met in our populations, which could confound this estimate (Liang et al. 2015). And this lack of equilibrium convergence serves as a possible explanation as to why migrations rates of 5 per generation did not produce our observed F_{ST} in simulated scenarios.

In our simulations, increasing the number of selected sites/QTLs and migration rates both failed to produce comparable F_{ST} estimates to what we find in the five B populations. Increasing migration rates did produce reductions in $F_{ST,}$ and it is likely that a drastic increase in rates of simulated migration per generation would produce something comparable to our observed values. However, once again, migration rates that high *every generation* do not seem likely, given that these populations are maintained independently and all migration is by definition accidental. It would also likely result in too much genetic variation being maintained, relative to the patterns in our genome-wide data, unless population sizes were also reduced. Our results could possibly be explained by parallel selection at a large number of loci, but more work would be required to test this hypothesis. And our initial findings suggest even this explanation may not be adequate.

Regarding the relative importance of selective sweeps, our results are not conclusive. However, the high levels of genetic variation maintained in these populations in the face of relatively small population sizes, which foster genetic drift and background selection (Mueller et al. 2013), together with the long-sustained selection which should

foster reduced genetic variation due to selective sweeps, seem difficult to reconcile with the idea of adaptation primarily driven by hard selective sweeps. If one's imagination extends so far as to suppose that such selective sweeps arise for a few alleles which are consistently favored by natural selection for far more than 1,000 generations, say for 100,000 generations, then our experiment does not test for the existence of such alleles. We doubt that any laboratory experiment with outbred metazoa will accomplish such a test in this century (Endler 1986).

REFERENCES

Barrick, J.E. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243-1247 (2009).

Boitard, S., Schlötterer, C., Nolte, V., Pandey R.V. & Futschik, A. Detecting selective sweeps from pooled next-generation sequencing samples. *Mol. Biol. Evol*. **29**, 2177-2186 (2012).

Boitard, S. et al. A. Pool-HMM: a Python program for estimating the allele frequency spectrum and detecting selective sweeps from next generation sequencing of pooled samples. *Mol. Ecol. Resour.* **13**, 337-340 (2013).

Burke, M.K. *et al*. Genome-wide analysis of a long-term selection experiment with *Drosophila*. *Nature* **467**, 587–592 (2010).

Burke, M.K. How does adaptation sweep through the genome? Insights from long-term selection experiments. *Proc. R. Soc.* **279**, 5029-5038 (2012).

Dobzhansky, T. & Wright, S. Genetics of natural populations. V. Relations between mutation rate and accumulation of lethals in populations of *Drosophila pseudoobscura*. *Genetics* **26**, 23-51 (1941).

Endler, J.A. *Natural selection in the wild.* Princeton University Press, New Jersey (1986).

Falconer, F.S. & Mackay, T.F.C. *Introduction to quantitative genetics.* Pearson, New York (1996).

Fiston-Lavier, A.S., Singh, N.D., Lipatov, M. & Petrov, D.A. *Drosophila melanogaster recombination rate calculator*. *Gene* **463**, 18-20 (2010).

Gillespie, *J. Population genetics: a concise guide*. Johns Hopkins University Press, Baltimore (2010).

Hartl, D.L. & Clark, A. *Principles of population genetics*. Sinauer, Massachusetts (1997).

Hedrick, P.W. *Genetics of populations.* Jones & Bartlett Learning Press, Massachusetts (2009).

Hermisson, J. & Pennings, P. S. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335–2352 (2005).

Ives, P. T. Further genetic studies of the South Amherstpopulation of *Drosophila melanogaster*. *Evolution* **24**, 507–518 (1970).

Kofler, R. *mimicree* https://sourceforge.net/p/mimicree/wiki/Manual/ (2015) (Date of access: 1/7/2016).

Kofler, R., Pandey, R.V. & Schlötterer, C. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **27**, 3435–3436 (2011a).

Kofler, R. *et al.* PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS ONE* **6**: e15925 (2011b).

Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*: 1303.3997v1 [q-bio.GN] (2013).

Li, H., & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

Li, H. *et al.* The sequence alignment / map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

Liang, M.A., Ya-Jie, J.J., & De-Xing, Z. Statistical measures of genetic differentiation of populations: rationales, history, and current states. *Curr. Zool.* **61**, 886-897 (2015).

Long, A.D., Liti, G., Luptak, A., &Tenaillon, O. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nat. Rev. Genet*. **16**, 567-582 (2015).

Mackay, T.F.C. *et al.* The *Drosophila melanogaster* genetic reference panel. *Nature* **482**, 173-178 (2012).

Maynard Smith, J. & J. Haigh. The hitch-hiking effect of a favorable gene*. Genet. Res.* **23**: 23-35 (1974).

Mueller, L.D., Joshi, A., Santos, M. & Rose, M.R. Effective population size and evolutionary dynamics in outbred laboratory populations of *Drosophila. J. Genet.* **92**, 349-361(2013)

Oleksyk, T. K., Smith, M. W. & O'Brien, S. J. Genome-wide scans for footprints of natural selection. *Philos. Trans. R. Soc. Lond*. *B* **365**, 185–205 (2010).

Orozco-terWengel, P. *et al*. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected traits. *Mol. Ecol.* **21**: 4931-41 (2012).

Pennings P. S., Hermisson J. Soft sweeps II: molecular population genetics of adaptation from recurrent mutation or migration. *Mol. Biol. Evol.* **23**, 1076–1084 (2006).

Przeworski, M., Coop, G. & Wall, J. D. The signature of positive selection on standing genetic variation. *Evolution* **59**, 2312–2323 (2005).

Rose, M.R. Laboratory evolution of postponed senescence in *Drosophila melanogaster. Evolution* **38**, 1004-1010 (1984).

Rose, M.R. & Charlesworth, B. A test of evolutionary theories of senescence. *Nature* **287**, 141-142 (1980).

Rose, M.R. & Charlesworth, B. Genetics of life-history in *Drosophila melanogaster* II. Exploratory selection experiments. *Genetics* **97**, 187-196 (1981).

Rose, M.R., Passananti, H.B. & Matos, M. *Methuselah Flies: A Case Study in the Evolution of Aging*. World Scientific Publishing, Singapore (2004).

Schlotterer, C. Kofler, R., Versace, E., Toberl, R. & Franssen, S.U. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity* doi: 10.1038/hdy.2014.86 (2014).

Smith, A.F.A., Hubley, R. & Green, P. *RepeatMasker Open-4.0* http://www.repeatmasker.org/ (2013-2015).

Teotonio H. & Rose, M.R. Variation in the reversibility of evolution. *Nature* **408**, 463-466 (2000).

Tenaillon, O. *et al.* The molecular diversity of adaptive convergence. *Science* **335**, 457-461 (2012).

Tobler R. *et al.* Massive habitat-specific genomic response in D. *melanogaster* populations during experimental evolution in hot and cold environments. *Mol Biol Evol.* **31**, 364-375 (2014).

Turner, T. L., Stewart, A. D., Fields, A. T., Rice, W. R. & Tarone, A. M. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet*. **7**: e1001336 (2011).

Zhou, D. et al. Experimental selection of hypoxia- tolerant *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **108**, 2349–2354 (2011).

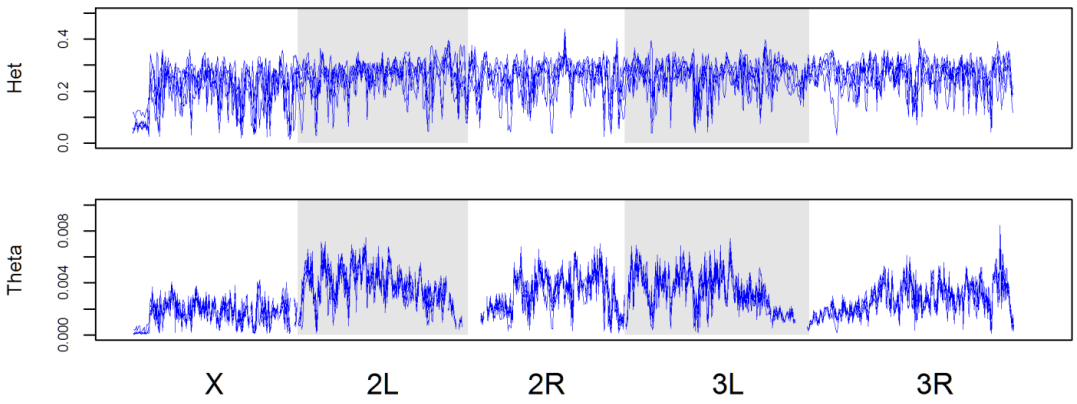

Figure 1.1. Genome-wide patterns of genetic variation 100kb windows.
Heterozygosity and Watterson theta (Θ) plotted across 100kb non-overlapping windows across all major chromosome arms for the 5 B populations. All replicates are shown.

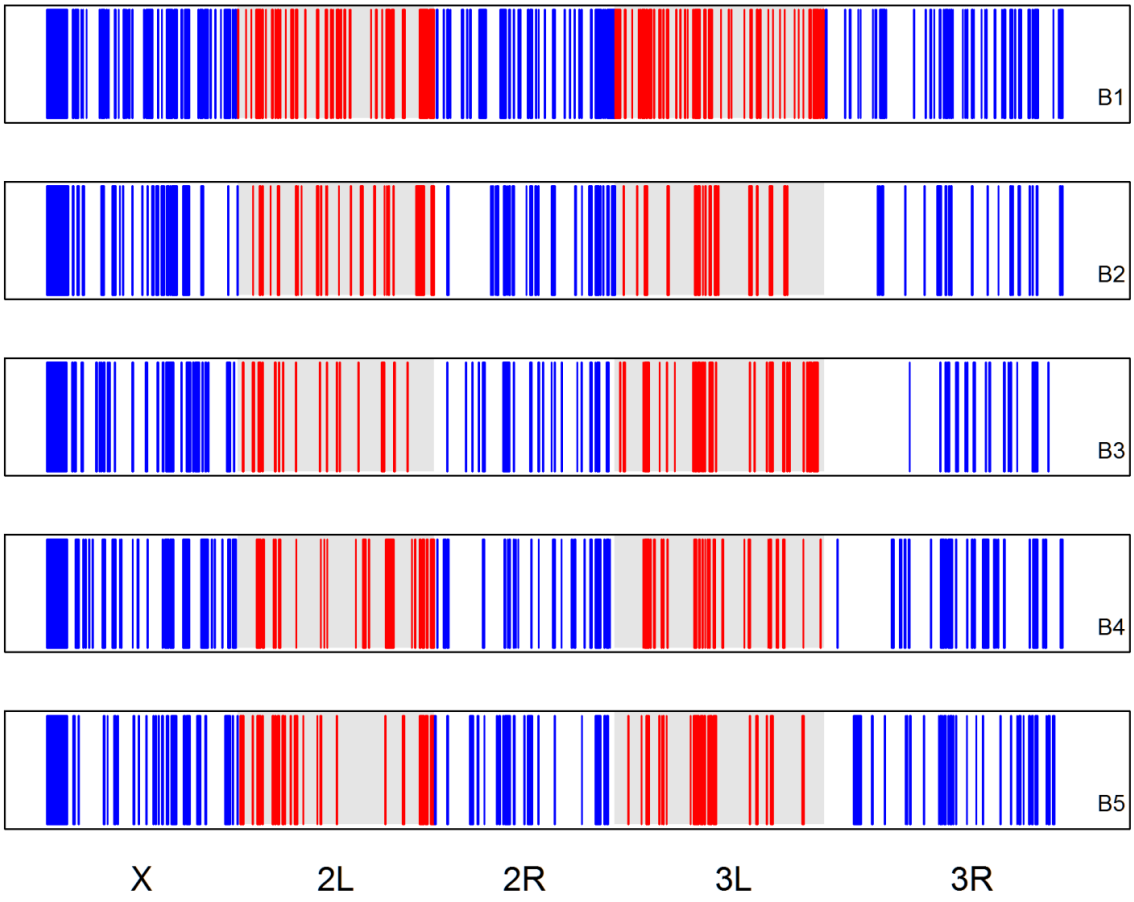

Figure 1.2. All selective sweeps detected. Regions across all major chromosome arms in the 5 B populations showing evidence for selection based on our analysis using Pool-Hmm. Each panel shows results from a different B population replicate. There is no significance to the color coding outside other than differentiating adjacent chromosome arms.

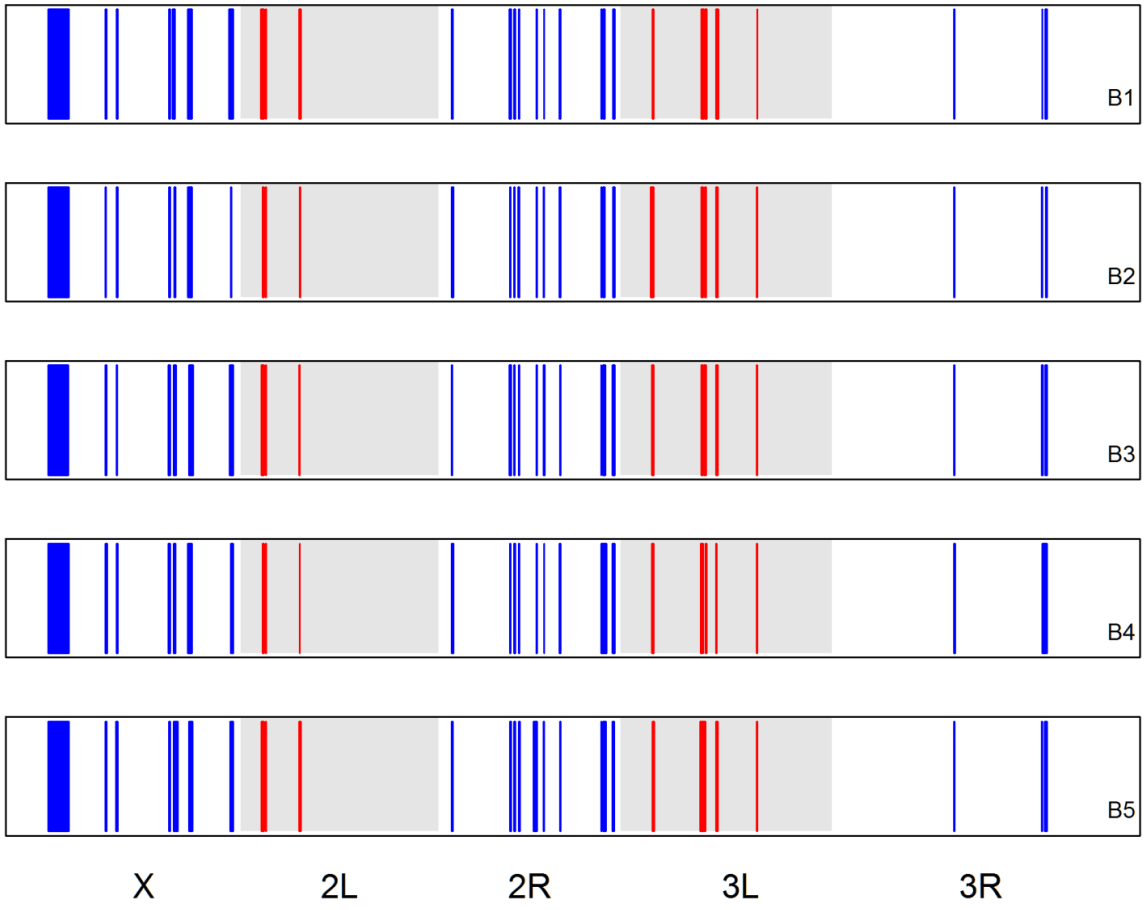

Figure 1.3. Overlapping selective sweeps. Overlapping regions across all major chromosome arms showing evidence for selection across all 5 B populations based on our analysis using Pool-Hmm. Each panel shows results from a different B population replicate as these regions do not perfectly overlap. There is no significance to the color coding other than differentiating adjacent chromosome arms.

Table 1.1. 95% confidence intervals for average F_{ST} and average heterozygosity for simulations with unconditionally beneficial alleles and overdominance. Confidence intervals for each scenario are based on the distribution of these values taken from 100 simulation runs where each run consists of 5 simulated populations.

Heterozygous Effect	Number of Sites	Selection Coefficient (s)	Mean Het	Variance Het	Mean F_{ST}	Variance F_{ST}
B populations	NA	NA	0.28	0.0024	0.08	0.0004
	NA	NA	0.29	0.0021		
	NA	NA	0.28	0.0024		
	NA	NA	0.27	0.0037		
	NA	NA	0.27	0.0037		
Neutral	NA	NA	0.22	$0.0031 \pm 7.0 \times 10^{-5}$	0.26	$0.0022 \pm 8.5 \times 10^{-5}$
Overdominance	20	0.03	0.20	$0.0039 \pm 5.3 \times 10^{-5}$	0.24	0.0039 ± 0.0002
		0.065	0.19	$0.0041 \pm 5.0 \times 10^{-5}$	0.24	0.0051 ± 0.0002
		0.1	0.18	$0.0042 \pm 4.4 \times 10^{-5}$	0.26	0.0066 ± 0.0003
		$0.03 < s < 0.1$	0.21	$0.0032 \pm 5.4 \times 10^{-5}$	0.24	$0.0028 \pm 9.0 \times 10^{-5}$
	30	0.03	0.22	$0.0034 \pm 8.0 \times 10^{-5}$	0.25	0.0027 ± 0.0001
		0.065	0.20	$0.0041 \pm 9.8 \times 10^{-5}$	0.27	0.0038 ± 0.0002
		0.1	0.15	0.0060 ± 0.0001	0.42	0.0108 ± 0.0005
		$0.03 < s < 0.1$	0.19	$0.0040 \pm 6.5 \times 10^{-5}$	0.27	0.0045 ± 0.0002
A1 Dominant	5	0.065	.20	0.0060 ± 0.0002	0.29	0.0055 ± 0.0005
	10	0.0325	.21	$0.0035 \pm 7.6 \times 10^{-5}$	0.28	0.0033 ± 0.0001
	20	0.01625	.19	0.0058 ± 0.0002	0.32	0.0055 ± 0.0002
A2 Dominant	5	0.065	.20	0.0055 ± 0.0002	0.28	0.0042 ± 0.0002
	10	0.0325	.20	$0.0037 \pm 6.1 \times 10^{-5}$	0.27	0.0033 ± 0.0001
	20	0.01625	.19	0.0050 ± 0.0001	0.30	0.0045 ± 0.0002
Codominant	5	0.065	.20	0.0060 ± 0.0002	0.29	0.0051 ± 0.0003

	10	0.0325	.20	$0.0040 \pm 6.4 \times 10^{-5}$	0.29	0.0041 ± 0.0002
	20	0.01625	.18	0.0061 ± 0.0002	0.33	0.0060 ± 0.0003

Table 1.2. Average genome wide F_{ST} and average heterozygosity for B populations and simulations with selection and migration. For the B populations, variance in heterozygosity and F_{ST} over 50kb windows is shown. For each simulated scenarios, 95% confidence intervals for variance in heterozygosity and F_{ST} over 50kb windows calculated from replicate simulation are shown.

Populations/Selection Scenario	Migration Rate	Mean Het	Variance Het	Mean F_{ST}	Variance F_{ST}
B populations	NA	0.28	0.0024	0.08	0.0004
	NA	0.29	0.0021		
	NA	0.28	0.0024		
	NA	0.27	0.0037		
	NA	0.27	0.0037		
Neutral	M=0	0.22	0.0025 ± 2.5 X 10^{-5}	0.24	0.0021 ± 6.5 X 10^{-5}
	M=1	0.22	0.0024 ± 2.6 X 10^{-5}	0.22	0.0020 ± 6.2 X 10^{-5}
	M=5	0.24	0.0020 ± 2.2 X 10^{-5}	0.17	0.0012 ± 3.7 X 10^{-5}
3 QTLs with 0.05 starting freq.	M=0	0.19	0.0037 ± 5.3 X 10^{-5}	0.24	0.0027 ± 0.0001
	M=1	0.20	0.0034 ± 4.8 X 10^{-5}	0.22	0.0023 ± 8.6 X 10^{-5}
	M=5	0.22	0.0029 ± 4.6 X 10^{-5}	0.16	0.0015 ± 6.6 X 10^{-5}
3 QTLs with 0.5 starting freq.	M=0	0.21	0.0025 ± 2.8 X 10^{-5}	0.24	0.0023 ± 6.6 X 10^{-5}
	M=1	0.22	0.0024 ± 2.5 X 10^{-5}	0.22	0.0021 ± 7.2 X 10^{-5}
	M=5	0.24	0.0020 ± 2.2 X 10^{-5}	0.16	0.0013 ± 4.3 X 10^{-5}
10 QTLs with 0.05 starting freq.	M=0	0.16	0.0048 ± 5.2 X 10^{-5}	0.26	0.0033 ± 0.0002
	M=1	0.17	0.0046 ± 6.1 X 10^{-5}	0.24	0.0030 ± 0.0002
	M=5	0.19	0.0039 ± 6.0 X 10^{-5}	0.17	0.0018 ± 0.0001
10 QTLs with 0.5 starting freq.	M=0	0.21	0.0026 ± 3.1 X 10^{-5}	0.24	0.0025 ± 9.6 X 10^{-5}
	M=1	0.21	0.0025 ± 2.7 X 10^{-5}	0.22	0.0022 ± 9.3 X 10^{-5}
	M=5	0.23	0.0021 ± 2.2 X 10^{-5}	0.17	0.0014 ± 5.6 X 10^{-5}
20 QTLs with 0.05 starting freq.	M=0	0.16	0.0050 ± 5.5 X 10^{-5}	0.27	0.0044 ± 0.0002

	M=1	0.17	0.0048 ± 5.4 X 10^{-5}	0.24	0.0040 ± 0.0002
	M=5	0.19	0.0043 ± 6.2 X 10^{-5}	0.18	0.0022 ± 0.0001
20 QTLs with 0.5 starting freq.	M=0	0.21	0.0027 ± 3.4 X 10^{-5}	0.24	0.0026 ± 0.0001
	M=1	0.21	0.0026 ± 2.7 X 10^{-5}	0.22	0.0022 ± 7.9 X 10^{-5}
	M=5	0.23	0.0021 ± 2.3 X 10^{-5}	0.17	0.0015 ± 5.9 X 10^{-5}

Table 1.3. Number of regions where selection was detected using Pool-HMM method with different per site transition probabilities (*q*).

	B1			B2			B3			B4			B5		
q=	10^{-9}	10^{-10}	10^{-11}	10^{-9}	10^{-10}	10^{-11}	10^{-9}	10^{-10}	10^{-11}	q10^{-9}	10^{-10}	10^{-11}	10^{-9}	10^{-10}	10^{-11}
2L	68	61	53	41	33	29	30	26	22	40	34	29	39	35	31
2R	70	62	58	44	42	37	40	36	31	38	33	26	36	32	29
3L	79	62	54	32	31	29	52	43	38	42	38	34	32	31	29
3R	75	67	64	37	32	25	27	24	22	51	48	44	49	42	37
X	70	61	53	50	45	43	49	41	35	52	45	43	39	35	32
Total	362	313	282	204	183	163	198	170	148	223	198	176	195	175	158

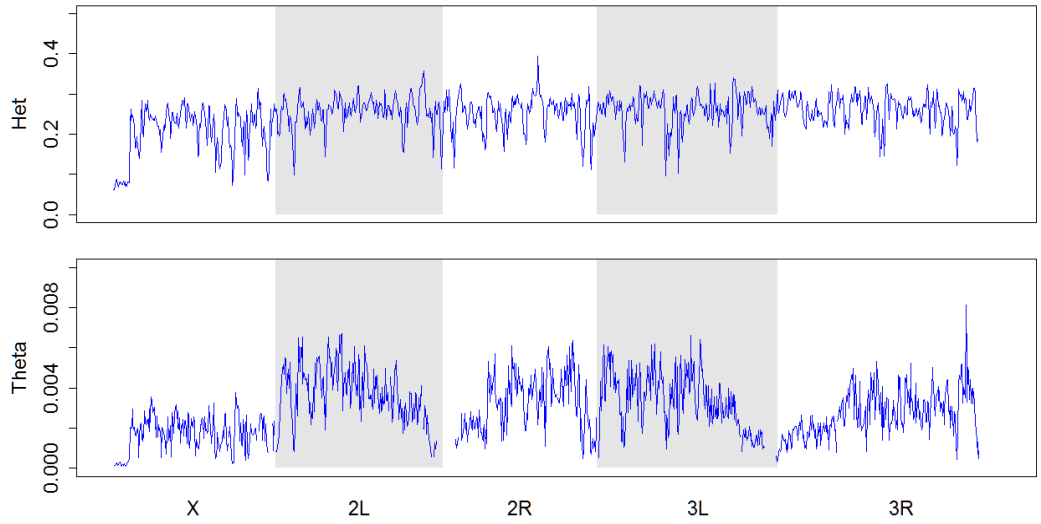

Figure S1.1. Mean genome-wide patterns of genetic variation 100kb windows. Mean heterozygosity and Watterson theta (Θ) plotted across 100kb non-overlapping windows across all major chromosome arms for the 5 B populations.

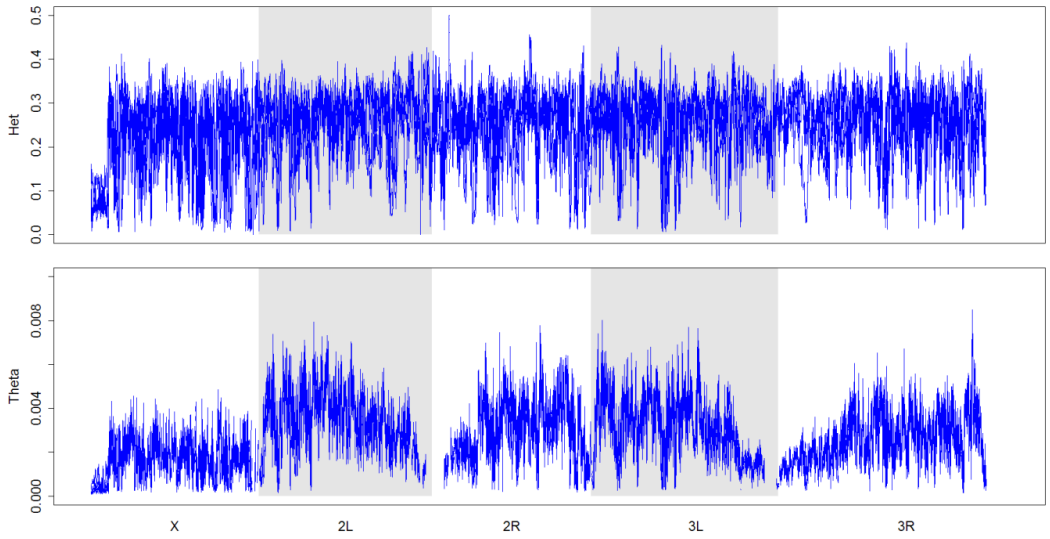

Figure S1.2. Genome-wide patterns of genetic variation 30kb windows.
Heterozygosity and Watterson theta (Θ) plotted across 30kb non-overlapping windows across all major chromosome arms for the 5 B populations.

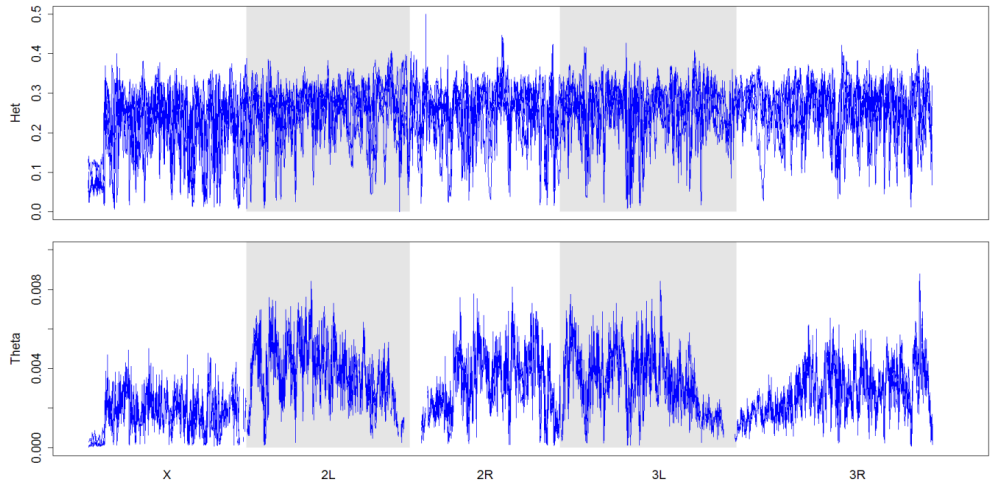

Figure S1.3. Genome-wide patterns of genetic variation 50kb windows.
Heterozygosity and Watterson theta (Θ) plotted across 50kb non-overlapping windows across all major chromosome arms for the 5 B populations.

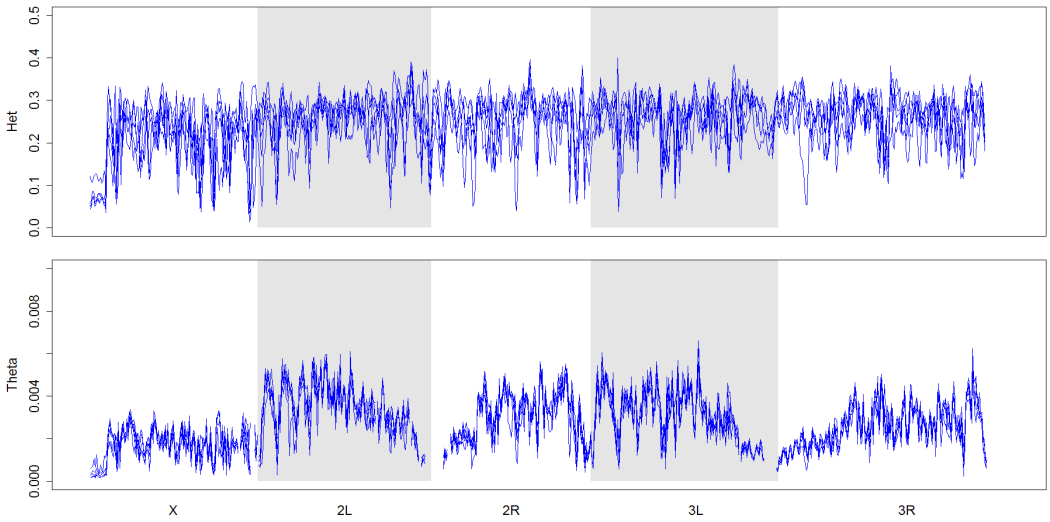

Figure S1.4. Genome-wide patterns of genetic variation 150kb windows.
Heterozygosity and Watterson theta (Θ) plotted across 150kb non-overlapping windows across all major chromosome arms for the 5 B populations.

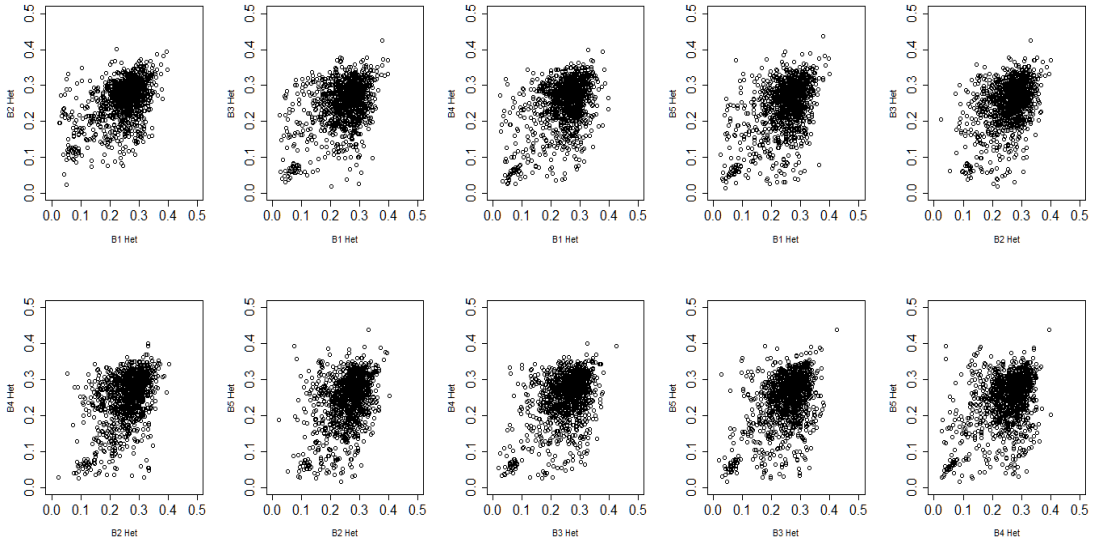

Figure S1.5. Pair-wise comparisons of heterozygosity. Values calculated over 100kb windows for all 5 B populations.

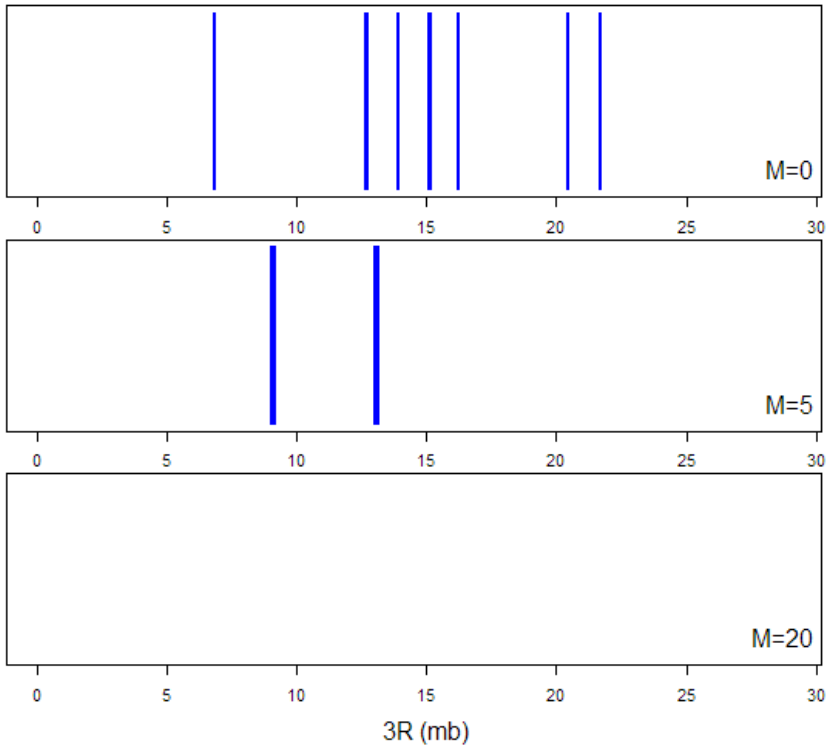

Figure S1.6. Pool-HMM results for drift simulation. Pool-HMM results for chromosome arm 3R from simulations featuring neutrally evolving populations. Regions where selection was detected by the Pool-HMM are shown in blue. Results are shown for a population with no migration (top), a population from a group of 5 where there were 5 migration events per generation, and a population from a group of 5 where there were 20 migration events per generation.

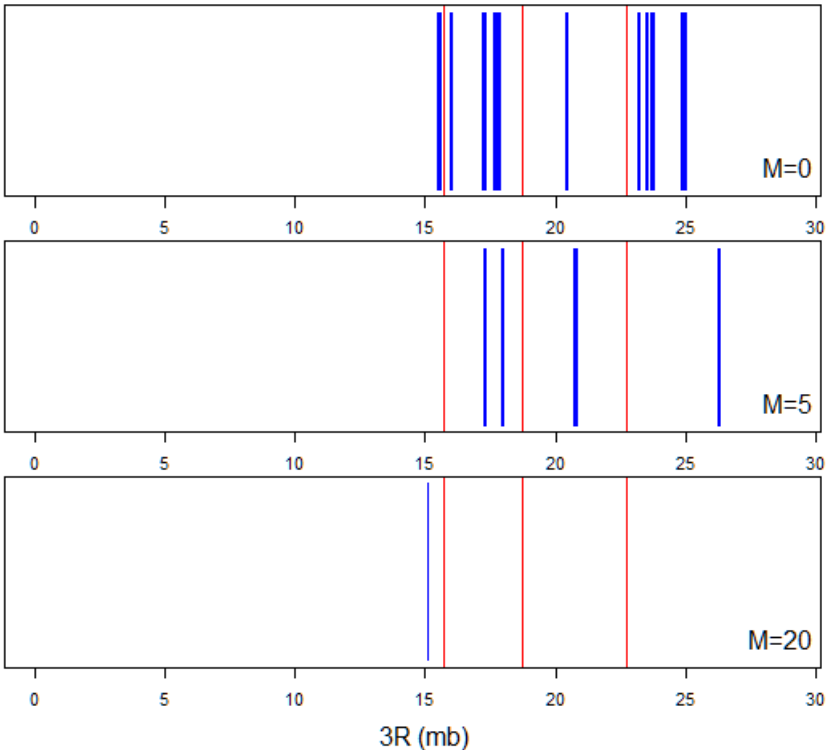

Figure S1.7. Pool-HMM results 3 selected QTL low starting frequency. Pool-HMM results for chromosome arm 3R from simulations with 3 selected QTL starting at low frequencies (0.05). The locations of the selected QTLs are indicated by red lines, and regions where selection is detected by Pool-HMM are in blue. Results are shown for a population with no migration (top), a population from a group of 5 where there were 5 migration events per generation (middle), and a population from a group of 5 where there were 20 migration events per generation (bottom).

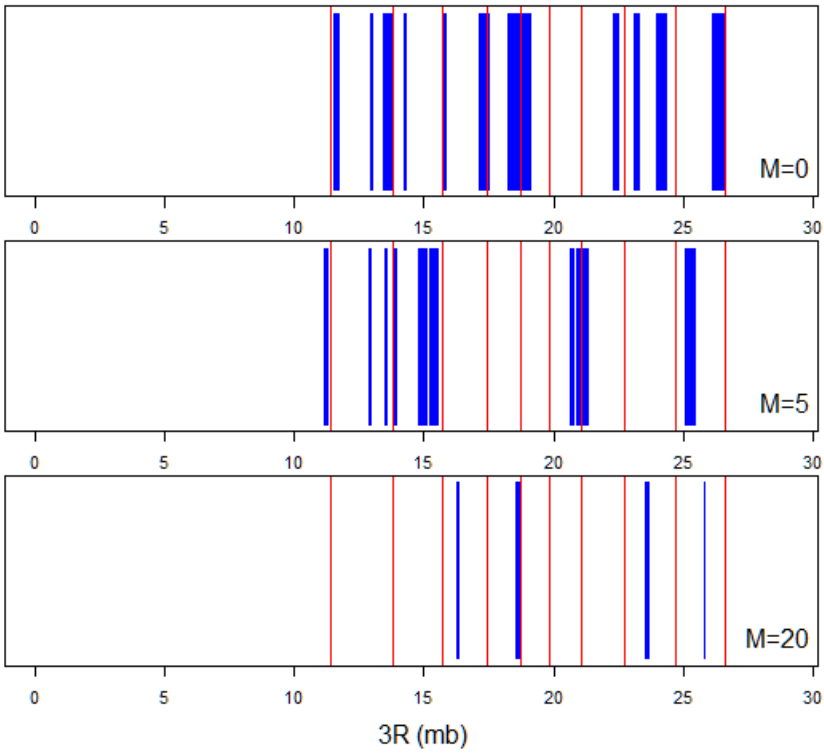

Figure S1.8. Pool-HMM results 10 selected QTL low starting frequency. Pool-HMM results for simulations with 10 selected QTL starting at low frequencies (0.05). The locations of the selected QTLs are indicated by red lines, and regions where selection is detected by Pool-HMM are in blue. Results are shown for a population with no migration (top), a population from a group of 5 where there were 5 migration events per generation (middle), and a population from a group of 5 where there were 20 migration events per generation (bottom).

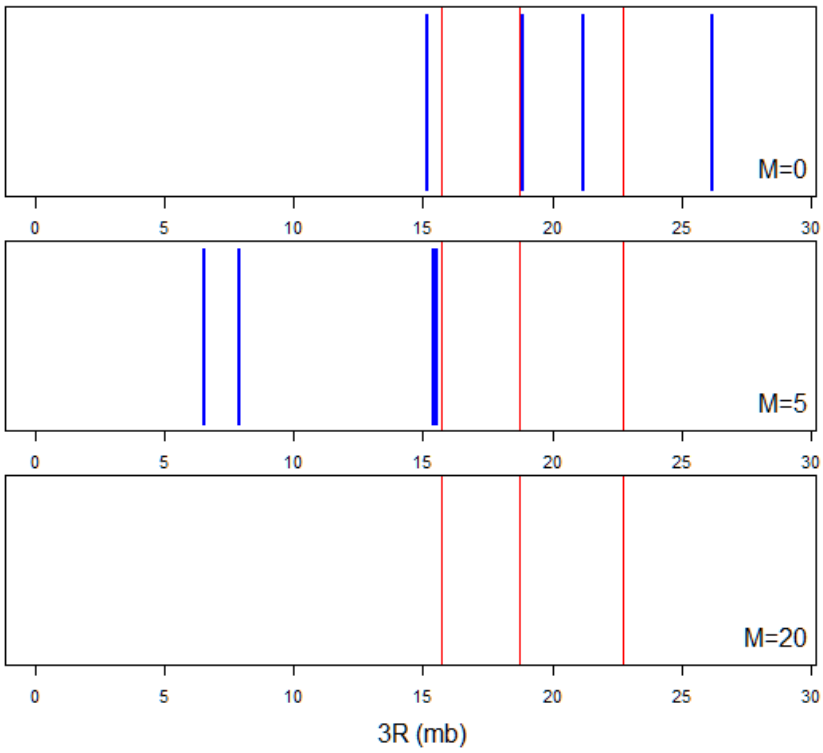

Figure S1.9. Pool-HMM results 3 selected QTL high starting frequency. Pool-HMM results chromosome arm 3R from simulations with 3 selected QTL starting at high frequencies (0.5). The locations of the selected QTLs are indicated by red lines, and regions where selection is detected by Pool-HMM are in blue. Results are shown for a population with no migration (top), a population from a group of 5 where there were 5 migration events per generation (middle), and a population from a group of 5 where there were 20 migration events per generation (bottom).

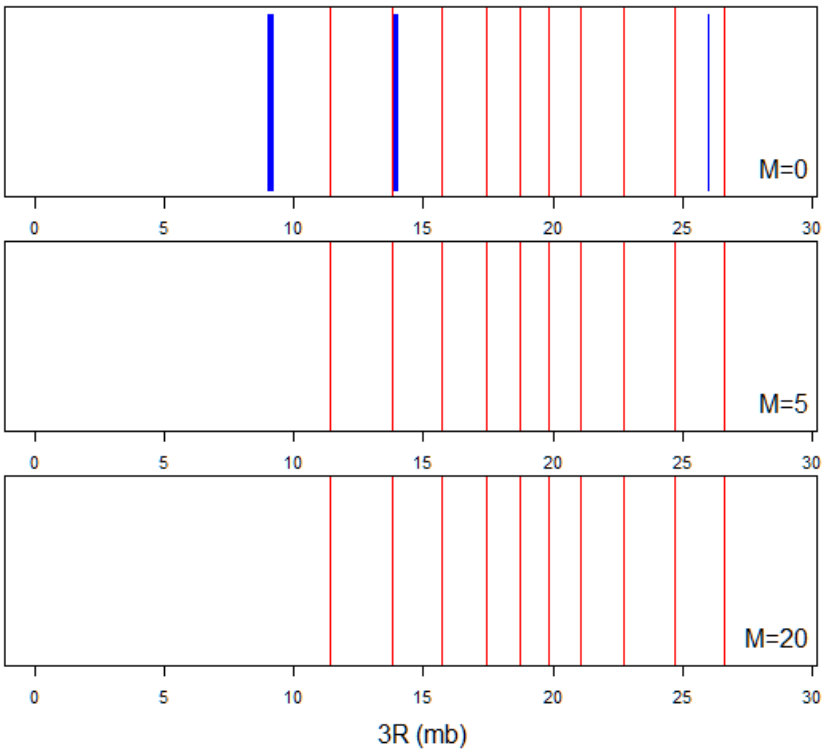

Figure S1.10. Pool-HMM results 10 selected QTL high starting frequency. Pool-HMM results for simulations with 10 selected QTL starting at high frequencies (0.5). The locations of the selected QTLs are indicated by red lines, and regions where selection is detected by Pool-HMM are in blue. Results are shown for a population with no migration (top), a population from a group of 5 where there were 5 migration events per generation (middle), and a population from a group of 5 where there were 20 migration events per generation (bottom).

Table S1.1. 100 kb regions that consistently have values of Θ less than 0.001 across all 5 B populations.

Chromosome	Start of region
2L	50000
2L	21350000
2L	21750000
2R	19150000
3L	150000
3R	50000
3R	150000
3R	250000
3R	550000
3R	1450000
3R	8250000
3R	27850000
X	250000
X	350000
X	450000
X	550000
X	650000
X	750000
X	850000
X	950000
X	1050000
X	1150000
X	1250000
X	1350000
X	1450000
X	1550000
X	1650000
X	1750000
X	1850000
X	1950000
X	2050000
X	2150000
X	2250000
X	2350000
X	7350000
X	11450000
X	14150000
X	16350000
X	16450000

X	16550000
X	16650000

Table S2.2. 100 kb regions that consistently have values of heterozygosity less than 0.2 across all 5 B populations.

Chromosome	Start of region
2L	2505390
2L	22705390
2L	22805390
2R	19110083
2R	20310083
3L	3719855
3L	9419855
3L	9819855
3L	24019855
3R	14200284
3R	14300284
X	120552
X	220552
X	320552
X	420552
X	520552
X	620552
X	720552
X	820552
X	920552
X	1020552
X	1120552
X	1220552
X	1320552
X	1420552
X	1520552
X	1620552
X	1720552
X	1820552
X	1920552
X	2020552
X	2120552
X	2220552
X	2320552
X	3720552
X	14120552
X	14620552
X	16420552
X	16620552

Table S1.3. Results from Pool-HMM being applied to 100kb windows samples from our neutral simulations. Shows the total number of instances where selection was detected across all 100kb regions taken from each set of neutral simulations.

Scenario	Migration Rate	Number of Simulated Populations	Total Number of Instances where Selection was Detected
Neutral	M=0	300	2
	M=1	300	2
	M=5	300	0

CHAPTER 2

Genomics of Parallel Experimental Evolution in *Drosophila*

ABSTRACT

What are the genomic foundations of adaptation in sexual populations? We address this question using fitness-character and whole-genome sequence data from 30 Drosophila laboratory populations. These 30 populations are part of a nearly forty-year laboratory radiation featuring three selection regimes, each shared by ten populations for up to 837 generations, with moderately large effective population sizes. Each of three sets of ten populations that shared a selection regime consist of five populations that have long been maintained under that selection regime, paired with five populations that had only recently been subjected to that selection regime. We find a high degree of evolutionary parallelism in fitness phenotypes when most-recent selection regimes are shared, as in previous studies from our laboratory. We also find genomic parallelism with respect to the frequencies of single-nucleotide polymorphisms, transposable elements, insertions, and structural variants, which was expected. Entirely unexpected was a high degree of parallelism for linkage disequilibrium. The evolutionary genetic changes among these sexual populations are rapid and genomically extensive. This pattern may be due to segregating functional genetic variation that is abundantly maintained genome-wide by selection, variation that responds immediately to changes of selection regime.

INTRODUCTION

Genome-wide sequencing of experimentally-evolved populations has emerged as a powerful method for parsing the genetic underpinnings of adaptation (Burke et al. 2010; Turner et al. 2011; Tenaillon et al. 2012; Schlötterer et al. 2015). An emerging pattern in this research is a contrast between the genomics of experimental evolution in asexual and sexual populations. Initial adaptation in the most common asexual paradigm, serially-cultured Escherichia coli, features sequential selective sweeps of new mutations, at least over the first several thousand generations of laboratory selection (e.g., Barrick et al. 2009, Tenaillon et al. 2012, Maddamsetti et al. 2015). By contrast, initial adaptation in the most common sexual paradigm, outbred laboratory Drosophila melanogaster, apparently depends on moderate changes in the allele frequencies of standing genetic variation, rather than selective sweeps, at least for as many as 600 generations (Burke et al. 2010; Turner et al. 2011; Burke 2012; Orozco-ter Wengel et al. 2012; Tobler et al. 2014). Here we present data for fitness characters and whole-genome, pooled, DNA sequences of 30 Drosophila laboratory populations taken from a large laboratory radiation of outbred populations (Rose et al. 2004; Mueller et al. 2013).

In this study, our primary goal is to establish whether the emerging contrast between asexual and sexual populations is sustained over a wide range of evolutionary durations, from dozens to nearly 1,000 generations of sustained selection regimes in sexual populations. Our second goal is to determine the degree to which phenotypic and genomic parallelism occurs between populations that share recently-imposed versus long-sustained selection regimes. Our third goal is to probe the evolutionary genetic mechanisms underlying parallelism within each set of ten populations that share a recent selection

regime. The terms parallelism and convergence are often confused in the literature (Fong et al. 2005; Arendt and Reznick 2007), and in any case are terms of shifting usage. One such usage is that "convergence" refers to a similar phenotype evolving by different genetic mechanism among distantly related populations and species (Arendt and Reznick 2007), while "parallelism" refers to the emergence of similar functional, structural, and genomic features of populations that diverged from a common ancestor (Elias and Tawfik 2012). In this usage, our experimental system is well qualified to examine such "parallelism," as all 30 study populations share a moderately distant ancestor by the standards of experimental evolution: the Ives population (vid. Ives 1970; Rose et al. 2004; Burke et al. 2016).

There have been a number of experimental evolution studies of phenotypic convergence or parallelism in Drosophila (Service et al. 1988; Teotonio and Rose 2000; Matos and Avelar 2001; Teotonio and Rose 2001; Matos et al. 2002; Matos et al. 2004; Teotonio et al. 2009). Here we employ Drosophila stocks that feature five-fold replicated groups of outbred populations that have been subjected to parallel sequences of selection regimes for decades (Rose et al. 2004; Burke et al. 2016). Three selection regimes were repeatedly imposed on these five-population groups, here called "A-type", "B-type", and "C-type" (Figure 2.1A). These selection treatments differ chiefly with respect to the length of their discrete generations, which are 10, 14, and 28 days, respectively. Fifteen of these populations were created decades ago and have long been subjected to either A, B, or C-type selection (the ACO, B, and CO populations, respectively). Matched to them are 15 populations recently derived from five common ancestral "O" populations and since subjected to one of A, B, or C-type selection (the AO, BO, and nCO populations). In total, this experimental system features 30 populations subsequently assessed using pooled genome-

wide sequencing and focal fitness assays, structured as three groups of ten populations each sharing the same recent selection history (Figure 2.1B).

MATERIALS AND METHODS

Experimental system

The laboratory phylogeny used for this experiment consisted of 30 D. melanogaster populations that underwent selection for more than three decades, with five replicate populations maintained for each of six different evolutionary histories. The estimated effective population size for each population is near 1000 for populations maintained with reproduction in either vials or cages (Mueller et al. 2013). For the 30 populations studied here, we had three selection treatments which differed chiefly with respect to the length of their discrete generations: the duration from egg-collection that began each generation to the egg-laying that started the following generation. A-type selection requires rapid larval development, with flies newly emerged from rearing vials transferred to population cages for egg-laying, in order to start the next discrete generation at about 10 days of age. B-type selection involves population maintenance exclusively in rearing vials, with laying adults harvested at 14 days for a few hours of egg-laying to start the next discrete generation. C-type selection is like A-type selection, except that adults are collected at 14 days from egg, and then these adults are kept in cages for 12 days with unyeasted plates of medium, followed by two days of yeasted plates on which egg-laying occurs.

Fifteen populations of A, B, and C-type populations were created long ago with respect to generation number: five “ACO” populations that had undergone 737 generations of A-type selection at the time of the experiments reported here; five “B” populations that had undergone 837 generations of B-type selection; and five “CO” populations that had 290

generations of C-type selection. Corresponding to them are 15 populations all derived from five "O-type" populations (vid. Rose, 1984) that were cultured from eggs laid by females 9-10 weeks old: five "AO" populations given 146 generations of A selection; five "BO" given 121 generations of B selection; and five "nCO" given 37 generations of C selection. Their derivation is described in detail in Burke et al. (2016). These generation numbers for the three selection regimes are all calibrated to the generation at which flies were sampled for pooled DNA sequencing.

Phenotypic assay

The flies assayed phenotypically were taken from the same generation as those we used for pooled DNA sequencing. Assays of fecundity and age-specific survival were conducted in cages holding adults during the 19 day interval between day 9 (from egg) and day 28 (from egg). This time period includes the longest duration that any adult fly is allowed to live in our present stock system, which no longer includes the O populations. Fitness components were calculated from the data collected from the assay cages during this 19 day interval. As is our normal practice, all phenotypic assays were performed after two generations of common-garden rearing using 14-day B-type culture (vid. Rose et al., 2004).

To test for phenotypic differentiation between newly derived and long standing replicates of the same treatment, we tested effects of selection on fecundity over 3-4 consecutive ages. The observations consisted of fecundity at a particular age (t) but within a small age interval (k=1,2,…,m). Within each interval, mortality or fecundity rates were modeled by a straight line and allowing selection regime (j= 1 (ACO or B or CO), 2 (AO or BO or nCO)) to affect the intercept of that line but not the slope. However, slopes were

allowed to vary between intervals. Populations (i= 1,…,10) were assumed to contribute random variation to these measures. With this notation the fecundity at age-t, interval-k, selection regime-j and population-i, is y_{ijkt} and is described by,

$$yijkt = \alpha + \beta_k + \delta_j\gamma_j + (\omega + \pi_k\delta_k)t + \delta_k\delta_j\mu_{jk} + c_i + \varepsilon_{ijkt}$$

where $\delta_s = 0$ if s=1 and 1 otherwise and c_i, and ε_{ijkt} are independent standard normal random variables with variance $\sigma c2$ and σ_ε^2 respectively. The effects of selection on the intercept are assessed by considering the magnitude and variance of both γ_φ and μ_{jk}.

DNA extraction and sequencing

Genomic DNA was extracted from samples of 120 female flies collected from each of the 30 individual populations ($ACO_{1\text{-}5}$, $AO_{1\text{-}5}$, $CO_{1\text{-}5}$, $nCO_{1\text{-}5}$, $B_{1\text{-}5}$, $BO_{1\text{-}5}$) using the Qiagen/Gentra Puregene kit, following the manufacturer's protocol for bulk DNA purification. The 30 gDNA pools were prepared as standard 200-300 bp fragment libraries for Illumina sequencing, and constructed such that each 5 replicate populations of a treatment (e.g., $ACO_{1\text{-}5}$) were given unique barcodes, normalized, and pooled together. Each 5-plex library was run on individual PE100 lanes of an Illumina HiSEQ 2000 at the UNC High Throughput Sequencing Facility. Resulting data were 100 bp paired-end reads. Each population was sequenced twice; data from both runs were combined for some analyses as described below. Combining reads from two independent sequencing runs likely alleviate the effects of possible bias introduced from running all replicates for each population in the same lane. Moreover, preliminary comparisons between analyses resulting from only the

first run and analyses from the combined data did not yield substantive differences in our results (data not shown).

Single Nucleotide Polymorphism (SNP) analysis

Read mapping and pre-processing

We first trimmed the reads to remove low quality bases using a script provided in the PoPoolation software package (Kofler et al. 2011a). We then mapped reads with BWA (version 0.7.6) (Li and Durbin 2009) against the D. melanogaster reference genome (release 5.51) with the following mapping parameters: -n 0.01 (error rate), -o 2 (gap opening), -d 12 and -e 12 (gap length), and -l 150 to effectively disable the seed option. We then converted the resulting alignment files to SAM format using the BWA sampe command. We filtered the SAM files for reads mapped in proper pairs with a minimum mapping quality of 20 and converted them to the BAM format using SAMtools (Li et al. 2009). The rmdup command in SAMtools was then used to remove potential PCR duplicates. The two BAM files from each population's two sequencing runs were merged using BAMtools to maximize coverage (Barnett et al. 2011). These merged BAM files were then all combined in the mpileup format once again using SAMtools. Using PoPoolation2 (Kofler et al. 2011b), the resulting mpileup was converted to a "synchronized" files, which is a format that allele counts for all bases in the reference genome and for all populations being analyzed. We used RepeatMasker 4.0.3 (http://www.repeatmasker.org) to create a gff file to mask low complexity regions of the D. melanogaster genome version 5.51.

Heterozygosity analysis

We calculated and plotted heterozygosity across the five major chromosome arms to see if we could find any evidence of selective sweeps and to determine if there was

convergence in overall patterns of variation. To do this, SNPs were first called across all 30 populations used in this study from our synchronized file. SNPs where discarded if coverage in any of the populations was less than 20X or greater than 500X. We also required a minimum minor allele frequency of 2% across all eight populations. Based on these parameters, ~1.01 million SNPs were identified across the major chromosome arms. The average SNP coverage at each across our 30 populations ranged from 28X to 108X, with all but two of our populations (CO_5 and B_3 at 28X and 31X respectively) and having coverage greater than 30X (Table S2.3 for more detailed coverage information). A SNP table with major and minor allele counts for each SNP in each population was then generated. Using these counts, heterozygosities were calculated and plotted over 100kb non-overlapping windows. We also performed t-tests comparing mean genome-wide heterozygosities between different groups of populations.

We also took steps to identify potential bias in our SNP calling procedure given that average coverage varied across the 30 populations. In particular, the CO_5 and B_3 sequence data have much lower average coverage than the other populations. To test for possible bias, we repeated our SNP calling procedure with the added exception that a minimum coverage of 20 was relaxed to 10 for the CO_5 and B_3 populations; it was unchanged for the remaining 28 populations. This ultimately resulted in an additional ~300 thousand SNPs being identified, suggesting that the inclusion of these low coverage populations combined with our minimum coverage requirement of 20 is indeed a source of bias. However, as allele frequency estimates suffer at low coverages, we chose to be more conservative and maintain this requirement for SNPs included in our analyses.

F_{ST} estimates

F_{ST} estimates for replicate populations were obtained using the formula: $F_{ST}= (H_T-H_S)/H_T$ where H_T is heterozygosity based on total population allele frequencies, and H_S is the average subpopulation heterozygosity in each of the replicate populations (Hedrick, 2009). F_{ST} estimates were made at every polymorphic site in the data set for a given set of replicate populations. This was done to quantify the level of similarity between replicates of our 6 sets of populations. The mean genome wide F_{ST} for the 6 sets of populations are reported in supplementary table S7.

SNP differentiation

PoPoolation2 was used to obtain measures of SNP differentiation between newly derived and long standing populations within and between treatments (Kofler et al. 2011b). More specifically, we used this software package to perform Cochran-Mantel-Haenzsel tests of differentiation to compare SNP frequencies between our various groups of replicate populations. The specific comparisons performed were as follows: ACO vs AO, ACO vs B, AO vs BO, ACO vs CO, AO vs nCO, B vs BO, B vs. CO, BO vs nCO, CO vs. nCO, A vs C (all A-types versus all C-types), A vs B, and B vs C. For each comparison, relevant populations were paired by their replicate number (e.g. ACO1 with AO1, ACO2 with AO2, and so on), which reflects patterns ancestry in the case of all but the five B populations.

The CMH test was performed at each site polymorphic across our 30 populations. Results for all positions not found in our SNP table were then discarded; thus all positions failing to meet our SNP calling criteria as stated above were removed. To generate null distributions for p-values generated by each comparisons (i.e., distributions of these p-values associated with a null expectation of genetic drift rather than selection), we used a

permutation approach. For a given comparison, the relevant populations were randomly assigned to one of two groups and the CMH test was then performed at each polymorphic position in the shuffled data set. This was done a 1000 times and the smallest p-value was recorded. The quantile function in R was then used to define thresholds that define the genome-wide false-positive rate, per site, at 5%. This was done for each the 12 comparisons we performed.

To give a concrete example of the permutation procedure, we will focus on the $ACO_{1\text{-}5}$ vs $CO_{1\text{-}5}$ comparison. For a given permutation, 5 of the 10 populations were randomly assigned to group X and the remaining to group Y. A permuted data set could look something like group X = CO_1, ACO_2, ACO_3, CO_3, CO_4 and group Y = CO_5, ACO_1, ACO_4, ACO_5, CO_2. When the CMH test is run, the populations would be paired based on their order in groups X and Y (eg. CO_1-CO_5, ACO_2-ACO_1, ACO_3-ACO_4, etc.). Given that the pairings matter and we have 10 populations, this gives 3,628,800 possible permutations. For each of the shuffled data sets, the CMH test was performed between groups X and Y at each polymorphic position and the smallest p-value generated across the entire genome was recorded. Once we had a list of the smallest p-values generated across 1000 permutations, we used the quantile function in R to define thresholds that give a genome-wide false-positive rate, per site, of 5%. That is, our test gives us a Type I error rate of 0.05 for each and every site considered significantly differentiated. Note that this test is not designed to minimize Type II errors across the genome, making it undoubtedly statistically conservative.

Linkage Disequilibrium (LD) analysis

We calculated linkage disequilibrium using LDx, a method which uses an approximate maximum likelihood approach to estimate LD (r^2) from pooled resequencing data (Feder et al. 2012), using HPC resources provided by the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. To view the complete scripts used for this analysis, see https://github.com/k8hertweck/flyPopGenomics. To prepare data for analysis in LDx, we called SNPs on each merged bam file using mpileup and filtered the output using default options (samtools mpileup -uIf dmel.RELEASE5 *.bam | bcftools view -v snps | vcfutils.pl varFilter > *.flt.vcf). We than ran LDx using default options, except for adjusting the insert size to 200 and minimum read depth to 20 (perl LDx.pl -l 20 -h 100 -s 200 -q 20 -a 0.1 -i 11 *.sam *.flt.vcf > *.flt.out). Because LDx compares SNPs within read pairs, we only report LD calculations for distances of 200 bp or less. Distances less than 11 bp were also discarded because of small sample sizes.

To evaluate differences in patterns of LD decay between our populations, we fit the data to a biexponential model in R using the "SSbiexp" function and the "nlme" package (Pinheiro et al. 2015). Data from each chromosome (2, 3 and X) was handled independently. We chose to use a biexponential model after evaluating quadratic, cubic, and quartic models. To do this, we split the data into a training set (80% of the data) and a test set (remaining 20%). We then calculated the mean squared error (MSE) for each model and found that the biexponential model had the lowest MSE.

We initially partitioned variation in LD into the fixed effects of selection regime and whether or not the populations were long standing (ACO,B, or CO) or newly derived (AO, BO, nCO) and the random effects of population. However, we found that whether or not the

populations were long standing or newly derived did not have an effect on parameter estimates. So this was dropped from the model. The random effects over populations are due to both sampling and genetically-based differences that arise due to genetic drift. We tested models with population variation in subsets of parameters and with a constant within-population variation. The model chosen had the lowest Akaike and Bayesian information criterion (Pinheiro and Bates 2000).

For each chromosome, we constructed confidence intervals based on model predictions with a coverage level that applied to all observed points. For each population we had maximum likelihood parameters estimates and their covariance matrix estimates, which were assumed to have normal distributions. From these distributions, we drew samples of the parameter vectors,$\tilde{\mu}_i$, (i= 1,…,m). For each sampled parameter vector, we made predictions $f(\tilde{\mu}_i, t)$, for all values of t. From these m predictions, we generated order statistics, $f^j(\tilde{\mu}_i, t)$, where $f^1(\tilde{\mu}_i, t)$ is the smallest predicted value at t and $f^m(\tilde{\mu}_i, t)$ is the largest. If there are k-points that we want to include in a simultaneous interval, then from the Bonferroni inequality (Miller 1966) the confidence level is elevated to 1-0.05/k. From the order statistics, we then used $f^l(\tilde{\mu}_i, t)$ as the lower confidence limit and $f^u(\tilde{\mu}_i, t)$ as the upper confidence limit where, l = round (m0.05/(2k)) and u=(m+1-l). For our purposes, m= 10,000 and k= 38. We have LD estimates for distances of 10 to 200bp, but here we only used distances that were multiples of 5.

RESULTS

Phenotypic analysis

We performed assays of fecundity and survival during the same generations as those used to collect samples for DNA sequencing, testing for functional parallelism

between long-standing and newly-derived populations (kx, Figure 2.1C-E; for lx, mx see Figure S2.1, see supplementary Tables S2.1 and S2.2 for p-values). We found statistically significant phenotypic differentiation between groups of populations sharing their most recent selection regime for just one fitness-character: newly-derived nCO populations had superior reproductive output to long-standing CO populations during a single time interval (Figure. 2.1E). Overall, the fitness-character results show a lack of evolutionary differentiation between long-standing and newly-derived populations sharing the same distal selection regime. This rapid loss of differentiation is consistent with previous Drosophila studies on the experimental evolution of functional characters (Teótonio and Rose 2000; Rose et al. 2004; Fragata et al. 2014).

Patterns of genetic variation

Given the lack of phenotypic differentiation with shared recent selection regimes, we aimed to identify potential underlying genomic factors that similarly lack differentiation. Our average sequencing coverage varied between populations, but was ≥ 50X for all populations except CO_5 (25X) and B_3 (29X) (Table S2.3). Genomic analysis focused on single nucleotide polymorphism (SNP) variation. We identify ~1.01million SNPs across the major chromosome arms.

Extensive heterozygosity is maintained in our populations, despite hundreds of generations in the laboratory, as previously found by Burke et al. (2010) for the five ACO populations and the five CO populations. In terms of heterozygosity calculated from our SNP data, we find few large regions where genetic variation has been depleted when heterozygosity is calculated over 100kb windows (Figure 2.2). Regions where heterozygosity has been reduced to low levels are predominately found in the A-type

populations. This result is robust to reductions in window size (see Figures S2.2 and S2.3 for 50k and 30kb results). Mean heterozygosity is also lower in populations subjected to A-type selection (Table S2.4), which could be due to more intense A-type selection. For heterozygosity calculated from our SNP data, these differences were found to be statistically significant based on t-tests comparing mean genome wide heterozygosity between groups of populations (see Table S2.5 for p-values). We find that more variation is maintained in C-type populations than both A-type and B-type populations, with more in B-type populations than A-type.

As for patterns of similarity in variation across replicates, we find a high degree of similarity between replicate populations with parallel evolutionary histories, as indicated by mean genome wide F_{ST} estimates that are all less than 0.10 from SNP data (Table S2.6). Given the range of these values (0.041-0.087), we suggest that Wright (1978) would have described the variation between replicates within each selective treatment as "small". This pattern is recapitulated by visual examination of heterozygosity for replicate populations (Figure S2.4). We also display this finding by plotting the variance in allele frequencies, per SNP, for each of the selection treatments (Figure S2.5). Observed variances in raw SNP frequencies are very low across all treatments, implicating parallel evolution of standing variants among the five replicates of each treatment. While we do not find any evidence for widespread reductions in SNP heterozygosity across our 100 kb windows, there are a number of local depressions in heterozygosity in all of our populations. We also find many localized depressions in SNP heterozygosity, defined as 100kb windows with heterozygosity less than 0.2, which are consistent across given groups of replicate populations (Tables S2.7-S2.12). More regions like this are found in our populations

subjected to A-type selection (ACO: 105 regions, AO: 161 regions), than populations subjected to B-type (B: 59, BO: 11) and C-type selection (CO: 5, nCO: 8). This is suggestive of selection acting at parallel at sites within these regions (vid. Oleksyk et al. 2010). This result is consistent with other Drosophila work showing that adaptation in experimentally-evolved populations is driven by selection on standing genetic variation (Burke et al. 2010; Turner et al. 2011; Orozco-ter Wengel et al. 2012; Tobler et al. 2014).

Previous work with these populations sequenced only ACO and CO populations, and pooled DNA across replicates such that direct observations of parallelism within evolutionary histories at the nucleotide level were not possible (Burke et al. 2010). Burke et al. (2010) did sequence a single replicate population (ACO_1) and allele frequencies in this single replicate were highly similar to allele frequencies in the entire pool of ACO flies. The present work corroborates this observation and expands upon it considerably. While widespread migration between replicates of each selection treatment would also produce the patterns we observe here, we contend that evolutionary parallelism at the SNP level is the more likely underlying scenario. For further discussion of how our results compare to those from Burke et al. (2010), see Figure S2.6.

SNP differentiation

To formally and systematically assess levels of differentiation between the six groups of five-fold replicated populations, we performed Cochran-Mantel-Haenszel (CMH) tests comparing SNP frequencies between and among groups that shared distal selection regimes. We find lower levels of differentiation among newly-derived and long-standing populations recently subjected to the same type of selection, versus comparisons involving

populations recently subjected to different types of selection as indicated by more significantly differentiated variants in the latter (Figures S2.3 and S2.4, and Figure S2.7).

We performed nine statistically-conservative comparisons of SNP differentiation between populations that did not share recent selection regimes. In six of these tests, dozens to hundreds of significant differentiated SNPs are found (Figure 2.3, Figure. S2.7, Table 2.1). The two comparisons between groups of five B-type and five C-type groups did not yield significantly differentiated SNPs. We do, however, find significantly differentiated SNPs when we compare all 10 B-type populations with all 10 C-type populations (Figure 2.3). By contrast, across the three comparisons of populations that shared a recent selection regime (ACO vs. AO; B vs. BO; CO vs. nCO), but had different evolutionary histories, we detect no significantly differentiated SNPs (Figure 2.4).

In our between-treatment comparisons, we find more significantly differentiated SNPs in our comparisons between long-standing populations (e.g. ACO vs CO) versus comparisons between newly-derived populations (e.g. AO vs nCO) (Figure S2.7 and Table 2.1). This finding is consistent with previous simulations of experimental evolution in sexual populations, which suggest that increasing number of generations under selection increases the power to detect SNPs underlying responses to directional selection (Baldwin-Brown et al. 2014; Kofler and Schlötterer 2014).

When we increase replication by treating our longstanding and newly-derived populations that share a recent section regime as equivalent replicates (e.g. all ten A-types, merging ACO with AO populations), we detect more significantly differentiated SNPs. This finding is again consistent with previous simulations of experimental evolution, which demonstrate a strong impact of varying the number of evolving replicates on the

experimental detection of SNPs underlying responses to directional selection (Baldwin-Brown et al. 2014; Kofler and Schlötterer 2014). It is also consistent with the results of a similar analysis of yeast experimental evolution (Burke et al. 2014).

Linkage disequilibrium

A key question in studies of sexual laboratory populations with low to moderate effective population sizes is level of linkage disequilibrium (LD). If there are high levels of LD, the number of genomic sites targeted by selection may be far smaller than the number of SNPs exhibiting statistically significant differentiation. We characterized linkage in our populations using LDx (Feder et al. 2012) and fit these estimates to a biexponential model. Populations subjected to different selection regimes had statistically distinct patterns of LD decay, but newly-derived and long-standing populations subjected to the same selection did not (Figure 2.5). Furthermore, LD was higher in A-type populations compared to B and C-type populations. This pattern cannot be explained by number of generations, because long-standing ACO populations had many more generations for recombination to break down LD compared to C-type and BO populations. In fact, LD was lowest in the C-type populations, which have the fewest generations under laboratory cultivation. It appears that LD (like SNP, TE, and SV frequency differentiation) is highly parallel, depending primarily on recent selection history. LD patterns may be determined by relative intensities of selection at loci across the entire genome. In particular, consistently higher levels of LD in the A-type populations may be due to more intense selection, as suggested by the low number of individuals surviving to reproduction during this type of selection's initial generations. However, formal simulation would be required to support this

conjecture, as we lack the haplotype sequencing data required to characterize long range LD (cf. Franssen et al. 2015).

DISCUSSION

The results of our functional and genomic analyses show that experimental evolution of moderately outbred Drosophila involves reproducible and extensive changes across the fruit fly genome. Marked functional and genomic differentiation between newly-derived populations that do not share recent selection regimes is comparable to that found between long-standing populations that do not share selection regimes, despite large disparities in number of generations under divergent selection. This finding supports the contention that convergent, or "parallel" depending on favored usage, evolution can result from polymorphisms that were shared in an ancestral population (cf. Stern, 2013), as all selected lines in this study were derived from a common ancestral population (Rose et al. 2004; Burke et al. 2016).

This consistent pattern across six evolutionary histories supports the hypothesis that outbred sexual populations rapidly respond to selection in a reproducible manner, because they maintain functional genetic variation at many sites across their genomes. We observe clear evolutionary differentiation in response to selection that is not associated with complete local elimination of genetic variation near sites of genomic change. This provides evidence that evolutionary change in sexual populations is not driven by hard selective sweeps on newly arisen beneficial mutations of large effect (Maynard Smith and Haigh 1974; Burke 2012), even over periods of more than 800 generations. These dynamics differ from those observed in long-term evolution experiments with E. coli, which are dominated by hard selective sweeps, clonal interference, and clonal replacement (e.g.

Maddamsetti et al. 2015), and which also find different alleles driving change in different replicate populations (Woods et al. 2006, Tenaillon et al. 2012). Similar results to ours have surfaced repeatedly in genomic studies of Drosophila lab evolution (Turner et al. 2011, Orozco-ter Wengel et al. 2012; Tobler et al. 2014), though those studies featured notably less replication than our study, fewer generations of selection, and notably less-extensive genomic responses to selection. We suggest that the genomic foundations for the experimental evolution of outbred sexual populations are different in kind from those of strictly clonal paradigms of experimental evolution (cf. Barrick et al. 2009, Tenaillon et al. 2012; Maddamsetti et al. 2015), in which selective sweeps by newly-arisen mutants are the chief determinants of adaptation. The scale of our study underscores the following conclusions about the evolution of outbred sexual populations: (i) their evolution can be fast; (ii) their evolution can be repeatable from nucleotide to fitness, thanks to standing genetic variation; (iii) their adaptation does not wait for new functional mutations and subsequent hard selective sweeps.

It might be proposed that a single haplotype is the target of selection for each of the A, B, and C selection regimes. But if that were the case, the rapid convergence on A, B, and C phenotypes and genotypes by the recently derived A, B, and C lines (AO, BO, and nCO, respectively) suggests that the selection coefficients associated with these three selection regimes are large in magnitude. That in turn requires that the 15 populations long-subjected to these three selection regimes (ACO, B, and CO) should be approaching fixation at many sites across the genome, a pattern that is not apparent in the heterozygosity data for the B and CO populations, at least. Therefore, we suggest that our results indicate that most of the genomic sites under selection in our study are undergoing a shift between

stable balanced polymorphisms, not selective sweeps toward fixation of a single genome-wide haplotype.

Any laboratory evolution study faces the challenging question of its applicability to the evolution of real-world populations in nature. Our point of view on this question is that we do not think that it is likely that sexual populations in nature generally undergo stable selection regimes for up to 1,000 generations. Thus our study is not intended to directly model what occurs in natural populations. Instead, our study is intended to use extreme and unusual evolutionary histories to test competing evolutionary genetic hypotheses. Most such hypotheses are hard to test using the shifting, short-term, and ambiguous data available from most studies of evolution in nature. Our strategy has long been based on strong-inference experimentation (cf. Platt, 1966), which requires laboratory control of culture regimes over long periods, as in the classic LTEE study of Lenski et al. (1991). Thus, for example, if hard selective sweeps were the predominant mode of adaptation to the selection regimes we used, then the 15 populations that featured up to 1,000 generations of sustained selection should have produced numerous genomic regions with negligible heterozygosity. As we did not observe such patterns, we consider this a refutation of the hypothesis that adaptation predominantly proceeds by hard selective sweeps in sexual populations.

References

Arendt J, Reznick D. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. 23(1):26-32.

Baldwin-Brown JG, Long AD, Thornton KR. 2014. The power to detect quantitative trait loci using resequenced, experimentally evolved populations of diploid, sexual organisms. Mol. Biol. Evol. 31(4):1040–1055.

Barnett DW, Garrison EK, Quinlan AR, Stromberg MP, Marth GT. 2011. BamTools: A C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 27:1691-1692.

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. 2009. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature 461:1243–1274.

Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. 2010. Genome-wide analysis of a long-term evolution experiment with Drosophila. Nature 467:587–590.

Burke MK. 2012. How does adaptation sweep through the genome? Insights from long-term selection experiments. Proc Roy Soc B. 279:5029–5038.

Burke MK, Liti G, Long AD. 2014. Long standing genetic variation drives repeatable experimental evolution in outcrossing populations of Saccharomyces cerevisiae. Mol Biol Evol. 32:3228-3239.

Burke MK, Barter TT, Cabral LG, Kezos JN, Phillips MA, Rutledge GA, Phung KH, Chen RH, Nguyen HD, Mueller LD, Rose MR. 2016. Rapid convergence and divergence of life-history in experimentally evolved Drosophila Melanogaster. Evolution 70:2085-2098.

Elias M, Tawfik DS. 2012. Divergence and convergence in enzyme evolution: parallel evolution of paraoxonases from quorum-quenching lactonases. J. Biol Chem. 287(1):11-20. doi: 10.1074/jbc.R111.257329.

Feder AF, Petrov DA, Bergland AO. 2012. LDx: estimation of linkage disequilibrium from high-throughput pooled resequencing data. PLoS ONE. 7:doi:10.1371/journal.pone.0048588.

Fragata I, Simoes P, Lopez-Cunha M, Lima M, Kellen B, Barbaro M, Santos J, Rose MR, Santos M, Matos M. 2013. Laboratory selection quickly erases historical differences. PLoS ONE 9:e96227 10.1371/journal.pone.0096227

Franssen SU, Volte N, Tobler R, Schlötterer C. 2015. Patterns of linkage disequilibrium and long range hitchhiking in evolving experimental Drosophila melanogaster populations. Mol Biol Evol. 32:495-509.

Ives, PT. 1970. Further studies of the South Amherst population of Drosophila melanogaster. Evolution 24: 507—518.

Kofler R, Orozco-terWengel P, De Maio N, Pandey RV, Nolte V, Futschik A, Kosiol C, Schlötterer C. 2011a. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS ONE 6:e15925

Kofler R, Pandey RV, Schlötterer C. 2011b. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27:3435-3436.

Lenski RE, Rose MR, Simpson SC, and Tadler SC. 1991. Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations. Amer. Nat. 138:1315-1341.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment / map format and SAMtools. Bioinformatics 25:2078-2079.

Maddamsetti R, Lenski RE, Barrick JE. 2015. Adaptation, Clonal Interference, and Frequency-Dependent Interactions in a Long-Term Evolution Experiment with Escherichia coli. Genetics 200: 619-631.

Matos M, Simoes P, Duarte A, Rego C, Avelar T, Rose MR. 2004. Convergence to novel environment: Comparative method versus experimental evolution. Evolution 58(7):1503-1510.

Matos M, Avelar T. 2001. Adaptation to the laboratory: Comments on Srgo and Partride. Am Nat. 158:655-656.

Matos M, Avelar T, Rose MR. 2002. Variation in the rate of convergent evolution: adaptation to the laboratory environment in Drosophila subobscura.J Evol Biol. 15:673-683.

Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favorable gene. Genet Res. 23:23-35.

Miller RG. 1966. Simultaneous statistical inference. New York: McGraw-Hill.

Muller LD, Joshi A, Santos M, Rose MR. 2013. Effective population size and evolutionary dynaics in outbred laboratory populations of Drosophila. J Genet. 92:349-361.

Oleksyk TK, Smith MW, O'Brien SJ. 2010. Genome-wide scans for footprints of natural selection. Philos Trans R Soc Lond. 365:185-205.

Orozco-terWengel P, Kapun M, Nolte V, Kofler R, Flatt T, Schlötterer C. 2012. Adaptation of Drosophila to a novel laboratory environment reveals temporally heterogeneous trajectories of selected traits. Mol Ecol. 21: 4931-4941.

Pinheiro JC, Bates DM. 2000. Mixed-effects models in S and S-PLUS. New York: Springer.

Pinheiro JC, Bates DM, DebRoy S, Sarkar D, R Core Team. 2015. nlme: linear and nonlinear mixed effects models. R pckage version 3.1-122 http://CRAN.R-project.org/package=nlme.

Platt JR. 1964. Strong inference. Science 146:347-353.

Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. 2006. Tests of parallel molecular evolution with Escherichia coli. Proc Natl Acad Sci. 103: 9107-9112.

Rose MR, Passananti HB, Matos M. 2004. Methuselah Flies: A case study in the evolution of aging. Singapore: World Scientific Publishing.

Schlötterer C, Kofler R, Versace E, Tobler R, Franseen SU. 2015. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. Herdity 114:331-440.

Stern DL. 2013. The genetic causes of convergent evolution. Nature Review: Genetics 14:751-764.

Tenaillon O, Rodriguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, Gaut BS. 2012. The molecular diversity of adaptive convergence. Science 335:457-461.

Rose MR and Teotónio H. 2000. Variation in the reversibility of evolution. Nature 408:463-466.

Tobler R, Franssen SU, Kofler R, Orozco-terWengel P, Nolte V, Hermisson J, Schlötterer C. 2014. Massive habitat-specific genomic response in D. melanogaster populations during experimental evolution in hot and cold environments. Mol Biol Evol. 31:364-375.

Turner TL, Steward AD, Fields AT, Rice WR, Tarone AM. 2011. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in Drosophila melanogaster. PLoS Genet. 7:e10001336.

Wright, S. 1978. Evolution and Genetics of Populations. Volume 4: Variability Within and Among Natural Populations. (Chicago and London: The University of Chicago Press).

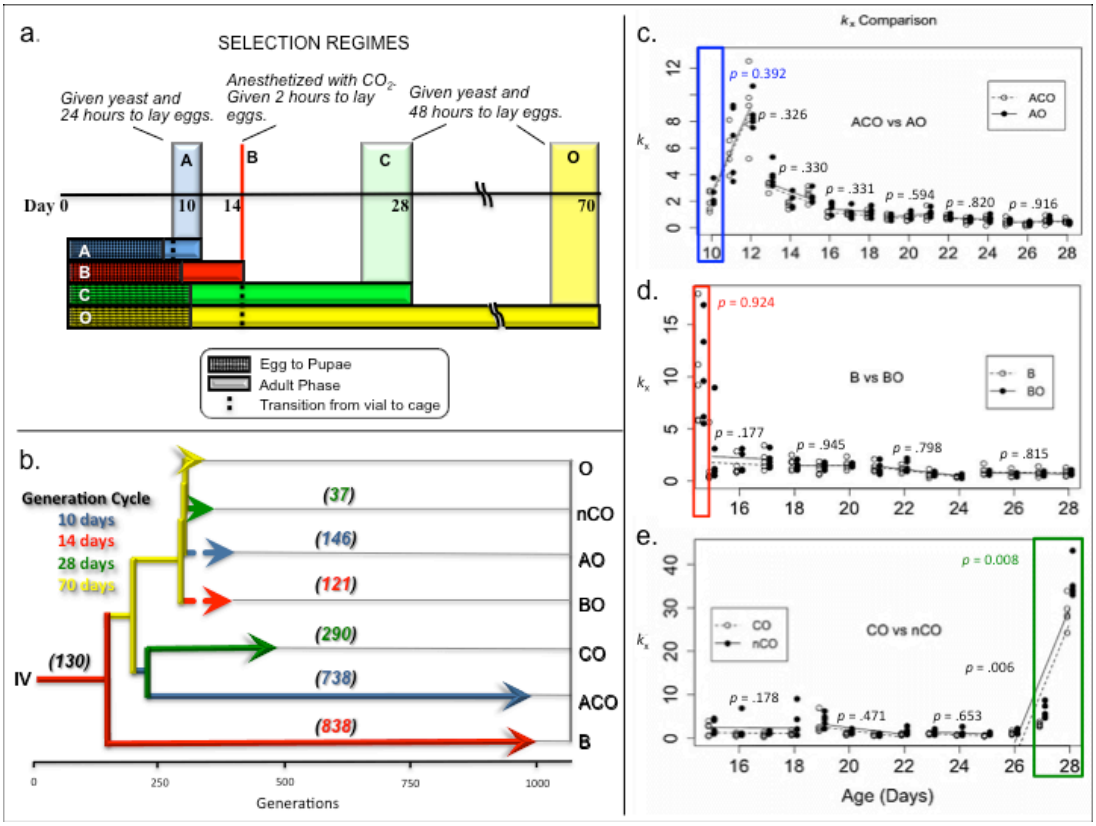

Figure 2.1. Experimental design, fecundity, and survival. (A) protocols for selection regimes, (B) evolutionary history of experimental populations and number of generations under selection regime for six groups of five-fold replicated populations. (C-E) comparison of kx (kx = lx mx, where lx is probability of survival to age x and mx is fecundity at age x) between long-standing and newly derived populations over a 19 day interval.

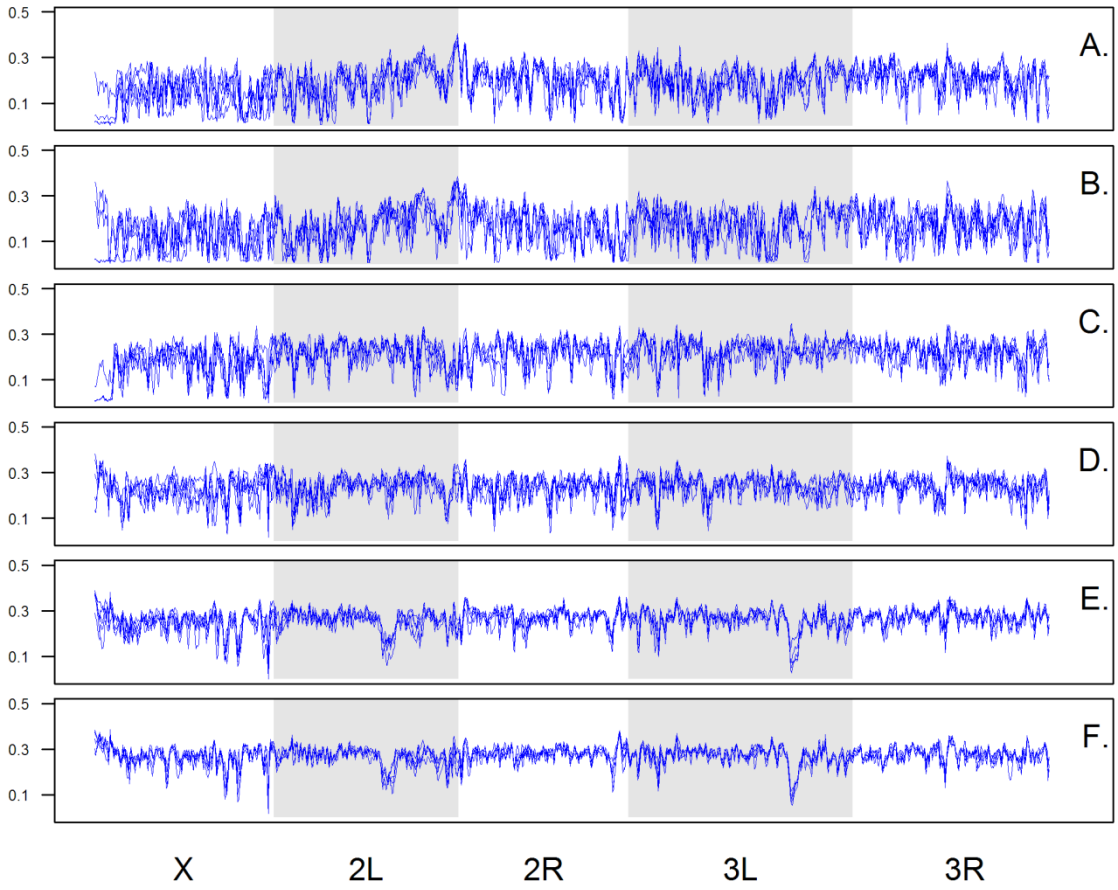

Figure 2.2. **Genome-wide heterozygosity 100kb windows.** Heterozygosity calculated from SNP data plotted across 100kb non-overlapping windows across all major chromosome arms for $ACO_{1\text{-}5}$ (A), $AO_{1\text{-}5}$ (B), $B_{1\text{-}5}$ (C), $BO_{1\text{-}5}$ (D), $CO_{1\text{-}5}$ (E), $nCO_{1\text{-}5}$ (F). Each panel shows results from all five replicate populations plotted individually.

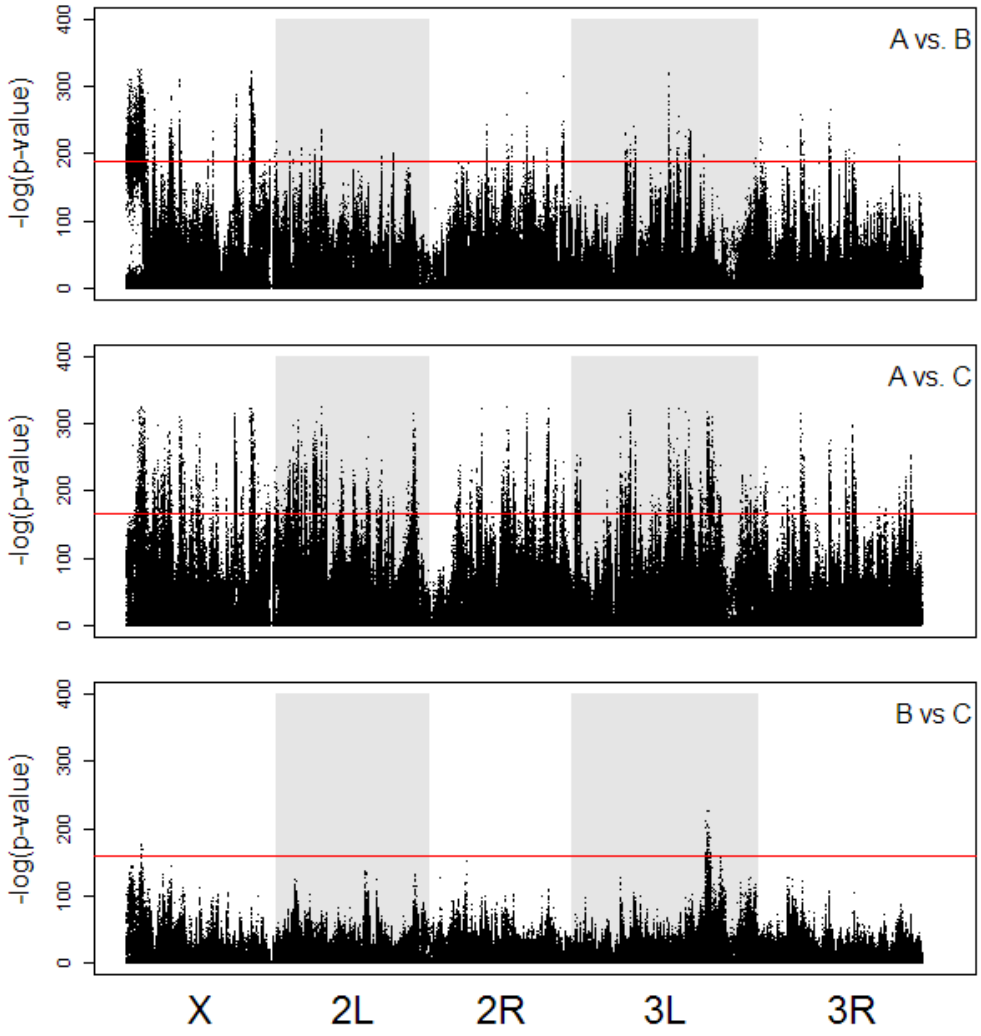

Figure 2.3. CMH test results for between selection regime comparisons. Results from CMH tests comparing SNP frequencies between populations that are from different selection treatments. For these comparisons, we grouped populations based purely on their most recent selection regime. For example, we compared all ten A-types (ACO and AO populations) with all ten B-types (B and BO populations). Results are plotted as –log(p-values) across all major chromosome arms. Significance thresholds for each comparison are derived from permutation tests and are indicated by a red line.

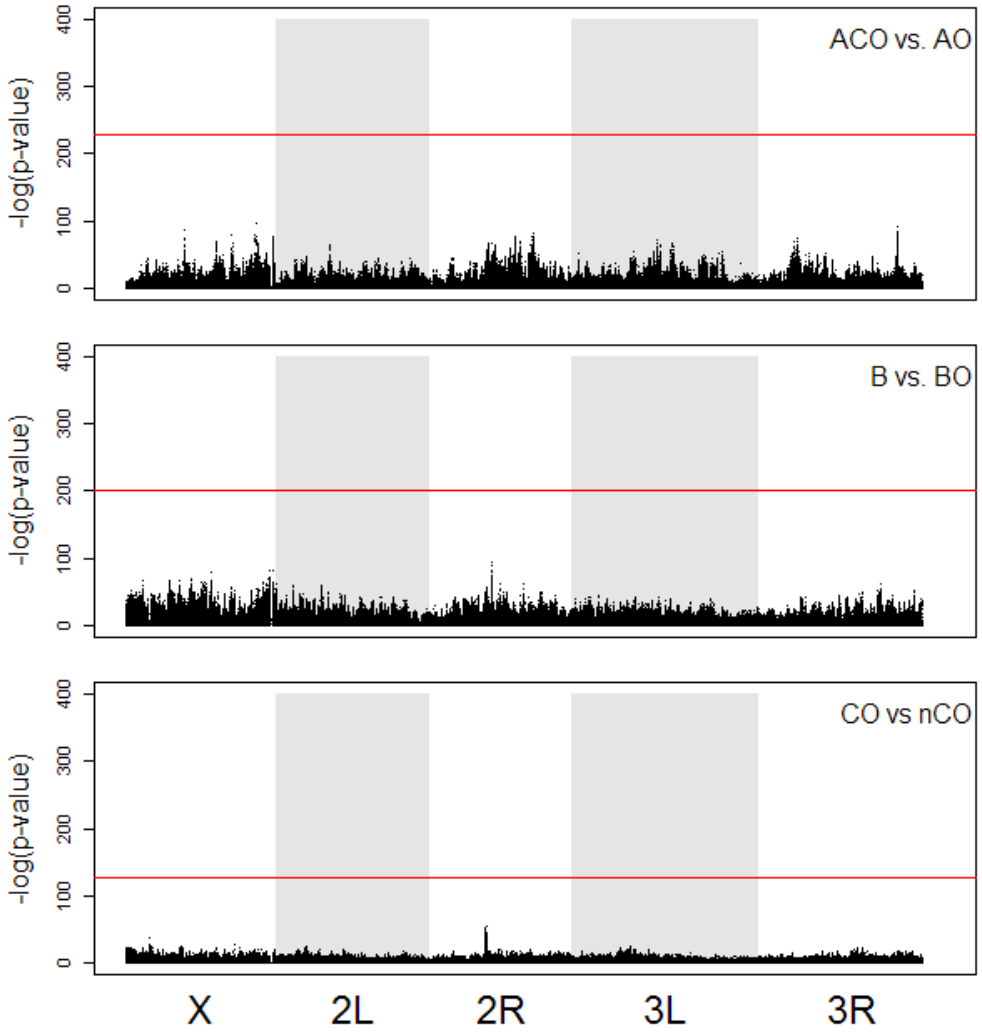

Figure 2.4. CMH test results for within selection regime comparisons. Results from CMH tests comparing SNP frequencies between populations that share a recent selection regime. Results are plotted as –log (p-values) across all major chromosome arms. Significance thresholds for each comparison are derived from permutation tests and are indicated by a red line.

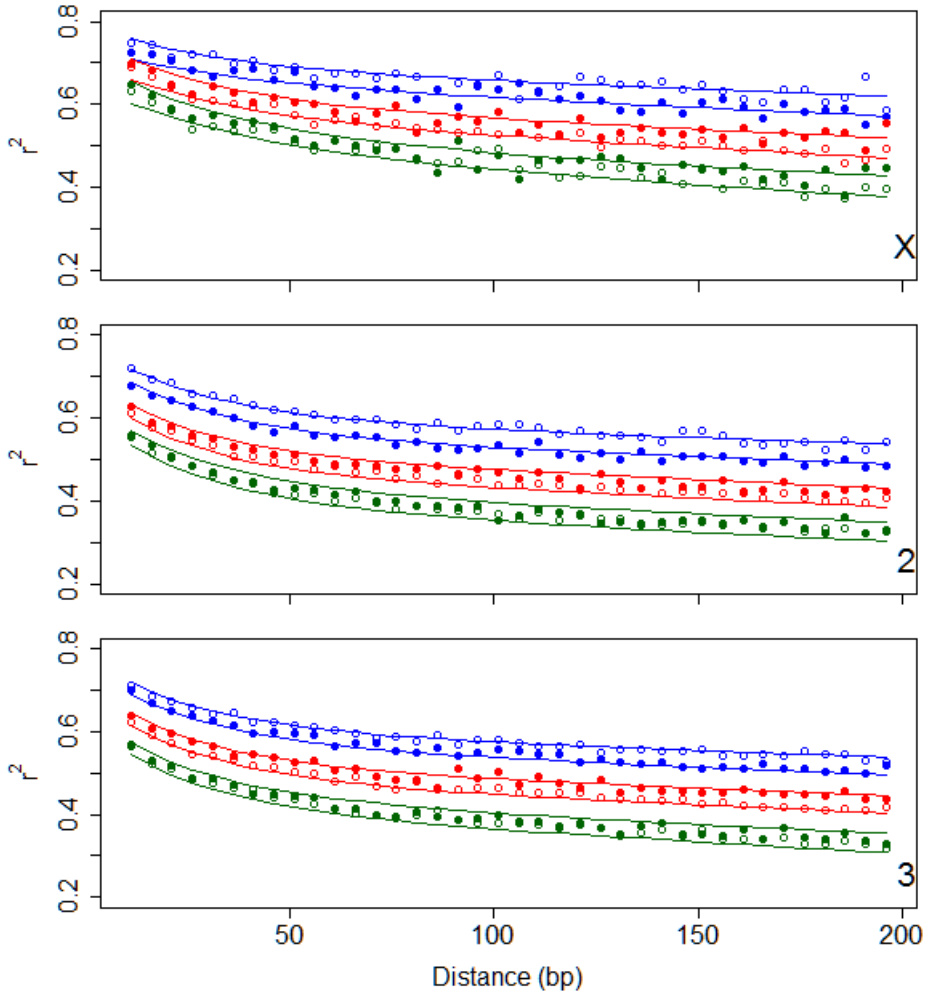

Figure 2.5. LD results. Mean linkage disequilibrium (r2) estimates from LDx for chromosomes X, 2, and 3, plotted against distance from a focal SNP for six groups of populations. Closed circles are estimates for long-standing populations (ACO, B, and CO); open circles are estimates for newly derived populations (AO, BO, and nCO). A, B, and C-type populations are blue, red, and green respectively. Colored lines represent simultaneous confidence intervals based on predictions from the biexponential model to which the data were fitted.

Table 2.1. Counts of significantly differentiated SNPs from CMH tests comparing frequencies between and among selective regimes.

Comparison	Significance Threshold*	Number of Significant SNPs
ACO vs CO	5.67 x 10-145	412
AO vs nCO	1.64 x 10-166	143
A vs C	8.21 x 10-167	10109
ACO vs B	1.25 x 10-158	1146
AO vs BO	2.77 x 10-149	42
A vs B	8.56 x 10-189	4152
CO vs B	1.11 x 10-183	0
nCO vs BO	1.78 x 10-120	0
B vs C	6.40 x 10-159	64
ACO vs AO	2.91 x 10-230	0
B vs BO	2.18 x 10-201	0
CO vs nCO	1.70 x 10-126	0

*Significance thresholds for each test were determined based on our permutation approach to correcting for multiple comparisons.

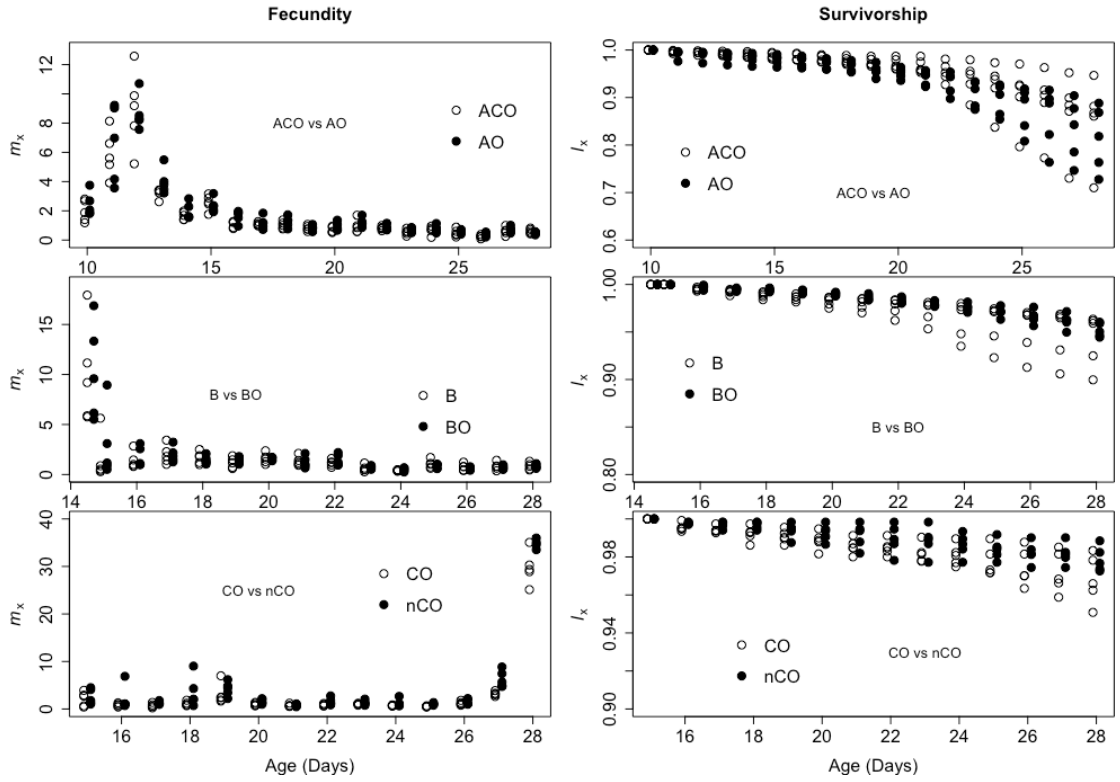

Figure S2.1: Fecundity and survivorship for all 30 Drosophila populations. Fecundity (mx) is measured as average eggs laid per surviving female and survivorship (lx) is measured as percent of each population alive at a given day.

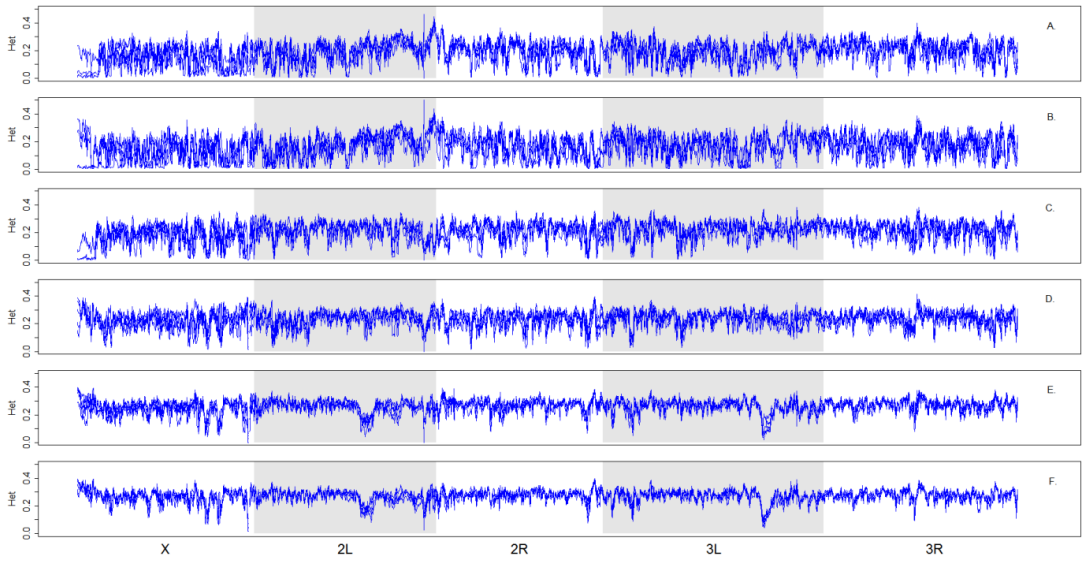

Figure S2.2. Genome-wide heterozygosity 50kb windows. Heterozygosity calculated from SNP data plotted across 50kb non-overlapping windows across all major chromosome arms for ACO1-5 (A), AO1-5 (B), B1-5 (C), BO1-5 (D), CO1-5 (E), nCO1-5 (F). Each panel shows results from all five replicate populations plotted individually.

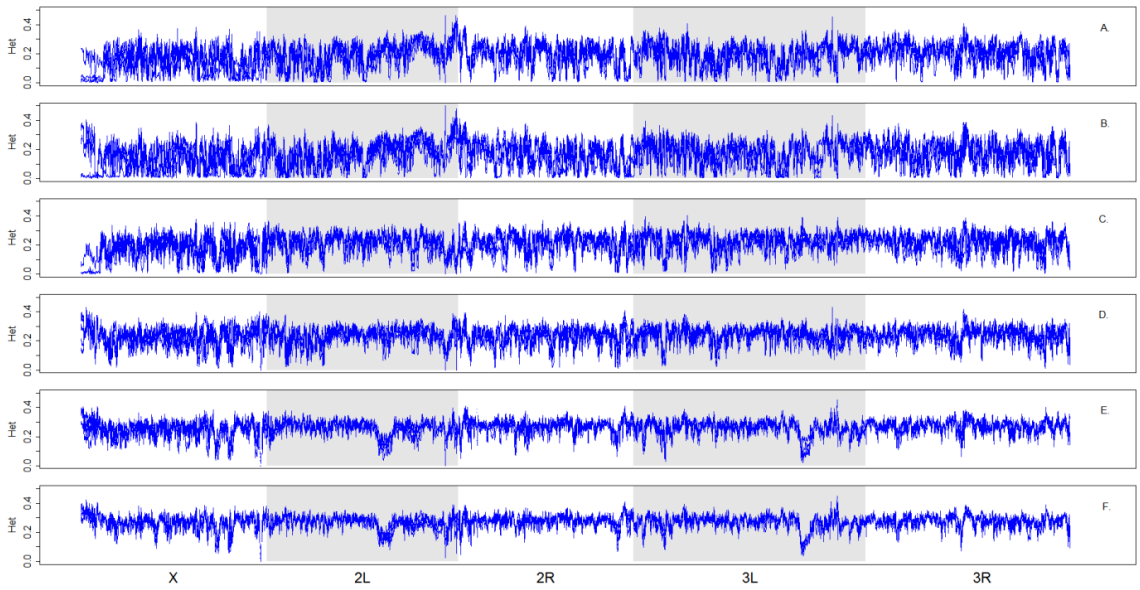

Figure S2.3. Genome-wide heterozygosity 30kb windows. Heterozygosity calculated from SNP data plotted across 30kb non-overlapping windows across all major chromosome arms for ACO1-5 (A), AO1-5 (B), B1-5 (C), BO1-5 (D), CO1-5 (E), nCO1-5 (F). Each panel shows results from all five replicate populations plotted individually.

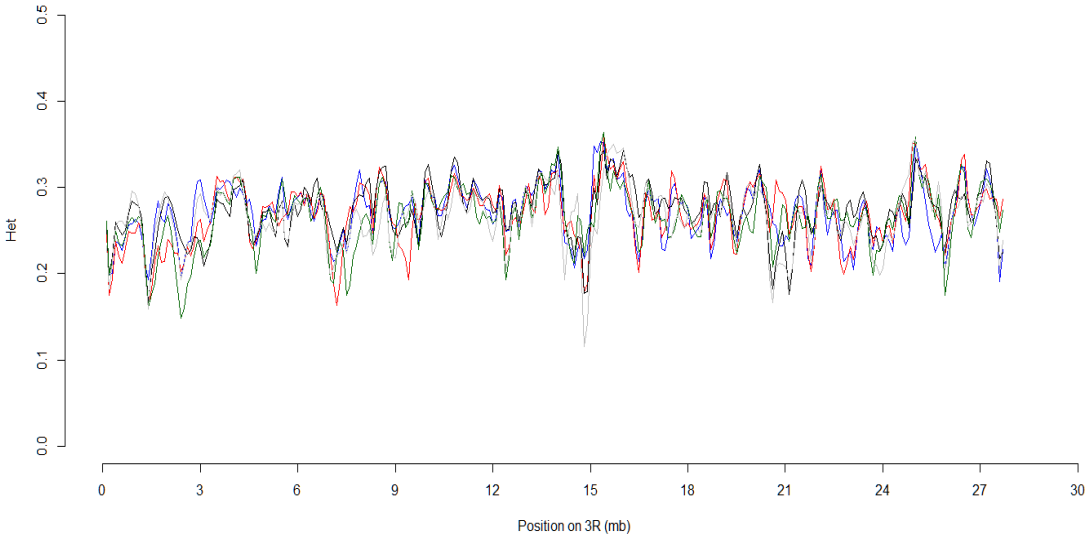

Figure S2.4. Chromosome 3R heterozygosity for CO replicates. Heterozygosity across chromosome arm 3R calculated from SNP data plotted across 100kb windows for the 5 CO replicates. Each color corresponds to an individual replicate.

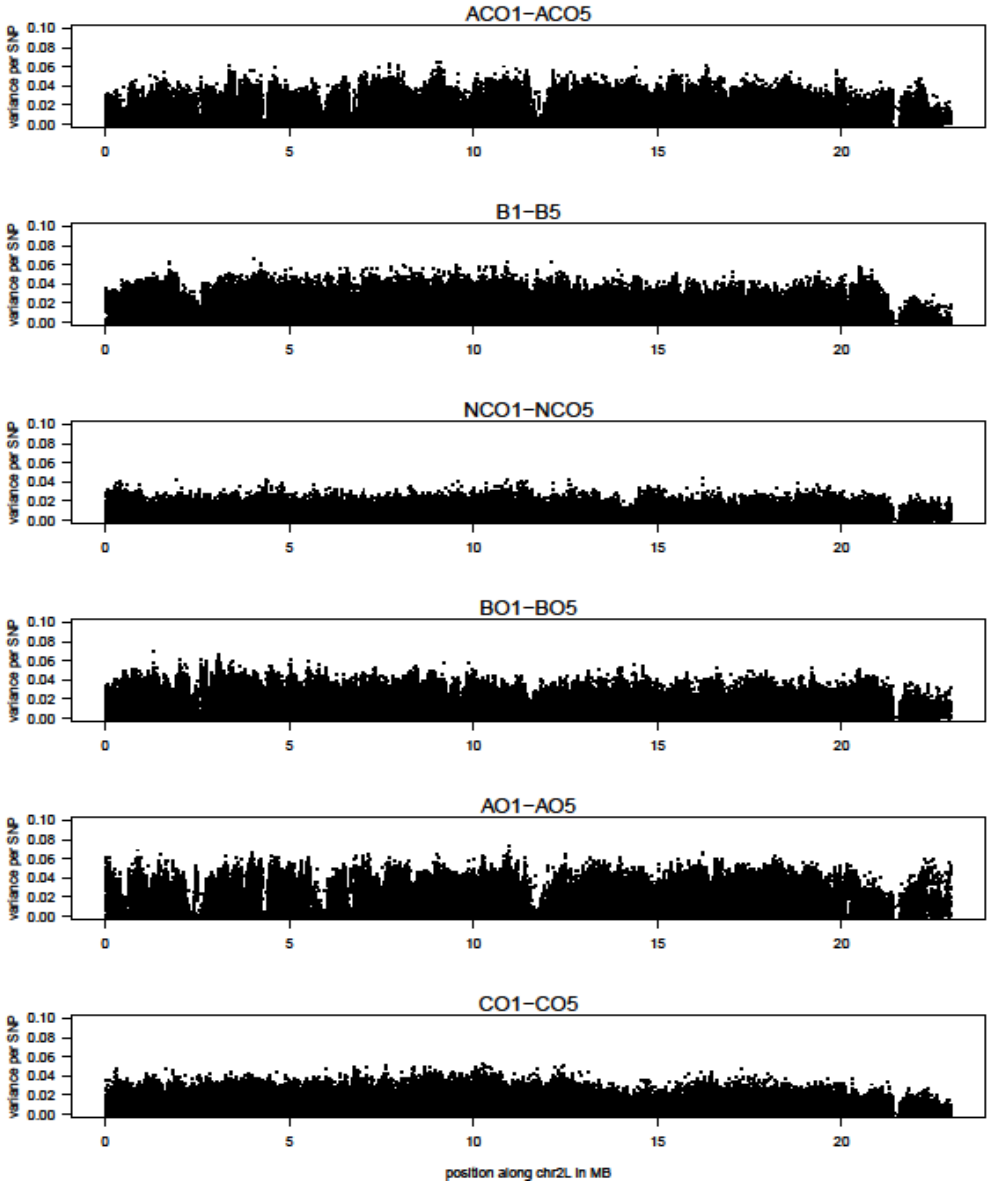

Figure S2.5. Variance in SNP frequencies across a single chromosome arm 2L. We present the variance among five allele frequency estimates (of the major allele) at every SNP position in each selection treatment. The maximum variance possible is 0.3 (e.g. if two of the five replicates were fixed for one allele and three were fixed for the other), and our observed variances are far lower than this. This pattern of low per-SNP variance is preserved across all the major chromosome arms of Drosophila, though we present results from 2L only for conciseness.

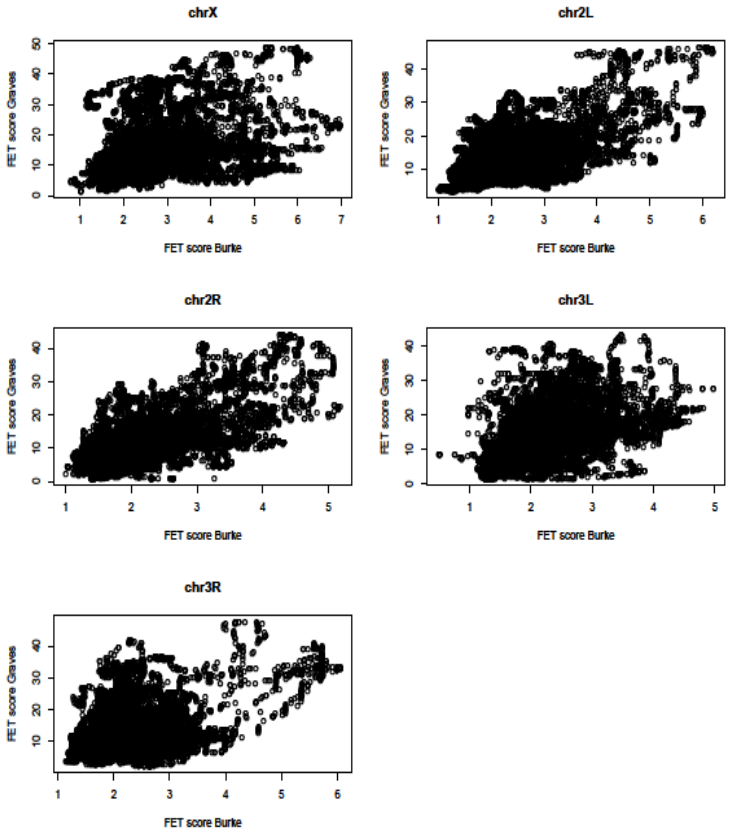

Figure S2.6. Comparison of differentiation results between Burke et al. 2010 and the current study. Burke et al. (2010) pooled genomic data across all five ACO populations and all five CO populations, and conducted Fisher's Exact Tests on the allele counts from these pools. The p-values from these tests were –log10 transformed and evaluated on sliding windows of 100KB every 2KB along each chromosome. The quantile score that only 5% of the transformed p-values exceeded was recorded for each window and here we call this the "FET score" for clarity. We employed the same methods as Burke et al. (2010) by pooling replicates, applying Fisher's Exact Tests, and evaluating the same genomic windows. FET scores from each study are plotted against one another for each major chromosome arm. FET scores are far higher in the current study because coverage was higher (in Burke et al. 2010, the ACO and CO populations had a mean coverage of 20X each while here we observe mean coverage to be 427 for the ACOs and 346 for the COs), increasing power. However, the two studies independently implicate many of the same windows as having elevated FET scores, and thus the same genes underlying differentiation, particularly for chromosomes 2 and 3.

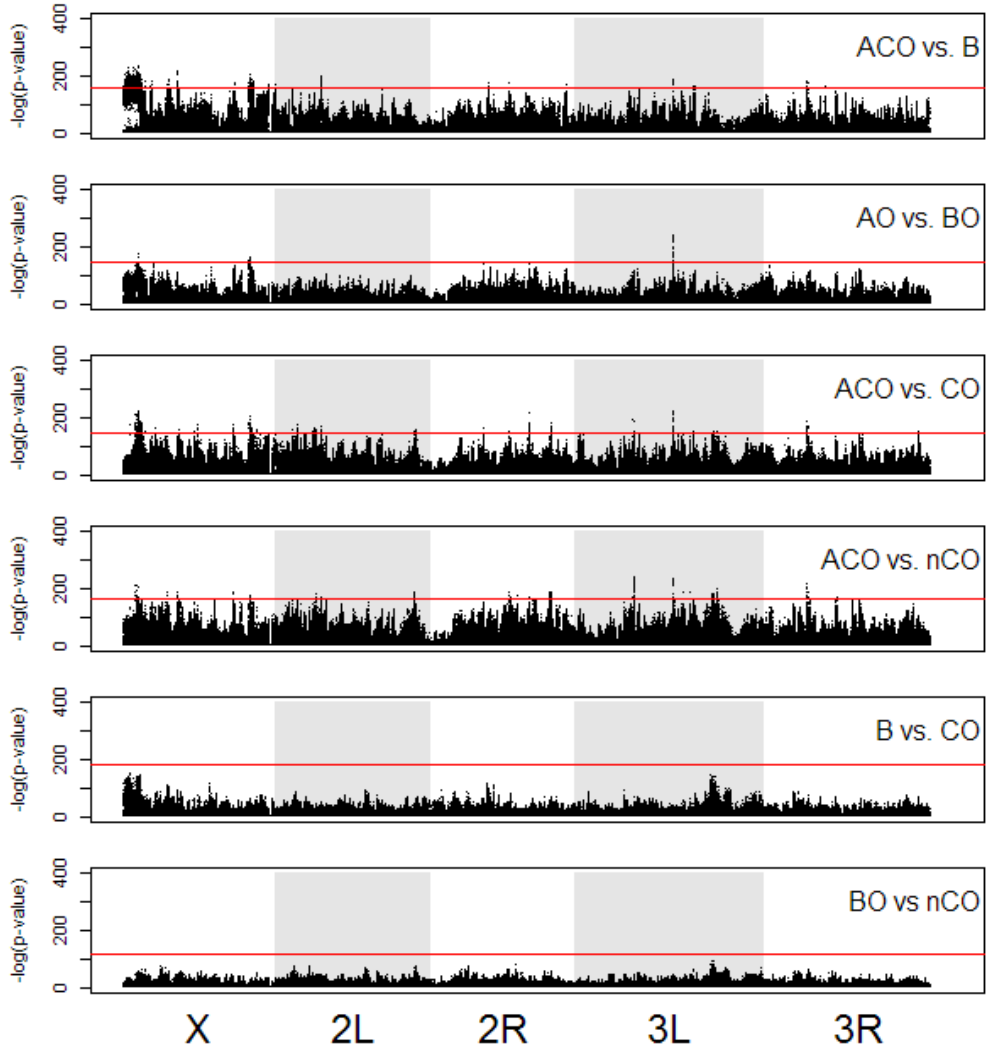

Figure S2.7. Results from 5 X 5 CMH tests between populations subjected to different selection regimes. Results are plotted as –log (p-values) across all major chromosome arms. Significant thresholds for each comparison are derived from permutation tests and are indicated by a red line.

Supplementary Methods: Tables S2.1 and S2.2

We tested for convergence between paired selection treatments (i.e. CO vs. nCO, AO vs. ACO, and B vs. BO) for effects of selection on fecundity and survivorship over 3-4 consecutive ages. The observation consisted of fecundity or survivorship at a particular age (t) but within a small age interval (k= 1,2,…,m). These age intervals were chosen to span the ages, such that all comparison populations still had live flies. Within each interval, fecundity rates or survivorship were modeled by a straight line and allowing selection regime (j= 1 (ACO or CO or B), j=2 (AO or nCO or BO)) to affect the intercept of that line but not the slope. However, slopes were allowed to vary between intervals. Populations (i= 1,…,10) were assumed to contribute random variation to these measures. With this notation, the fecundity at age-t, interval-k, selection regime-j and population-i, is y_{ijkt} and is describe by,

$$y_{ijkt}=\alpha+\beta_k+\delta_j\gamma_j+\omega+\pi_k\delta_k t+\delta_k\delta_j\mu_{jk}+c_i+\varepsilon_{ijkt}$$

where δ_s=0 if s=1 and 1 otherwise, and c_i and ε_{ijkt} are independent standard normal random variables with variance σ_c^2 and σ_ε^2 respectively. The effects of selection on the intercept are assessed by considering the magnitude and variance of both γ_j and $\mu_{jk.}$

To test for divergence, the six selection regimes were reclassified to three different categories: AO and ACO to A; CO and nCO to C; B and BO to B. The same equation was used as listed above to assess the effects of the three selection treatments. For both the convergence and divergence analysis, we used the Bonferroni correction to adjust the significance level for each pair-wise comparison made by dividing the significance level by the number of age interval used in the analysis (0.05 / n, where n is the number of age intervals used).

Table S2.1: Listed p-values testing for parallel evolution between our treatments. Intervals were used to break up the data into smaller portions for better analysis of the entire curve. Bonferroni correction was applied to each pair-wise comparison to determine significance by dividing the p-value by the number of intervals used in each analysis. Bolded and italicized values are significant. Non-significant values are congruent with parallel evolution.

Interval	ACO vs AO		B vs BO		CO vs nCO	
	Fecundity	Survivorship	Fecundity	Survivorship	Fecundity	Survivorship
1	0.29911	0.84882	0.18093	0.78532	0.18303	0.46469
2	0.29469	0.66125	0.96273	0.27102	0.47664	0.18277
3	0.30375	0.53089	0.80526	0.16135	0.65612	0.08773
4	0.53059	0.22566	0.79165	0.04422	0.00724	0.01746
5	0.73621	0.12249				
6	0.82034	0.05594				

Table S2.2: Listed p-values testing for divergent evolution between our treatment types. Intervals were used to break up the data into smaller portions for better analysis of the entire curve. Bonferroni correction was applied to each pair-wise comparison to determine significance by dividing the p-value by the number of intervals used in each analysis. Bolded and italicized values are significant. Significant values are congruent with divergent evolution.

Interval	A type vs B type		A type vs C type		B type vs C type	
	Fecundity	Survivorship	Fecundity	Survivorship	Fecundity	Survivorship
1	0.69848	0.07481	0.93819	0.06117	0.64227	0.92072
2	0.44601	0.01816	0.04757	0.00702	0.20379	0.69047
3	0.85409	2.427x10-5	0.72101	1.399x10-6	0.89306	0.31616
4	0.64908	2.579x10-11	1.066x10-14	2.538x10-13	8.782x10-12	0.02891

Table S2.3. Average SNP coverage across the genome for all populations used in this study.

Population	Replicate	Average Read Coverage
CO	1	47
	2	90
	3	107
	4	73
	5	28
ACO	1	82
	2	99
	3	104
	4	88
	5	52
AO	1	60
	2	108
	3	76
	4	85
	5	94
nCO	1	100
	2	59
	3	62
	4	95
	5	52
B	1	74
	2	56
	3	31
	4	93
	5	65
BO	1	84

	2	48
	3	47
	4	67
	5	56

Table S2.4. Mean genome-wide heterozygosities calculated from SNP data

	ACO	AO	B	BO	CO	nCO
Replicate 1	0.192	0.165	0.206	0.233	0.270	0.277
Replicate 2	0.182	0.160	0.229	0.238	0.267	0.274
Replicate 3	0.176	0.178	0.216	0.241	0.269	0.270
Replicate 4	0.216	0.153	0.216	0.241	0.260	0.275
Replicate 5	0.176	0.157	0.215	0.236	0.260	0.275
Mean	0.189	0.162	0.216	0.238	0.265	0.274

Table S2.5. T-test results comparing mean genome wide heterozygosity calculated from SNP data between different groups of populations.

Comparison	P-value
ACO vs AO	0.02059
ACO vs B	0.01625
ACO vs BO	0.002188
ACO vs CO	0.0002364
ACO vs nCO	0.0002474
AO vs B	1.629 X 10-05
AO vs BO	1.686 X 10-05
AO vs CO	7.286 X 10-07
AO vs nCO	4.282 X 10-06
B vs BO	0.00216
B vs CO	1.057 X 10-05
B vs nCO	2.348 X 10-05
BO vs CO	2.442 X 10-05
BO vs nCO	1.341 X 10-05
CO vs nCO	0.0125

Table S2.6. Mean genome wide F_{ST} across replicates for a given group of replicate populations calculated from SNP data.

Populations	Mean FST
ACO_{1-5}	0.062
AO_{1-5}	0.087
B_{1-5}	0.065
BO_{1-5}	0.058
CO_{1-5}	0.041
nCO_{1-5}	0.028

Table S2.7. 100 kb regions that consistently have values of heterozygosity less than 0.1 across all 5 ACO populations.

Chromosome	Start of Region	ACO1 Het	ACO2 Het	ACO3 Het	ACO4 Het	ACO5 Het
2L	2405390	0.03	0.05	0.12	0.07	0.02
2L	2505390	0.02	0.04	0.03	0.06	0.01
2L	2605390	0.04	0.03	0.08	0.06	0.06
2L	4305390	0.09	0.07	0.03	0.08	0.07
2L	4405390	0.12	0.10	0.08	0.13	0.07
2L	5305390	0.12	0.11	0.05	0.14	0.11
2L	5805390	0.04	0.03	0.04	0.14	0.01
2L	5905390	0.03	0.02	0.02	0.11	0.01
2L	6005390	0.02	0.06	0.07	0.14	0.03
2L	6705390	0.03	0.04	0.05	0.10	0.04
2L	6805390	0.05	0.06	0.06	0.09	0.04
2L	9905390	0.12	0.08	0.13	0.14	0.13
2L	11605390	0.05	0.05	0.01	0.12	0.05
2L	11705390	0.01	0.03	0.01	0.09	0.02
2L	11805390	0.01	0.01	0.01	0.08	0.01
2L	11905390	0.01	0.02	0.01	0.09	0.01
2L	12005390	0.07	0.08	0.06	0.14	0.07
2L	21405390	0.12	0.14	0.10	0.11	0.11
2R	1518404	0.08	0.10	0.14	0.06	0.09
2R	4318404	0.10	0.09	0.07	0.10	0.07
2R	4418404	0.01	0.03	0.04	0.07	0.05

2R	4518404	0.12	0.13	0.12	0.15	0.14
2R	7318404	0.06	0.07	0.11	0.09	0.08
2R	8818404	0.05	0.08	0.13	0.13	0.11
2R	8918404	0.08	0.13	0.09	0.15	0.13
2R	11318404	0.02	0.06	0.03	0.10	0.03
2R	11418404	0.03	0.08	0.03	0.09	0.03
2R	14318404	0.08	0.05	0.04	0.13	0.03
2R	14418404	0.07	0.04	0.04	0.14	0.01
2R	18718404	0.14	0.12	0.15	0.14	0.12
2R	18818404	0.10	0.07	0.10	0.10	0.08
2R	19118404	0.02	0.05	0.06	0.04	0.04
2R	19218404	0.07	0.07	0.04	0.05	0.04
2R	19318404	0.12	0.09	0.10	0.09	0.13
2R	20118404	0.02	0.07	0.02	0.10	0.06
2R	20218404	0.02	0.03	0.02	0.05	0.04
2R	20318404	0.01	0.02	0.02	0.04	0.02
2R	20418404	0.04	0.04	0.04	0.04	0.02
2R	20518404	0.08	0.06	0.06	0.06	0.05
2R	20618404	0.08	0.06	0.06	0.06	0.05
3L	3734721	0.12	0.11	0.08	0.10	0.12
3L	7934721	0.07	0.04	0.05	0.10	0.09
3L	8034721	0.08	0.05	0.08	0.10	0.08
3L	8134721	0.06	0.06	0.11	0.12	0.05
3L	8234721	0.05	0.08	0.13	0.14	0.05
3L	9734721	0.10	0.06	0.03	0.07	0.03
3L	9834721	0.10	0.06	0.01	0.08	0.02

3L	9934721	0.14	0.10	0.04	0.11	0.05
3L	16134721	0.07	0.08	0.06	0.09	0.01
3L	16234721	0.06	0.07	0.04	0.06	0.06
3L	16334721	0.13	0.15	0.08	0.09	0.13
3L	17034721	0.03	0.03	0.04	0.05	0.04
3L	17134721	0.04	0.06	0.06	0.09	0.06
3L	17234721	0.09	0.11	0.12	0.14	0.09
3L	17434721	0.10	0.05	0.14	0.14	0.04
3L	17534721	0.08	0.05	0.07	0.11	0.06
3L	17634721	0.03	0.15	0.03	0.11	0.09
3L	18134721	0.08	0.08	0.09	0.12	0.09
3L	18234721	0.05	0.04	0.07	0.05	0.04
3L	18334721	0.07	0.09	0.13	0.07	0.06
3L	18434721	0.07	0.13	0.12	0.08	0.08
3L	19134721	0.15	0.14	0.12	0.13	0.14
3L	19234721	0.10	0.08	0.10	0.07	0.07
3R	14003326	0.06	0.06	0.05	0.13	0.07
3R	24803326	0.06	0.08	0.14	0.07	0.10
3R	24903326	0.04	0.03	0.06	0.09	0.03
3R	26203326	0.11	0.10	0.07	0.12	0.09
3R	26303326	0.02	0.05	0.02	0.06	0.02
3R	26403326	0.02	0.02	0.04	0.07	0.02
3R	26503326	0.10	0.09	0.12	0.12	0.10
X	519797	0.01	0.03	0.01	0.13	0.01
X	1219797	0.02	0.04	0.02	0.15	0.02
X	1819797	0.02	0.03	0.02	0.08	0.02

X	1919797	0.01	0.03	0.01	0.14	0.01
X	2219797	0.02	0.04	0.02	0.14	0.02
X	2319797	0.02	0.04	0.02	0.14	0.02
X	2419797	0.02	0.04	0.02	0.14	0.02
X	2519797	0.01	0.03	0.01	0.14	0.01
X	2619797	0.01	0.03	0.01	0.13	0.01
X	2719797	0.04	0.05	0.03	0.14	0.06
X	3719797	0.02	0.02	0.01	0.10	0.01
X	3819797	0.02	0.02	0.04	0.11	0.02
X	8119797	0.10	0.04	0.04	0.12	0.05
X	9919797	0.10	0.15	0.08	0.14	0.05
X	10819797	0.09	0.09	0.12	0.15	0.08
X	14119797	0.12	0.04	0.08	0.07	0.10
X	14219797	0.08	0.05	0.03	0.05	0.10
X	14319797	0.12	0.10	0.03	0.11	0.10
X	15219797	0.03	0.10	0.03	0.14	0.13
X	15319797	0.04	0.06	0.02	0.09	0.07
X	18319797	0.08	0.03	0.13	0.07	0.08
X	18419797	0.09	0.04	0.05	0.10	0.08
X	18519797	0.05	0.04	0.03	0.11	0.04
X	18619797	0.02	0.03	0.02	0.14	0.02
X	18819797	0.02	0.02	0.02	0.14	0.06
X	18919797	0.05	0.03	0.02	0.12	0.06
X	19019797	0.06	0.05	0.02	0.13	0.05
X	19119797	0.03	0.05	0.02	0.12	0.03
X	19219797	0.06	0.08	0.06	0.14	0.04

X	19719797	0.11	0.04	0.15	0.13	0.03
X	19919797	0.11	0.05	0.14	0.14	0.07
X	20019797	0.08	0.05	0.09	0.12	0.04
X	20119797	0.08	0.07	0.07	0.12	0.03
X	20619797	0.07	0.08	0.03	0.11	0.12
X	20719797	0.05	0.03	0.05	0.12	0.03

Table S2.8. 100 kb regions that consistently have values of heterozygosity less than 0.1 across all 5 AO populations.

Chromosome	Start of Region	AO1 Het	AO2 Het	AO3 Het	AO4 Het	AO5 Het
2L	1205390	0.10	0.02	0.10	0.09	0.13
2L	1305390	0.13	0.05	0.09	0.09	0.12
2L	1505390	0.10	0.01	0.05	0.05	0.13
2L	1605390	0.06	0.03	0.03	0.08	0.11
2L	1705390	0.06	0.06	0.04	0.10	0.10
2L	1805390	0.03	0.04	0.03	0.06	0.11
2L	2205390	0.03	0.05	0.10	0.05	0.02
2L	2305390	0.02	0.02	0.05	0.02	0.02
2L	2405390	0.04	0.02	0.01	0.07	0.02
2L	2505390	0.04	0.01	0.01	0.06	0.01
2L	2605390	0.06	0.02	0.03	0.05	0.04
2L	3105390	0.08	0.14	0.13	0.11	0.12
2L	3705390	0.13	0.06	0.10	0.12	0.14
2L	4305390	0.06	0.01	0.02	0.07	0.02
2L	4405390	0.06	0.10	0.10	0.10	0.10
2L	5605390	0.04	0.06	0.10	0.15	0.12
2L	5705390	0.01	0.04	0.10	0.10	0.05
2L	5805390	0.01	0.02	0.08	0.06	0.02
2L	5905390	0.01	0.02	0.04	0.03	0.01
2L	6005390	0.02	0.03	0.02	0.05	0.06
2L	6505390	0.10	0.09	0.12	0.07	0.10
2L	6605390	0.01	0.05	0.05	0.01	0.04
2L	6705390	0.01	0.07	0.02	0.01	0.09

2L	7405390	0.10	0.05	0.13	0.07	0.12
2L	7605390	0.04	0.08	0.08	0.13	0.14
2L	9905390	0.11	0.09	0.09	0.13	0.10
2L	11605390	0.02	0.04	0.07	0.01	0.01
2L	11705390	0.01	0.02	0.05	0.01	0.01
2L	11805390	0.01	0.07	0.01	0.01	0.01
2L	11905390	0.01	0.10	0.01	0.01	0.01
2L	12005390	0.07	0.14	0.07	0.05	0.07
2L	12105390	0.15	0.14	0.13	0.12	0.14
2L	16505390	0.15	0.06	0.09	0.08	0.11
2L	21305390	0.09	0.10	0.09	0.07	0.12
2L	21405390	0.13	0.13	0.10	0.07	0.11
2R	4318404	0.08	0.09	0.07	0.09	0.08
2R	4418404	0.01	0.03	0.04	0.01	0.06
2R	8818404	0.13	0.13	0.12	0.10	0.07
2R	8918404	0.12	0.15	0.09	0.12	0.08
2R	9018404	0.08	0.11	0.10	0.12	0.08
2R	9118404	0.08	0.10	0.13	0.13	0.09
2R	11218404	0.07	0.05	0.10	0.08	0.13
2R	11318404	0.03	0.03	0.03	0.04	0.05
2R	11418404	0.07	0.01	0.01	0.02	0.03
2R	11518404	0.15	0.02	0.03	0.02	0.04
2R	14318404	0.03	0.04	0.05	0.07	0.04
2R	14418404	0.01	0.05	0.04	0.05	0.02
2R	14518404	0.09	0.14	0.12	0.09	0.07
2R	15218404	0.12	0.09	0.09	0.07	0.12

2R	15318404	0.08	0.05	0.06	0.04	0.04
2R	15418404	0.06	0.03	0.06	0.05	0.03
2R	15518404	0.03	0.08	0.05	0.07	0.08
2R	17618404	0.08	0.12	0.11	0.11	0.07
2R	17718404	0.12	0.07	0.08	0.08	0.05
2R	17818404	0.13	0.11	0.13	0.08	0.09
2R	18718404	0.06	0.09	0.10	0.07	0.08
2R	18818404	0.05	0.06	0.07	0.10	0.07
2R	18918404	0.10	0.06	0.13	0.12	0.08
2R	19018404	0.09	0.08	0.14	0.08	0.10
2R	19118404	0.01	0.05	0.03	0.02	0.06
2R	19218404	0.01	0.01	0.01	0.06	0.01
2R	19318404	0.01	0.04	0.06	0.10	0.07
2R	19418404	0.10	0.13	0.14	0.14	0.15
2R	19918404	0.08	0.09	0.13	0.12	0.11
2R	20018404	0.05	0.05	0.05	0.04	0.08
2R	20118404	0.01	0.01	0.07	0.02	0.02
2R	20218404	0.02	0.03	0.10	0.02	0.03
2R	20318404	0.08	0.03	0.08	0.01	0.02
2R	20418404	0.15	0.02	0.05	0.04	0.03
2R	20518404	0.15	0.02	0.05	0.07	0.06
2R	20618404	0.12	0.02	0.06	0.07	0.06
3L	3734721	0.07	0.09	0.11	0.11	0.09
3L	8034721	0.09	0.05	0.12	0.13	0.04
3L	8234721	0.09	0.02	0.10	0.10	0.01
3L	9734721	0.06	0.03	0.05	0.10	0.04

3L	10334721	0.13	0.07	0.10	0.12	0.10
3L	11034721	0.09	0.14	0.14	0.12	0.08
3L	14634721	0.04	0.14	0.06	0.08	0.08
3L	16134721	0.01	0.09	0.13	0.01	0.08
3L	16234721	0.01	0.08	0.02	0.06	0.06
3L	17034721	0.03	0.02	0.05	0.01	0.02
3L	17134721	0.02	0.02	0.02	0.01	0.02
3L	17234721	0.04	0.03	0.09	0.01	0.02
3L	17334721	0.05	0.02	0.14	0.03	0.01
3L	17434721	0.05	0.02	0.14	0.04	0.02
3L	17534721	0.06	0.06	0.08	0.03	0.02
3L	17634721	0.09	0.09	0.02	0.01	0.02
3L	17734721	0.14	0.11	0.01	0.02	0.03
3L	18134721	0.10	0.09	0.12	0.01	0.03
3L	18234721	0.02	0.03	0.12	0.05	0.03
3L	18334721	0.02	0.04	0.12	0.10	0.08
3L	18434721	0.02	0.03	0.10	0.13	0.11
3L	18534721	0.05	0.09	0.12	0.14	0.14
3L	21234721	0.10	0.11	0.13	0.12	0.13
3L	21434721	0.12	0.14	0.14	0.14	0.14
3L	21534721	0.11	0.15	0.14	0.14	0.14
3L	21634721	0.13	0.15	0.15	0.12	0.13
3L	21734721	0.12	0.07	0.13	0.03	0.06
3L	24034721	0.13	0.13	0.15	0.14	0.15
3R	4303326	0.11	0.13	0.11	0.10	0.12
3R	4403326	0.07	0.15	0.13	0.06	0.15

3R	8703326	0.06	0.07	0.09	0.03	0.12
3R	10003326	0.09	0.09	0.13	0.06	0.05
3R	10103326	0.13	0.06	0.13	0.08	0.02
3R	14003326	0.06	0.04	0.09	0.06	0.10
3R	14103326	0.08	0.05	0.14	0.09	0.13
3R	14203326	0.12	0.09	0.11	0.14	0.12
3R	14503326	0.09	0.12	0.15	0.10	0.11
3R	17403326	0.06	0.11	0.12	0.07	0.10
3R	17903326	0.14	0.12	0.10	0.13	0.13
3R	19903326	0.14	0.08	0.09	0.04	0.10
3R	20003326	0.14	0.09	0.12	0.05	0.13
3R	24803326	0.07	0.05	0.14	0.01	0.06
3R	24903326	0.02	0.01	0.04	0.02	0.01
3R	26203326	0.09	0.08	0.07	0.10	0.06
3R	26303326	0.01	0.01	0.01	0.03	0.01
3R	26403326	0.01	0.01	0.01	0.01	0.01
3R	26503326	0.08	0.09	0.08	0.07	0.07
3R	27603326	0.07	0.08	0.03	0.06	0.08
X	1819797	0.02	0.02	0.02	0.02	0.02
X	1919797	0.03	0.01	0.01	0.02	0.01
X	2019797	0.02	0.01	0.01	0.01	0.00
X	2119797	0.10	0.02	0.02	0.08	0.01
X	2519797	0.14	0.01	0.01	0.12	0.01
X	2619797	0.11	0.01	0.01	0.10	0.01
X	2719797	0.05	0.05	0.05	0.03	0.02
X	2819797	0.09	0.11	0.14	0.04	0.04

X	3719797	0.01	0.03	0.12	0.01	0.02
X	3819797	0.01	0.10	0.08	0.01	0.04
X	6819797	0.05	0.05	0.12	0.02	0.14
X	7919797	0.08	0.14	0.14	0.03	0.03
X	8019797	0.07	0.09	0.07	0.02	0.03
X	8119797	0.05	0.06	0.02	0.01	0.07
X	8219797	0.07	0.10	0.06	0.03	0.14
X	8619797	0.11	0.02	0.11	0.12	0.01
X	8719797	0.05	0.06	0.15	0.04	0.04
X	9019797	0.07	0.10	0.09	0.01	0.04
X	10219797	0.07	0.08	0.13	0.07	0.12
X	10519797	0.03	0.11	0.11	0.06	0.10
X	13419797	0.11	0.04	0.14	0.13	0.12
X	13519797	0.05	0.04	0.11	0.09	0.12
X	14119797	0.06	0.08	0.02	0.01	0.05
X	14219797	0.04	0.06	0.01	0.02	0.03
X	14319797	0.12	0.09	0.01	0.10	0.05
X	15119797	0.03	0.06	0.02	0.12	0.12
X	15219797	0.02	0.05	0.02	0.05	0.06
X	15319797	0.02	0.08	0.02	0.03	0.03
X	16219797	0.15	0.03	0.10	0.11	0.04
X	18419797	0.12	0.07	0.10	0.09	0.02
X	18519797	0.07	0.04	0.05	0.05	0.02
X	18619797	0.02	0.03	0.03	0.05	0.01
X	18719797	0.07	0.05	0.07	0.05	0.02
X	18819797	0.09	0.04	0.06	0.04	0.01

X	18919797	0.09	0.01	0.01	0.13	0.02
X	19019797	0.06	0.03	0.01	0.14	0.03
X	19119797	0.01	0.04	0.02	0.03	0.03
X	19219797	0.04	0.06	0.06	0.07	0.02
X	20119797	0.11	0.12	0.02	0.07	0.04
X	20219797	0.14	0.12	0.02	0.10	0.02
X	20619797	0.03	0.14	0.06	0.14	0.03
X	21719797	0.05	0.02	0.14	0.04	0.02

Table S2.9. 100 kb regions that consistently have values of heterozygosity less than 0.1 across all 5 B populations.

Chromosome	Start of Region	B1 Het	B2 Het	B3 Het	B4 Het	B5 Het
2L	2505390	0.04	0.06	0.02	0.09	0.07
2L	2605390	0.09	0.09	0.07	0.09	0.12
2L	17505390	0.03	0.14	0.15	0.05	0.15
2L	21405390	0.08	0.10	0.11	0.09	0.10
2L	21505390	0.06	0.07	0.07	0.11	0.08
2L	21605390	0.05	0.06	0.09	0.13	0.06
2L	21705390	0.07	0.08	0.10	0.11	0.08
2L	22105390	0.09	0.14	0.14	0.09	0.12
2L	22205390	0.08	0.14	0.13	0.09	0.11
2L	22605390	0.07	0.12	0.13	0.06	0.10
2L	22705390	0.06	0.11	0.12	0.05	0.08
2L	22805390	0.06	0.12	0.10	0.06	0.08
2R	318404	0.11	0.15	0.13	0.10	0.13
2R	8218404	0.12	0.13	0.09	0.13	0.06
2R	8718404	0.13	0.09	0.09	0.15	0.09
2R	19118404	0.06	0.05	0.07	0.02	0.02
2R	19218404	0.10	0.07	0.07	0.02	0.02
2R	20218404	0.09	0.15	0.08	0.09	0.13
2R	20318404	0.04	0.08	0.07	0.04	0.08
3L	3434721	0.12	0.12	0.10	0.09	0.07
3L	3534721	0.08	0.08	0.08	0.03	0.06
3L	3634721	0.07	0.08	0.09	0.03	0.06
3L	3734721	0.08	0.08	0.13	0.06	0.07
3L	9334721	0.14	0.14	0.06	0.04	0.04

3L	9434721	0.04	0.12	0.03	0.05	0.04
X	119797	0.01	0.07	0.01	0.01	0.01
X	219797	0.01	0.07	0.01	0.01	0.01
X	319797	0.01	0.08	0.01	0.01	0.01
X	419797	0.01	0.11	0.01	0.01	0.01
X	519797	0.01	0.12	0.01	0.01	0.01
X	619797	0.01	0.15	0.01	0.01	0.01
X	1219797	0.02	0.14	0.03	0.03	0.02
X	1319797	0.01	0.12	0.01	0.01	0.01
X	1419797	0.01	0.12	0.01	0.02	0.02
X	1519797	0.01	0.10	0.01	0.01	0.02
X	1619797	0.01	0.10	0.01	0.01	0.01
X	1719797	0.01	0.08	0.01	0.01	0.01
X	1819797	0.01	0.07	0.01	0.01	0.01
X	1919797	0.01	0.14	0.01	0.01	0.01
X	2019797	0.01	0.14	0.01	0.01	0.01
X	2119797	0.02	0.09	0.02	0.02	0.02
X	2219797	0.01	0.09	0.04	0.02	0.05
X	2319797	0.02	0.09	0.06	0.01	0.07
X	2419797	0.14	0.15	0.09	0.11	0.13
X	6619797	0.14	0.12	0.11	0.11	0.10
X	6719797	0.08	0.13	0.05	0.06	0.06
X	11819797	0.04	0.10	0.12	0.12	0.09
X	14119797	0.04	0.09	0.02	0.03	0.11
X	14219797	0.02	0.13	0.02	0.02	0.12
X	14319797	0.04	0.15	0.09	0.11	0.10

X	14619797	0.04	0.14	0.12	0.09	0.14
X	14719797	0.12	0.11	0.06	0.03	0.04
X	16419797	0.03	0.07	0.07	0.03	0.07
X	16519797	0.02	0.07	0.02	0.03	0.04
X	16619797	0.03	0.05	0.02	0.02	0.06
X	16719797	0.11	0.10	0.02	0.05	0.05
X	18219797	0.05	0.03	0.09	0.03	0.06
X	21619797	0.05	0.05	0.05	0.04	0.04
X	21719797	0.00	0.01	0.00	0.00	0.01

Table S2.10. 100 kb regions that consistently have values of heterozygosity less than 0.1 across all 5 BO populations.

Chromosome	Start of Region	BO1 Het	BO2 Het	BO3 Het	BO4 Het	BO5 Het
2L	2505390	0.07	0.11	0.13	0.12	0.09
2L	21505390	0.10	0.10	0.11	0.09	0.10
2L	21605390	0.13	0.14	0.12	0.12	0.08
3L	3434721	0.12	0.09	0.06	0.07	0.12
3L	3534721	0.14	0.05	0.10	0.10	0.12
3L	3634721	0.11	0.05	0.11	0.11	0.11
3L	3734721	0.08	0.14	0.11	0.15	0.08
3L	9834721	0.09	0.14	0.04	0.11	0.11
3R	24803326	0.09	0.11	0.07	0.13	0.12
3R	27603326	0.10	0.13	0.13	0.13	0.10
X	1819797	0.15	0.11	0.13	0.13	0.13

Table S2.11. 100 kb regions that consistently have values of heterozygosity less than 0.1 across all 5 CO populations.

Chromosome	Start of Region	CO1 Het	CO2 Het	CO3 Het	CO4 Het	CO5 Het
3L	20134721	0.13	0.07	0.06	0.07	0.04
X	16219797	0.15	0.09	0.09	0.11	0.09
X	16319797	0.15	0.08	0.09	0.12	0.08
X	21619797	0.09	0.07	0.05	0.10	0.05
X	21719797	0.08	0.05	0.03	0.07	0.00

Table S2.12. 100 kb regions that consistently have values of heterozygosity less than 0.1 across all 5 nCO populations.

Chromosome	Start of Region	nCO1 Het	nCO2 Het	nCO3 Het	nCO4 Het	nCO5 Het
3L	20134721	0.11	0.09	0.07	0.07	0.10
3L	20234721	0.12	0.10	0.07	0.06	0.08
3L	20334721	0.13	0.10	0.11	0.05	0.07
3L	20434721	0.13	0.10	0.13	0.08	0.09
3L	20534721	0.14	0.11	0.13	0.10	0.13
X	16519797	0.12	0.14	0.10	0.15	0.15
X	21619797	0.05	0.06	0.06	0.07	0.10
X	21719797	0.04	0.07	0.02	0.05	0.11

CHAPTER 3

Effects of Evolutionary History on Genome-wide Convergence in *Drosophila* Populations

Abstract

Previous E&R studies using the Rose lab's experimental system suggests that the response to selection is highly repeatable at both the phenotypic and genomic levels, and that evolutionary history has little impact. However, other studies suggest that even when the response to selection is repeatable phenotypically, evolutionary history can have significant impacts at the genomic level. Here we test two hypotheses that may explain this discrepancy. *Hypothesis* 1: Evolutionary history matters in populations where past generations were subjected to very intense selective pressures. *Hypothesis* 2: Previous intense selection does not produce historical effects, but other evolutionary mechanisms may. We test these hypotheses using *D. melanogaster* populations that were subjected to 260 generations of intense selection for desiccation resistance and have since been under relaxed selection for the past 230 generations. Our genomic analysis found no signs of genetic fixation, and only limited evidence of genetic differentiation between the previously desiccation selected population and their controls. Given other studies using these populations found that relaxed selection has erased the extreme phenotypic differentiation previously found between these populations and their controls, we conclude that evolutionary history does not generically play a major role in shaping trajectories of adaptation during Drosophila experimental evolution.

INTRODUCTION

The combination of experimental evolution and next generation sequencing has become established as a powerful means for studying the genetics of adaptation and testing major tenets of population genetic theory (Long et al. 2015; Schlötterer et al. 2015). Studies featuring populations of fruit flies, *Drosophila melanogaster*, suggest that adaptation in sexually reproducing populations is fueled by selection on standing genetic variation, and is largely characterized by a lack of genetic fixation (Teotónio et al. 2009; Burke et al. 2010; Turner et al. 2011; Orozco-terWengel et al. 2012; Tobler et al. 2014; Huang et al. 2014; Franssen et al. 2015). The apparent lack of fixation is even seen in long-term experiments nearing a thousand generations of selection (Phillips et al. 2016). Moreover, work with outcrossing populations of *Saccharomyces cerevisiae* has shown that adaptation is still primarily driven by standing genetic variation even at much larger effective population sizes than what is currently seen in experiments featuring *D. melanogaster* (Burke et al. 2014).

In accordance with these findings, evolution in outbred populations is rapid and highly repeatable when newly derived *D. melanogaster* populations are subjected to the same selection regimes as long-standing populations (Burke et al. 2016; Graves et al. 2017). It takes only dozens of generations for newly derived populations to converge on long-standing populations at both the genomic and phenotypic levels, even when long-standing populations have previously undergone hundreds of generations of selection (Burke et al. 2016; Graves et al. 2017). These findings suggest that phenotypes and patterns of genetic variation are primarily shaped by most recent selection regime, and

that evolutionary history, prior to the recent selection regime, has little discernible impact. However, this runs contrary to evidence from experimental evolution work using *Drosophila suboscura* derived from wild populations at contrasting European latitudes (Simões et al. 2017). Their findings indicate that evolution is predictable at the phenotypic level, but differences in where source populations originate can have significant effects on outcomes at the genetic level, suggesting that evolutionary history does play a role when it comes to repeatability at the genomic level. The idea that the degree to which populations return to ancestral phenotypic values and allele frequencies is at least in part contingent on evolutionary history is also supported by reverse experimental evolution studies (Teotónio and Rose 2000; Teotónio et al. 2009). However, it should be noted that the authors of these reverse experimental evolution studies were unable to rule out the possibility that complete reversion would have occurred in all populations if given more time.

A possible resolution to why evolutionary history appears to play a role in some experiments but not others follows from Graves et al. (2017). In addition to the aforementioned results, the authors found evidence that more intense selection regimes lead to significantly greater losses of genetic variation compared to milder selection regimes. While they did not observe widespread fixation within any of the populations studied, fixation seemed possible provided a sufficiently intense selection regime. Therefore, the finding that phenotypes and patterns of genetic variation are almost exclusively shaped by most recent selection regime in Burke et al. (2016) and Graves et al. (2017) could be due to the fact that none of the populations studied were exposed to sufficiently intense selection regimes in their evolutionary histories. Presumably, strong selective pressures could potentially result in widespread fixation of alleles favored by

such selection. In these cases, given that adaptation in sexual experimental evolution is primarily fueled by standing genetic variation, the widespread fixation could have significant impact on how experimental populations respond to new selective pressures, and their ability to revert to ancestral states when moved back to ancestral conditions.

Here we test two hypotheses. Hypothesis 1: evolutionary history matters in populations where past generations were subjected to very intense selective pressures. Hypothesis 2: previous intense selection does not produce historical effects, in the absence of other factors such as inbreeding or chromosomal rearrangement.

We test these hypotheses using a group of *D. melanogaster* populations that were historically subjected to intense selection for desiccation resistance, TDO $_{1\text{-}5}$, and their controls, TSO $_{1\text{-}5}$. These populations were known as D and C respectively during active selection and were first described in Rose et al. (1992). The D populations were intensely selected for desiccation resistance for about 260 generations, and afterward were renamed as TDO, and maintained on a 21 day (T for "Three-week") relaxed culture selection for the past ~230 generations. The C populations were moderately selected for starvation resistance for about 260 generations in parallel with the D populations, serving as controls for the D populations, and were later renamed as TSO, and maintained under the same culture selection regime as the TDO populations. The extreme functional differentiation (i.e. carbohydrate content, water loss rates, and water content) previously seen between these two groups was achieved using environments so inimical to survival that only a small percentage (10-20%) of each generation survives selection (Rose et al. 1992; Gibbes et al. 1997; Djawdan et al. 1998; Archer et al. 2003). We have called this intense selection

paradigm “culling selection” in the past, and it represents one of the most extreme protocols used in Drosophila experimental evolution (Rose et al. 1990).

Findings from Phillips et al. (in review) show that there is currently no significant differentiation in relation to starvation and desiccation resistance between the TSO and TDO, but there are significant differences in mean longevity (Figure 3.1). Thus, relaxed selection has erased the extreme differences in starvation and desiccation resistance observed during selection (Djawdan et al. 1998; Archer et al. 2003; Phelan et al. 2003), but there is still evidence of previously observed longevity differences (Rose et al. 1992). (Figure 3.1). Here we pair these findings with an examination of patterns of genomic differentiation between the two group to test our hypotheses. With Hypothesis 1, large impacts of evolutionary history are in fact due to exposure to intense selection. If this hypothesis is correct, we would expect to find significant genomic differentiation between the TDO and TSO populations even after ~230 generations of relaxed selection in the former. If Hypothesis 2 is correct, we should not find such differentiation.

Materials and Methods

Populations

This experiment used large, outbred lab populations of *Drosophila melanogaster* derived from a population sampled by P.T. Ives from South Amherst, Massachusetts (Ives, 1970). The experimental stocks used in this study were derived from a set of 5 populations that had been selected for late reproduction ($O_{1\text{-}5}$). The $O_{1\text{-}5}$ populations were derived from the Ives stock in February 1980 (Rose et al. 1984). In 1988, two sets of populations were

derived from the O_{1-5} populations. One set (D_{1-5}) were selected for desiccation resistance while the other set (C_{1-5}) were maintained to control for desiccation resistance selection. The C_{1-5} populations were handled like the D_{1-5} populations, except flies were given nonnutritive agar instead of desiccant (Rose et al. 1992). In 2005, these populations were relaxed from selection and kept on a 21-day culture regime to the present day. Under this new regime, the D populations have been renamed to TDO, and the C populations to TSO. In total, the TDO populations underwent ~260 generations of selection for desiccation resistance, and ~230 generations of relaxed selection.

Populations were reared on a banana-molasses diet for stock maintenance and for experimental assays. The banana-molasses media is composed of the following ingredients per 1L distilled H_20: 13g Apex® Drosophila agar type II, 120g peeled, ripe banana, 40mL light Karo® corn syrup, 40mL dark Karo® corn syrup, 50mL Eden®organic barley malt syrup, 32g Red Star® active dry yeast, 2g Sigma-Aldrich® Methyl 4-hydroxybenzoate (anti-fungal), and 42 mL EtOH. Stocks are maintained on a 24-hour light cycle and kept at room temperature (24°C ± 1°C).

DNA extraction and sequencing

Genomic DNA was extracted from samples of 200 female flies collected from each of the 10 individual populations (TSO_{1-5} and TDO_{1-5}) using the Qiagen©/Gentra Puregene© kit, following the manufacturer's protocol for bulk DNA purification. The 30 gDNA pools were prepared as standard 200-300 bp fragment libraries for Illumina sequencing, and constructed such that each 5 replicate populations of a treatment (e.g., TSO_{1-5}) were given unique barcodes, normalized, and pooled together. Libraries were run across PE100 lanes

of an Illumina HiSEQ 2000 at the UCI Genomics Highthroughput Sequencing Facility. Resulting data were 100 bp paired-end reads. Each population was sequenced twice; data from both runs were combined for some analyses as described below. Combining reads from two independent sequencing runs likely alleviate the effects of possible bias introduced from running all replicates for each population in the same lane.

Genomic Analysis

Mapping of reads

Reads were mapped to the *D. melanogaster reference genome* (version 6.14) using bwa mem with default settings (BWA version 0.7.8) (Lir and Durbin 2009). The resulting SAM files were filtered for reads mapped in proper pairs with a minimum mapping quality of 20, and converted to the BAM format using the view and sort commands in SAMtools (Li et al. 2009). The rmdup command in SAMtools was then used to remove potential PCR duplicates. As each population was sequenced twice, there were two bam files corresponding to each population at this stage. BAMtools was used to combine pairs corresponding to the same populations (Barnett et al. 2011). Average coverage was above 70X for all populations except TSO_3, which was 67X (Table S3.1). Next, SAMtools was used to combine the 10 bam files into a single mpileup file. Using the PoPoolation2 software package (Kofler et al. 2011), these files were converted to "synchronized" files, which is a format that allele counts for all bases in the reference genome and for all populations being analyzed. We then used RepeatMasker 4.0.3 (http://www.repeatmasker.org) to create a gff file detailing low complexity regions in the *D. melanogaster* reference genome. The regions were then masked in our sync file once again using PoPoolation2.

SNP Variation

A SNP table was created using the sync file mentioned above. We only considered sites where coverage was between 30X and 200X, and for a site to be considered polymorphic we required a minimum minor allele frequency of 2% across all 10 populations. All sites failing to meet these criteria were discarded. To assess broad patterns of SNP variation in TSO and TDO populations, heterozygosity was calculated and plotted over 150kb non-overlapping windows directly from the major and minor counts in our SNP table. A t-test was also performed to compare mean heterozygosity between the two groups of populations. To assess how closely replicate populations resembled one another, F_{ST} estimates were also obtained using the formula: $F_{ST}= (H_T-H_S)/H_T$ where H_T is heterozygosity based on total population allele frequencies, and H_S is the average subpopulation heterozygosity in each of the replicate populations (Hedrick 2009). F_{ST} estimates were made at every polymorphic site in the data set for a given set of replicate populations.

SNP Differentiation

We used two different methods to assess SNP differentiation in the TSO and TDO populations. First, we used the CMH test as implemented in the PoPoolation2 software package to compare SNP frequencies between the TSO and TDO populations. As the findings of Wiberg et al. (2017) indicate that coverage variation can impact statistical results, we subsampled to a uniform coverage of 50X across the genome for each population using scripts provided in the PoPoolation2 software package. During this process, all positions with coverage less than 50X or greater than 200X were discarded. The subsampling procedure involved calculating the exact fraction of the allele frequencies at

each site, and linearly scaling them to our target coverage of 50X. In addition to these coverage requirements, we only considered sites polymorphic if they had a minor allele frequency of 2% across all ten populations. In total, the resulting subsampled sync file contained ~1.2 million SNPs spread across the major chromosome arms. CMH tests were then performed at every polymorphic site between the TSO and TDO populations. To correct for multiple comparisons, we used the permutation approach featured in Graves et al. (2017). Briefly, populations were randomly assigned to one of two groups and the CMH test was then performed at each polymorphic position in the shuffled data set to generate null distributions of p-values. This was done a 1000 times, and each time the smallest p-value generated was recorded. The quantile function in R was then used to define thresholds that define the genome-wide false-positive rate, per site, at 5%.

In addition to the CMH test, we also used the quasibinomial GLM approach recommended by Wiberg et al. (2017). Their findings suggest this approach has lower false positive and higher true positive rates than the CMH test. However, it should be noted that the permutation derived significance threshold used in our CMH tests are more stringent than anything featured in their analysis. The test was implemented using scripts provided by Wiberg et al. (2017). A .sync file was once again the primary input file, and we used the same SNP calling criteria outlined above (minimum coverage of 50X per population, maximum of 200X per population, and a minimum minor allele frequency of 2% across all 10 populations). Coverage was once again scaled to 50X to minimize the effect of coverage variation on our results. As counts of zero can lead to problems when implementing this approach, a count of 1 was added to each allele whenever a zero was encountered. In terms of establishing significance thresholds, another reported benefit of quasibinomal GLMs is

that they produce the expected uniform distribution of p-values under the null hypothesis which allows for standard method of correcting for multiple comparisons (Wiberg et al. 2017). As a result, to correct for multiple comparisons we used two common approaches, the Bonferroni correction and the q-value method (Storey and Tibshirani 2003; Storey et al. 2015). We chose to use the Bonferroni correction and the q-value methods as Wiberg et al. (2017) found them to be the most and least conservative approach, respectively.

Results

Heterozygosity and F_{ST}

We do not see any large regions where heterozygosity has been completely expunged, and this result is robust to reductions in window size (Figure 3.2, Figures S3.1 and S3.2). However, there are some notable depressions consistent across replicates that may be indicative of soft sweeps. Mean heterozygosity in the TSO populations ranges from 0.24 to 0.26, and 0.26 to 0.27 in the TDO populations (Table S3.2). Based on a t-test comparing the two sets of means, heterozygosity is significantly higher in the TDO populations (p-value = 0.001). Mean F_{ST} in the TSO populations is 0.04 and 0.07 in the TDO populations, which indicates there is a high degree of similarity between replicates.

SNP differentiation

We find little evidence of SNP (single nucleotide polymorphism) differentiation between the TDO and TSO populations (Figure 3.3). Based on our Cochran-Mantel-Haenszel (CMH) test results, we find a total of 17 sites with p-values that exceed our permutation derived significance threshold (Figure 3.3A). These 17 sites correspond to

three regions, two on chromosome 3L and one on the X chromosome. However, we find no signs of significant SNP differentiation using the quasibinomial GLM method. This is true using both the Bonferroni correction, and the less conservative q-value approach to correct for multiple comparisons (Figure 3.3B-C).

Within the significantly differentiated regions detected using the CMH test, we find a total of seven genes (Table 3.1). Six of the seven genes are located on chromosome arm 3L, while the remaining gene is located on chromosome X. For genes *CR42860*, *CR45802*, and *CR34047*, there is little to no information about their molecular and biological functions. Gene CG42355 is associated with sperm chromatin condensation, but not much else is presently known. Genes *sallimus* (*sls)* and *zormin* have been well documented to be associated with the development of the striated muscle sarcomeres (Bullard et al. 2005; Burkart et al. 2007; Orfanos et al. 2015). *Sls* expression is necessary for myoblast fusion, and the inevitable development of myoblasts into muscle fibres. Sallimus, protein derived from the *sls* gene, aides in aligning thin filaments side-by-side and in anti-parallel direction, which nucleates Z-disc formation in developing myofibrils (Bullard et al. 2005. Sallimus also binds to thin filaments (i.e. actin), aiding in balancing the two halves of the sarcomere. Protein zormin can be also be found near the Z-disc and M-line of the muscles (Orfanos et al. 2015). These filaments connect the Z-disc with the ends of the thick filaments. The size of these proteins, and the extensibility of their binding, affects the elasticity and stiffness of muscles (Burkart et al. 2007). These properties dictate muscle contraction, stiffness, and performance. The final gene, *CG32649*, located on the X chromosome is associated with ubiquinone biosynthesis and mitochondrial electron transport. There are two human orthologs, *COQ8A* and *COQ8B*, linked to *CG32649*. Mutations at these two ADCK genes can

lead to primary coenzyme Q-10 deficiency and nephrotic syndrome, respectively (Mollet et al. 2008; Ashraf et al. 2013).

DISCUSSION

Our genomic results combined with the phenotypic results presented in Phillips et al. (in review) indicate that ~230 generations of relaxed selection were enough for the previously desiccation selected TDO populations to largely converge on the TSO controls at both the phenotypic and genomic levels. The TDO populations do not show any signs of significantly enhanced survival in desiccating environments compared to the TSO populations, despite extreme differences in desiccation resistance prior to the relaxation of selection (Rose et al. 1992; Gibbs et al. 1997; Djawdan et al. 1998; Archer et al. 2007). There is also no longer any evidence of increased starvation resistance in the TDO populations, which was a trait previously found to be correlated with enhanced desiccation resistance (Djawdan et al. 1998).

As shown in Phillips et al. (in review), there are still clear signs of phenotypic differentiation with regards to longevity between the TDO and TSO populations. This arises notwithstanding the absence of extensive genomic differentiation detected between these two sets of populations. One possible explanation for this comes from the somewhat limited replication featured in this study. If we consider the genomic analysis of differentiation between A and C populations of Graves et al. (2017), we find that there is a general reduction in the ability of genomic analysis to detect differentiation when only ten populations total are compared as two groups of five. In Graves et al. (2017) thousands of differentiated sites were detected when comparing all ten A-type populations to all ten C-

type populations, compared to hundreds of sites when groups of five were compared to one another. This finding is also supported by theoretical studies examining the power of evolve and re-sequence studies to detect causal variants (Balwdin-Brown et al. 2014; Kofler and Schlötterer 2014). As we are limited to five replicates per treatment in this study, it is possible that we simply do not have the statistical power to detect the genomic differentiation underlying this residual phenotypic differentiation.

Although our selection protocol for desiccation resistance is relatively extreme, compared to other selection regimes we have used we do not find any clear evidence of it having a lasting impact on levels of genetic variation (Rose et al. 2004). Our genomic analysis did not yield any evidence of widespread fixation in the TDO populations. As such, the findings of our genomic analyses suggest that even when a moderately outbred experimental population's evolutionary history involves prolonged periods of intense selection, it does not have lasting or irreversible effects on patterns of genetic variation. This runs contrary to the findings of Simões et al. (2017), and even past studies using other populations related to the TDO and TSO populations (Teotónio and Rose 2000; Teotónio et al. 2009). However, it should be pointed out that none of those studies involved evolutionary histories of the duration used here. Our present findings are more in line with those of Graves et al. (2017), where most recent selection regime was the primary force shaping patterns of genetic variation.

Our findings also suggest that ~230 generations of relaxed selection were enough to erase any signs of widespread genetic differentiation between the TDO and TSO populations. CMH tests comparing SNP frequencies between the two groups of populations

did yield some significantly differentiated sites. However, this was limited to 17 sites as compared to the dozens to thousands of differentiated sites typically detected in *Drosophila* experimental evolution studies (Burke et al. 2010; Turner et al. 2011; Orozco-terWengel et al. 2012; Tobler et al. 2014; Huang et al. 2014; Franssen et al. 2015). There were a total of 7 genes associated with these sites, but none of these candidate genes had clear connections to desiccation resistance (See Results for details). Additionally, the quasibinomial GLM approach to detecting significantly differentiated SNP's advocated by Wiberg et al. (2017) did not detect any significant SNP differentiation between the TDO and TSO populations. Taken together, these findings indicate that ~230 generations of relaxed selection were enough to eliminate most signs of meaningful SNP differentiation between the TDO populations and their controls, further suggesting that evolutionary history does not play a significant role in *Drosophila* experimental evolution over sufficiently long periods.

Our findings about the role of evolutionary history in shaping patterns of genetic variation are not entirely conclusive however. For instance, while we have phenotypic data for the TDO populations prior to the relaxation of selection, we do not have any genomic data from this period. As such, we cannot directly show that relaxing selection resulted in significant shifts in patterns of genetic variation. We also cannot directly compare current levels of SNP differentiation between the TDO and TSO populations to what they were during the height of TDO selection for desiccation resistance. However, assuming past experimental evolution studies featuring genome-wide comparisons between experimentally evolved *Drosophila* populations are broadly applicable, these results nevertheless suggest patterns of genetic variation and differentiation in *Drosophila* experimental evolution are not dramatically impacted by evolutionary history.

Cumulatively, our findings suggest that extreme selection does not have major long-lasting impacts on genomic and phenotypic differentiation in *Drosophila* experimental evolution. We are able to detect some signs of genetic differentiation and residual differences in mean longevity when comparing the TDO and TSO populations, but nothing on the order of what is usually found between selected and control populations in *Drosophila* experimental evolution Burke et al. 2010; Turner et al. 2011; Orozco-terWengel et al. 2012; Tobler et al. 2014; Huang et al. 2014; Franssen et al. 2015). And there is no reason to believe these differences would dramatically impact how these populations respond to future selection. As such, we conclude that evolutionary history does not play a major role in trajectories of adaptation during *Drosophila* experimental evolution, if the duration of these experiments is sufficiently long and populations are not inbred at any point.

References

Archer MA, Phelan JP, Beckman, KA, Rose MR. Breakdown in correlations during laboratory evolution. II. Selection on stress resistance in *Drosophila* populations. Evolution. 2003;57: 536-543.

Archer, M.A., Bradley, T.J., Mueller, L.D. & Rose, M.R. Using experimental evolution to study the functional mechanisms of desiccation resistance in *Drosophila melanogaster.* Physiol Biochem Zool. 2007;80: 386-398.

Ashraf S, Gee HY, Woerner S, Xie LX, Vega-Warner V, Lovric S. et al. ADCK4 mutations promote steroid-resistant nephrotic syndrome through CoQ10 biosynthesis disruption. J Clin Invest. 2013;123:5179--5189.

Baldwin-Brown JG, Long AD, Thornton KR. The power to detect quantitative trait loci using resequenced, experimentally evolved populations of diploid, sexual organisms. Mol Biol Evol. 2014;31:1040–1055.

Barnett DW, Garrison EK, Quinlan AR, Stromberg MP, Marth GT BamTools: a C ++ API and toolkit for analyzing and managing BAM files. Bioinformatics. 2011;27:1691–1692.

Bullard B, Burkart C, Labeit S, Leonard K. The function of elastic proteins in the oscillatory contraction of insect flight muscle. J. Muscle Res. Cell Motility. 2005;26:479-485.

Burkart C, Qiu F, Brendel S, Benes V, Haag P, Labeit, S, et al.Modular proteins from the Drosophila sallimus (sls) gene and their expression in muscles with different extensibility. J Mol Biol. 2007:367:953--969.

Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. Nature. 2010;467:587–590.

Burke MK, Liti G, Long AD. Long standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. Mol Biol Evol. 2014;32:3228–3239.

Burke MK, Barter TT, Cabral LG, Kezos JN, Phillips MA, Rutledge GA, et al. Rapid convergence and divergence of life-history in experimentally evolved *Drosophila Melanogaster*. Evolution 2016;70:2085–2098.

Djawdan M, Chippindale AK, Rose MR, Bradley TJ. Metabolic reserves and evolved stress resistance in *Drosophila melanogaster.* Physiol Zool. 1998;71:584-594.

Franssen SU, Volte N, Tobler R, Schlötterer C. Patterns of linkage disequilibrium and long range hitchhiking in evolving experimental *Drosophila* melanogaster populations. Mol Biol Evol. 2015;32:495–509.

Gibbs, A.G., Chippindale, A.K. & Rose, M.R. Physiological mechanisms of evolved desiccation resistance in *Drosophila melanogaster*. J Exp Biol. 1997;200:1821-1832.

Graves JL, Hertweck KL, Phillips MA, Han MV, Cabral LG, Barter TT, et al. Genomics of parallel experimental evolution in Drosophila. Mol Biol Evol. 2017;34:831-842.

Hedrick PW. Genetics of populations. Massachusetts: Jones & Bartlett Learning Press; 2009.

Huang Y, Wright SI, Agrawal AF. Genome-wide patterns of genetic variation within and among alternative selective regimes. PLoS Genet. 2014;10, e1004527.

Kofler R, Pandey RV, Schlötterer C. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics. 2011;27:3435–3436.

Kofler R, Schlötterer C. A guide for the design of evolve and resequencing studies. Mol Biol Evol. 2014;31:474–483.

Long AD, Liti G, Luptak A, Tenaillon O. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. Nat. Rev. Genet. 2015;16:567–582.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–2079.

Mollet J, Delahodde A, Serre V, Chretien D, Schlemmer D, Lombes A, et al. CABC1 gene mutations cause ubiquinone deficiency with cerebellar ataxia and seizures. Am J Hum Genet. 2008;82:623–630.

Orfanos Z, Leonard K, Elliott C, Katzemich A, Bullard, B, Sparrow J. Sallimus and the dynamics of sarcomere assembly in Drosophila flight muscles. J Mol Biol. 2015;427:2151-2158.

Orozco-ter Wengel P, Kapun M, Nolte V, Kofler R, Flatt T, Schlötterer C. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected traits. Mol Ecol. 2012;21:4931–4941.

Phelan JP, Archer MA, Beckman KA, Chippindale AK, Nusbaum TJ, Rose MR. Breakdown in correlations during laboratory evolution. I. Comparative analyses of *Drosophila* populations. Evolution. 2003;57: 527-535.

Phillips MA, Long AD, Greenspan ZS, Greer LF, Burke MK, Bryant V, et al. Genome-wide analysis of long-term evolutionary domestication in *Drosophila melanogaster.* Sci. Rep. 2016. doi:10.1038/srep39281

Phillips MA, Rutledge GA, Kezos JN, Talbott A, Matty S, Arian H, et al. Effects of evolutionary history on genome wide and phenotypic convergence in Drosophila populations. In review BMC Genomics.

Rose MR. Laboratory evolution of postponed senescence in *Drosophila melanogaster*. Evolution. 1984;38:1004–1010.

Rose MR, Graves JL, Hutchinson EW. 1990. The use of selection to probe patterns of peliotrophy in fitness characters. In: *Insect Life Cycles*, Gilbert F, editor. New York: Spring-Verlag; 1990. p. 29-42.

Rose MR, Vu LN, Park, SU, Graves JL. Selection for stress resistance increases longevity in *Drosophila melanogaster*. Exp Gerontol. 1992;27:241-250.

Rose MR, Passananti HB, Matos M. Methuselah flies: a case study in the evolution of aging. Singapore:World Scientific Publishing; 2004.

Schlötterer C, Kofler R, Versace E, Tobler R, Franseen SU. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. Heredity 2015;114:331–440.

Simões P, Fragata I, Seabra SG, Faria GS, Santos MA, Rose MR, et al. 2017. Predictable phenotypic, but not karyotypic, evolution of populations with contrasting initial history. Sci Rep. 2017;71:913.

Storey JD, Tibshirani, R. Statistical significance for genomewide studies. Proc Natl Acad Sci. 2003:100:9440–9445

Storey JD, Bass A, Dabney A, Robinson D. q-value: Q-value estimation for false discovery rate control. R package version 2.2.2. 2015. Retrieved from http://github.com/jdstorey/qvalue(Accessed 07 April 2017).

Teotónio H, Rose MR. Variation in the reversibility of evolution. Nature 2000;408:463–466.

Teotónio H, Chelo IM, Bradiá M, Rose MR, Long AD. Experimental evolution reveals natural selection on standing genetic variation. Nat Genet. 2009;41:251–257.

Tobler R, Franssen SU, Kofler R, Orozco-ter Wengel P, Nolte V, Hermisson J, Schlötterer C. Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. Mol Biol Evol. 2014;31:364–375.

Turner TL, Steward AD, Fields AT, Rice WR, Tarone AM. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. PLoS Genet.2011; 7:e10001336.

Wiberg RAW, Gaggiotti OE, Morrissey MB, Ritchie MG, Johnson L. Identifying consistent allele frequency differences in studies of stratified populations. Methods Ecol Evol. 2017;8:1899-1909.

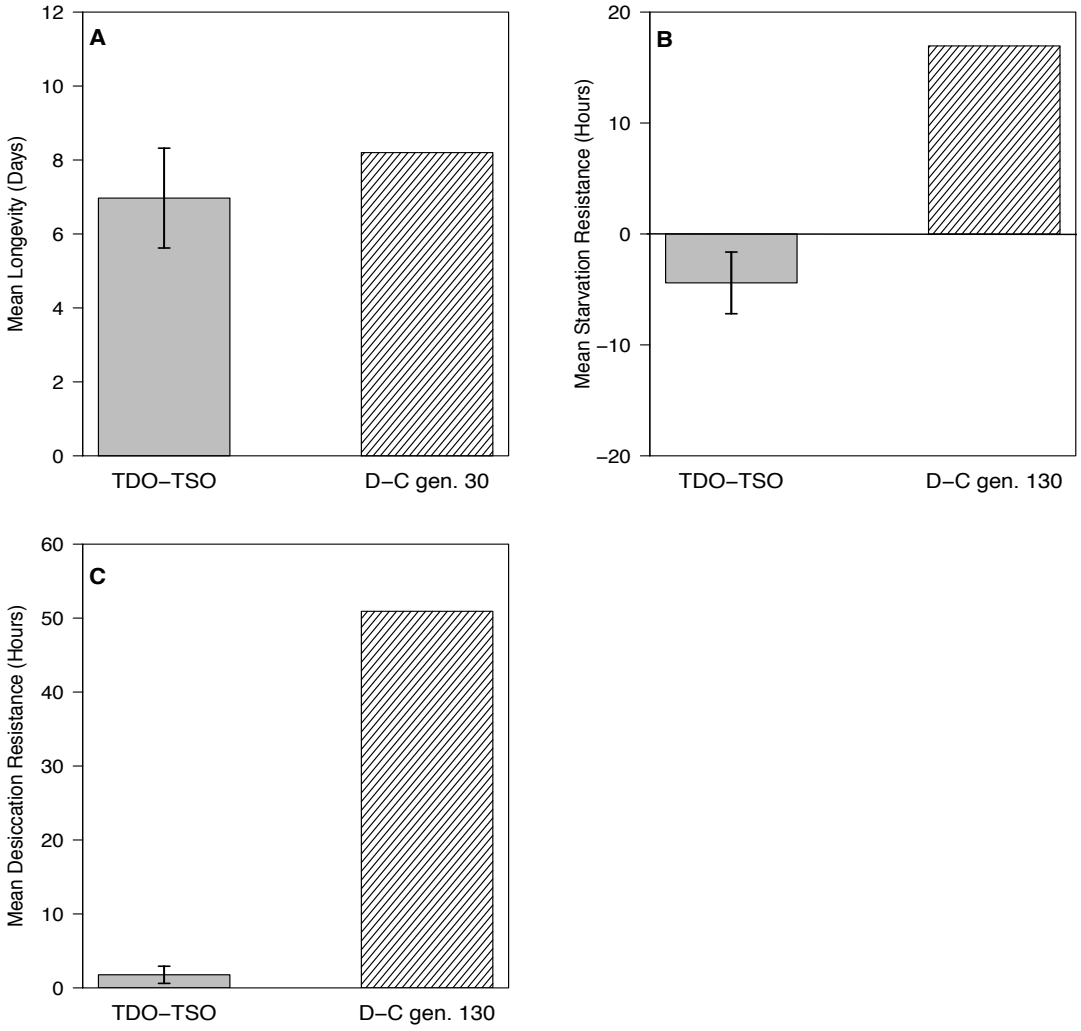

Figure 3.1. Historical and current starvation resistance, desiccation resistance and mean longevity data from females in the desiccation selected and control lines. (**a**) Difference in average longevity between the selected and control populations. Differences are significant in both comparisons shown. Difference in mean longevity was highest early in desiccation selection (Rose et al. 1992), and then decreases to near zero toward the end of selection (Archer et al. 2003). (**b**) Difference in mean survival when flies were subjected to starvation conditions. No significant difference between TSO and TDO populations, but starvation resistance was significantly different during selection with the greatest differences being found at generation 130 (Djawdan et al. 1998; Phelan et al. 2003). (**c**) Difference in mean survival when flies were subjected to desiccation conditions. No significant difference between TSO and TDO populations, but desiccation resistance was significantly different during selection with the greatest differences being found at generation 130 (Archer et al. 2003). Error bars are mean ± 1 SEM.

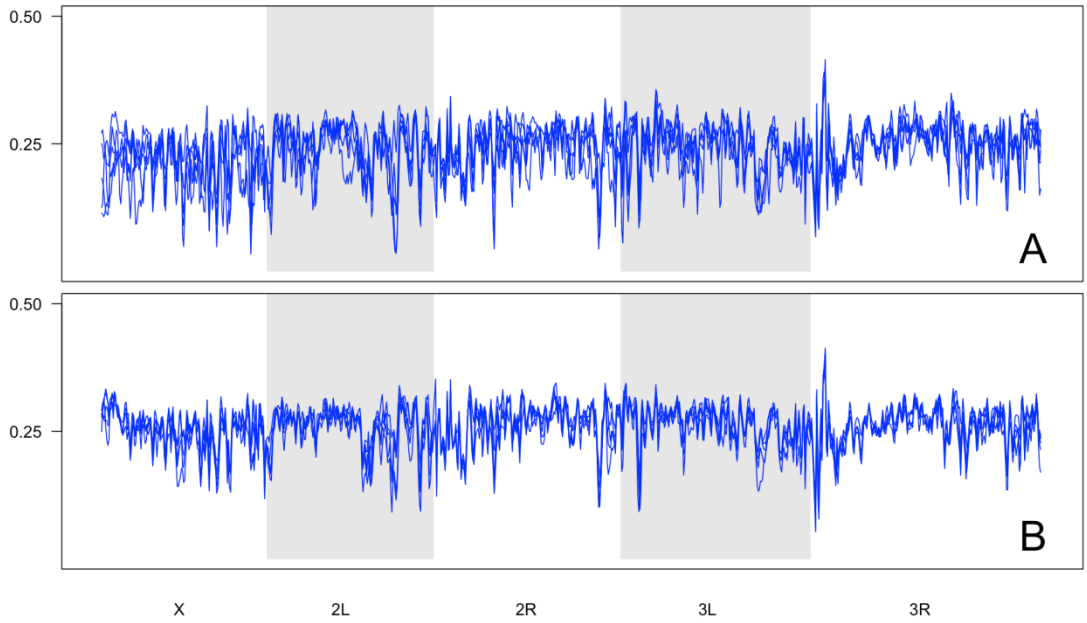

Figure 3.2. Genome-wide heterozygosity 150kb windows. Heterozygosity in the TSO **(a)** and TDO **(b)** populations plotted over 150-kb windows across all major chromosome arms. All replicates are shown for each population.

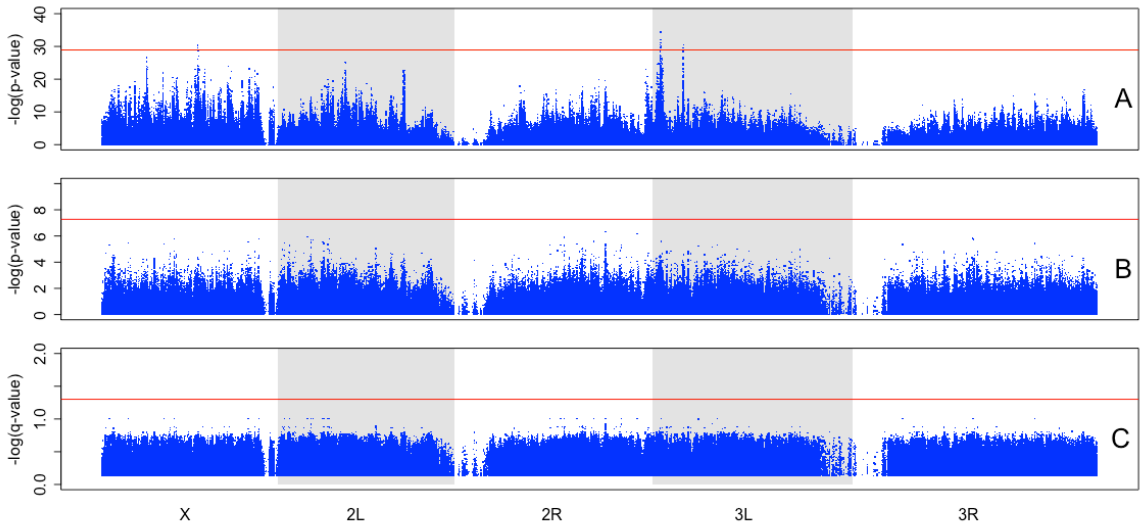

Figure 3.3. CMH and quasibinomial GLM comparison of SNP frequencies between TDO and TSO populations. Results from statistical test comparing SNP frequencies in the TDO and TSO populations. (**a**) Results from CMH tests plotted along all major chromosome arms as –log(p-values). Our permutation derived significance threshold is shown in red. (**b**) Results from quasibinomial GLM approach plotted as -log(p-values), and our Bonferroni corrected significance threshold is shown in red. (**c**) Results from quasibinomial GLM converted to q-values, and plotted –log(q-values). A 0.05 false discovery rate threshold was used, as shown in red.

Table 3.1. Genes located in regions found to be significantly differentiated based on our CMH test comparing SNP frequencies in the TDO and TSO populations.

Gene	Location	Association	Molecular Function	Biological Process
CG42355	3L: 2037371-2038224.	Unknown	Unknown	Sperm chromatin condensation.
sls (Sallimus)	3L: 2039681-2115611	Protein necessary for myoblast fusion; determinant of resting elasticity of striated muscle sarcomeres (myofibril stiffness); regulates mitochondrial respiration in sarcomere	Structural constituent of muscle; Actin binding; Protein binding.	Chromosome organization; skeletal muscle organ development; regulation of immune system process; mesoderm development; chromosome condensation; locomotion; somatic muscle development; myotube differentiation; visceral muscle development; striated muscle tissue development; regulation of multicellular organismal process.
CR42860	3L: 2088166-2089626	Unknown	Unknown	Unknown
Zormin	3L: 2117466-2151700	Found in the Z-disc and the M-line of muscles. Affects elasticity and stiffness of sarcomeres.	Protein binding; actin binding	
CR45802	3L: 2118498-2119567	Unknown	Unknown	Unknown
CR34047	3L: 5098376-5099795	Unknown	Unknown	Unknown
CG32649	X: 12898768-12901114	Ubiquinone biosynthesis; CoQ8A and CoQ8B human orthologs	Protein kinase activity.	Mitochondrial electron transport, ubiquinol to cytochrome c.

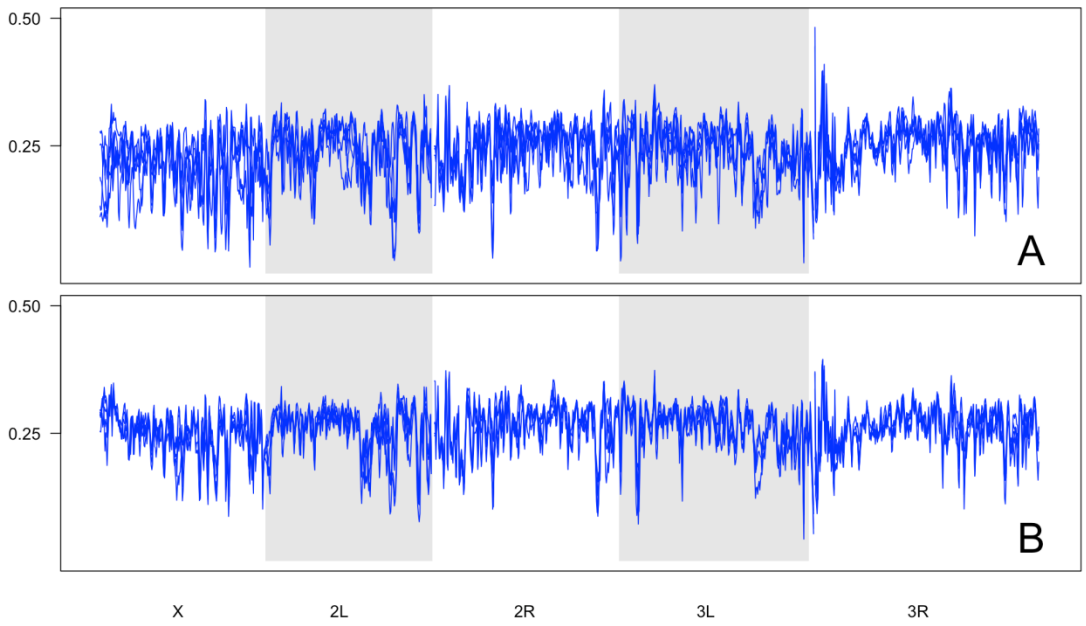

Figure S3.1. Genome-wide heterozygosity 100kb windows. Heterozygosity in the TSO *(A)* and TDO (*B)* populations plotted over 100-kb windows across all major chromosome arms. All replicates are shown for each population.

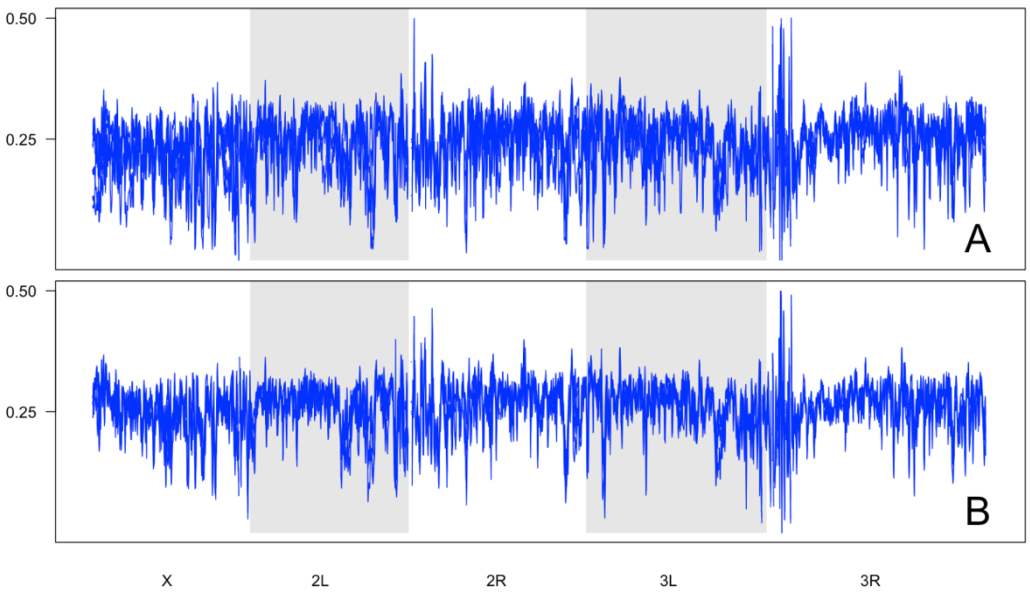

Figure S3.2. Genome-wide heterozygosity 50kb windows. Heterozygosity in the TSO *(A)* and TDO (*B)* populations plotted over 50-kb windows across all major chromosome arms. All replicates are shown for each population.

Table S3.1. Average read coverage across the genome for all populations used in this study.

Population	Replicate	Average Read Coverage
TSO	1	87
	2	83
	3	78
	4	80
	5	70
TDO	1	96
	2	81
	3	67
	4	87
	5	78

Table S3.2. Mean genome-wide heterozygosities calculated from SNP data

	TSO	TDO
Replicate 1	0.24	0.27
Replicate 2	0.25	0.27
Replicate 3	0.24	0.26
Replicate 4	0.25	0.26
Replicate 5	0.26	0.27
Mean	0.25	0.27

CHAPTER 4

The Relationship Between Effective Population Size and Power to Detect Causal Variants in *Drosophila* Experimental Evolution

ABSTRACT

As the number of E&R studies aimed at understanding the genetic architecture of complex traits in sexually reproducing populations grows, it is increasingly important to understand how experimental designs impact strength of inference. At present, empirical and theoretical studies have shown that a high degree of replication is essential, and experiments using sexually reproducing populations need not be longer than ~40 generations in duration. Here we empirically asses the importance of another potentially important experimental parameter, effective population size (N_e). We hypothesize that E&R studies featuring sexual populations with low N_e will have limited power to detect causal variants due to the increased strength of genetic drift. We test this hypothesis by analyzing DNA sequence data from two experiments featuring selection for starvation resistance in *Drosophila melanogaster*. One experiment was conducted using starvation selected and control populations with N_e of ~1000, while the other was conducted with starvation selected and control populations where the census size massively compressed. We find that at low N_e, genetic drift indeed plays a larger role in shaping patterns of genetic variation than what was seen in any of our previous work. Genome-wide comparisons of SNP frequencies in the high N_e experiment resulted in the identification of hundreds of candidate sites, while comparisons in the low N_e experiment failed to detect any candidate sites. It is also worth noting that both of these experiments feature over twice the level of replication seen in most published E&R studies. As such, we find support for our hypothesis and conclude that N_e is indeed a major determinant of statistical power in E&R studies.

INTRODUCTION

As the number of Evolve and Resequence (E&R) studies aimed at understanding the genetic architecture of complex traits in sexually reproducing populations grows, it is increasingly important to understand how experimental designs impact strength of inference. At present, replication has been shown to affect the ability to detect significantly differentiated regions of the genome (Burke et al. 2014; Graves et al. 2017). In accordance with theoretical studies examining the power of E&R studies to detect casual variants (Baldwin-Brown et al. 2014; Kofler and Schlötterer 2014), Graves et al. (2017) and Burke et al. (2014) found that reducing the number of replicates used in their analyses reduced the number of differentiated sites detected, and the ability to dissect genomic regions responding to selection. Graves et al. (2017) also suggest that selection experiments in moderately outbred populations need not be more than ~40 generations in duration, given the rate at which newly derived experimental populations converged on long-standing counterparts when subjected to the same selection regimes in their study. Here we seek to empirically assess the importance of another key experimental parameter that may impact the ability of E&R studies to detect causal variants, effective population size (N_e).

E&R featuring sexually reproducing populations studies have consistently found that adaptation is primarily fueled by selection on standing genetic variation (Teotonio et al 2009; Burke et al. 2010; Turner et al. 2011; Orzoco-terWengel et al. 2012; Tobler et al. 2014, Franssen et al 2014, Huang et al 2014; Burke et al. 2014; Graves et al. 2017). As reductions in N_e are associated with increased levels of genetic drift and reduced levels of genetic variation, small populations are theoretically expected to show a reduced response

to selection (Allendorf and Luikart 2007; Frankham 2005; Robertson 1960). Selection experiments have shown that this is indeed that case for phenotypes: reduced N_e has been shown to limit the phenotypic response to directional selection (Madalena and Robertson 1974; Marden et al. 1997; Weber and Diggins 1990). Given that adaptation in E&R studies with sexually reproducing populations is primarily fueled by standing genetic variation, it stands to reason that this reduced phenotypic response to selection at low N_e is due to genetic drift overwhelming selection.

E&R studies featuring sexual populations typically aim to detect causal genetic variants by the detection of either consistent patterns of genomic differentiation between groups of replicate populations subjected to different selection regimes, or consistent patterns of allele frequency change across generations in a single group of populations subjected to the same selection regime (Schlötterer et al. 2015; Long et al. 2015). If the reduced adaptive response to selection at low N_e is indeed due to genetic drift overwhelming the selection, detecting causal variants in E&R studies featuring populations with low N_e could pose a major statistical challenge. Instead of the highly parallel genomic response to selection across replicates that has come to characterize E&R studies with sexual populations (e.g. Burke et al., 2014; Graves et al., 2017), drift causing random reductions in the frequency or outright loses of beneficial alleles in some replicates should disrupt statistical consistency across replicates. High levels of genetic drift at low N_e could also produce large allele frequency changes at neutral sites that are consistent across replicates by chance, given the many statistical tests performed in genome-wide analyses, which would be difficult to distinguish from the action of selection on a per site basis. Therefore, *we hypothesize that E&R studies featuring sexual populations with low N_e will*

have limited power to detect causal variants, even if there is clear phenotypic differentiation, as well as sufficient replication and generations under selection.

Here we seek to test this hypothesis using genome-wide analysis of two experiments featuring selection for starvation resistance in *Drosophila melanogaster*. One of these experiments was conducted using N_e values of about 1,000, with ten control populations and ten selected populations (Kezos et al., in prep.). The other experiment was conducted with nine control populations and nine selected populations in which N_e had been intentionally compressed (Santos, 2018).

The ten starvation-selected "SCO" populations featured in Kezos et al. (in prep) have been undergoing selection since August 2010. This intense selection regime rapidly produced changes in body shape, body size, and stress resistance. The C-type populations ($CO_{1\text{-}5}$ and $nCO_{1\text{-}5}$) featured in Graves et al. (2017) serve as the controls for these populations. The small N_e populations, "pSB" and "pCB", were derived from the five "B" populations that have been maintained in the Rose lab since 1980 (Rose 1984). The B populations are maintained at a census size of ~2000 individuals. By contrast, census size was reduced to ~50 individuals in the pSB and pCB populations. The pSB populations were subjected to selection for starvation resistance for ~50 generations, while the pCB were maintained in parallel as controls. The response to starvation selection in these populations has been characterized by Santos (2018). Santos (2018) found that there were significant increases in starvation resistance in the pSB populations relative to the pCB populations, but the response to selection was significantly less than what was seen in populations maintained at higher N_e *that* were also subjected to same starvation selection protocol (Figure 4.1). To test our hypothesis, we assess our ability to detect significantly

differentiated genomic regions between the pSB and pCB populations using standard E&R approaches, with the genomic data from the SCO selection experiment serving as a contrasting point of reference with large N_e.

MATERIALS AND METHODS

Populations

The pSB (pSB_{1a-e}-pSB_{5a-e}) and pCB (pCB_{1a-e}-pCB_{5a-e}) populations were derived from five experimental populations of *D. melanogaster*, B_1-B_5, maintained in the Rose lab since 1980 (Rose 1984). The B populations are vial adapted and maintained on a 14-day life cycle. Each of the B population consists of ~2000 individuals and has an N_e of ~1000 based on estimates from demographic data (Mueller et al. 2013). A total of 10 populations were derived from each of the B population replicates, 5 pSB and 5 pCB per replicate. The pSB populations were subjected to selection for starvation resistance, while the pCB populations were maintained in parallel as controls. The pSB and pCB populations were maintained at a census size of ~50 individuals. At the time of derivation, the B populations had undergone ~795 generations of selection.

Starvation selection in the pSB populations consists of 5 steps. First, eggs from each of the starvation selected populations are collected and placed into a banana-based food medium and allowed to mature into adulthood over 14 days. On day 14, the flies are anesthetized using CO_2 and redistributed into vials of approximately 50 individuals each. These vials contain a high yeast banana-based food medium. After 3 days, the starvation selected populations (pSB) are placed in vials that contain an agar media, while the control populations are placed in vails containing the standard banana-based food medium. Agar is used for starvation conditions because it supplies water but does not provide any

nutritional value. For each of the starvation-selected populations, the number of flies in each vial is checked every 4 hours until 80% of individuals are dead. The surviving 20% are then used to breed the next generation. Following the starvation-selection phase, all flies are fed with high-yeasted banana medium for 3 days, and then their eggs are collected to found the next generation.

The SCO-a_{1-5} and SCO-b_{1-5} populations (28-day generation cycle) are ten populations intensely selected for starvation resistance. These populations were derived from the CO_{1-5} populations in August 2010, using the protocols published in Phelan et al. (2003). Briefly, after a two-week development period, flies are fed a high-yeast diet for three days before receiving a nonnutritive agar during the starvation period. Selected populations are then exposed to a nonnutritive agar until a 75-80% mortality threshold has been achieved. A three-day high yeast diet period follows the starvation period, and eggs are collected for the next generation. At the time of sequencing, the SCO populations had been subjected to ~75 generations of selection.

DNA Extraction and Sequencing for the Small N_e Populations

DNA was extracted and sequenced for 9 of the pSB and 9 of the pCB populations (Table S4.1). At the time of sequencing, the pSB and pCB populations had undergone ~50 generations of selection. DNA was extracted from samples of 200 female flies collected from each of the 18 populations chosen using Qiagen©/Gentra Puregene© kits, following the manufacturer's protocol for bulk DNA purification. The 18 gDNA pools were prepared as standard 200-300 bp fragment libraries for Illumina sequencing, and constructed such that each replicate populations of a treatment were given unique barcodes, normalized, and pooled together. Libraries were run across PE100 lanes of an Illumina HiSEQ 2000 at

the UCI Genomics Highthroughput Sequencing Facility. Resulting data were 100 bp paired-end reads. Each population was sequenced across three lanes; data from the three runs were combined for genomic analyses as described below. Samples were sequenced across three lanes to alleviate the effects of bias introduced from running entire sets of replicates on different lanes.

DNA Extraction and Sequencing for the SCO Populations

Genomic DNA was extracted from samples of 200 female flies collected from each of the 10 individual populations ($SCO_{1a,b}$-$SCO_{5a,b}$) using the Qiagen/Gentra Puregene kit, following the manufacturer's protocol for bulk DNA purification. The 30 gDNA pools were prepared as standard 200-300 bp fragment libraries for Illumina sequencing, and constructed such that each 5 replicate populations of a treatment were given unique barcodes, normalized, and pooled together. Libraries were run across PE100 lanes of an Illumina HiSEQ 2000 at the UCI Genomics Highthroughput Sequencing Facility. Resulting data were 100 bp paired-end reads.

Read Mapping for Small N_e and B populations

In addition to the pSB and pCB data, we also used data from their ancestral populations (B_{1-5}) published in Phillips et al. (2016). Samples used to generate the Phillips et al. (2016) data set were collected at generation ~785, which is ~10 generations before the pSB and pCB populations were derived. Fastq files from the three sets of populations were mapped to the D. melanogaster reference genome (version 6.14) using bwa mem with default settings (BWA version 0.7.8, Li and Durbin 2009). The resulting SAM files were filtered for reads mapped in proper pairs with a minimum mapping quality of 20, and converted to the BAM format using the view and sort commands in SAMtools (Li et al.

2009). The rmdup command in SAMtools was then used to remove potential PCR duplicates. As each population was sequenced across three lanes, there were three bam files corresponding to each population at this stage. BAMtools was used to combine files corresponding to the same populations (Barnett et al. 2011). Average coverage was above 50X for all populations except for a single pCB population, which was 48X (Table S4.1). Next, SAMtools was used to combine the 23 bam files into a single mpileup file. Using the PoPoolation2 software package (Kofler et al. 2011), these files were converted to "synchronized" files, which is a format that allele counts for all bases in the reference genome and for all populations being analyzed. RepeatMasker 4.0.3 (http://www.repeatmasker.org) was then used to create a gff file detailing low complexity regions in the D. melanogaster reference genome. The regions were then masked in our sync file once again using PoPoolation2.

Read Mapping for SCO and C-type Populations

In addition to the SCO DNA data described above, we also incorporated sequence data from the C-type populations from the Graves et al. (2017) data set. We mapped reads with BWA (version 0.7.8) (Li and Durbin 2009) against the *D. melanogaster* reference genome (version 6.14) using bwa mem with default settings. We filtered and sorted the resulting SAM files for reads mapped in proper pairs with a minimum mapping quality of 20 and converted them to the BAM using the view and sort commands in SAMtools (Li et al. 2009). The rmdup command in SAMtools was then used to remove potential PCR duplicates. Average coverage was above 40X or greater for all populations except SCO_{3a} and SCO_{5b} which were 29X and 26X respectively (Table S4.2). Next, the steps were repeated using the raw fastq files for the C-type populations, $CO_{1\text{-}5}$ and $nCO_{1\text{-}5}$, used in our Graves et

al. (2017) publication. This ultimately resulted in the creation of another 10 bam files. Bam files for all 20 populations were then combined into a single mpileup file using SAMtools. This mpileup file was then converted to a "synchronized" file using the PoPoolation2 software package (Kofler et al. 2011). This file displays allele counts for all positions in the reference genome for all populations in a succinct tab delimited format. RepeatMasker 4.0.3 (http://www.repeatmasker.org) to create a gff file containing simple sequence repeats found in the *D. melanogaster* genome version 6.14. These regions were then masked in the sync file.

Single Nucleotide Polymorphism (SNP) Variation for the Small N_e and B Populations

A SNP table was created using the sync file mentioned above. We only considered sites where coverage was between 30X and 200X, and for a site to be considered polymorphic we required a minimum minor allele frequency of 2% across all 23 populations. All sites failing to meet these criteria were discarded. To assess broad patterns of SNP variation in B, pSB, and pCB populations, heterozygosity was calculated and plotted over 150kb non-overlapping windows directly from the major and minor counts in our SNP table. To assess how closely replicate populations resembled one another, F_{ST} estimates were also obtained using the formula: $F_{ST} = (H_T - H_S)/H_T$ where H_T is heterozygosity based on total population allele frequencies, and H_S is the average subpopulation heterozygosity in each of the replicate populations (Hedrick 2009). F_{ST} estimates were made at every polymorphic site in the data set for a given set of replicates.

SNP Differentiation

The Cochran-Mantel-Haenzel (CMH) test as implemented in the PoPoolation2 software package was our primary means of assessing SNP differentiation between the pSB

and pCB populations. CMH test were performed at all positions meeting the SNP calling criteria stated above. When performing CMH tests, populations were paired based on evolutionary history (e.g. pSB_{1b} with pCB_{1b}, pSB_{1e} with pCB_{1e}, etc.). To correct for multiple comparisons, we used the permutation approach featured in Graves *et al.* (2017). Briefly, populations were randomly assigned to one of two groups and the CMH test was then performed at each polymorphic position in the shuffled data set to generate null distributions of p-values. This was done 1000 times, and each time the smallest p-value generated was recorded. The quantile function in R was then used to define thresholds that define the genome-wide false-positive rate, per site, at 5%. CMH tests were also performed comparing SNP frequencies in the B populations to the pSB, and pCB populations. These comparisons were done only using the "e" replicates of the pSB and pCB populations, once again populations were paired based on evolutionary history when performing the CMH tests.

The quasibionomial GLM approach advocated by Wiberg *et al.* (2017) was also used to asses SNP differentiation using the scripts they have made publicly available. As recommended by Wiberg et al. (2017), coverage at each position in our sync file was scaled to the effective sample size (n_{eff}) (Feder et al. 2012; Kolaczwoski et al. 2011). As counts of zero can lead to problems when implementing this approach, a count of 1 was added to each allele that otherwise had a count of 0. This approach was used to compare SNP frequencies between the pSB and pCB populations, and populations were once again paired based on evolutionary history. In terms of establishing significance thresholds, the quasibionomial GLM approach produces the expected uniform distribution of p-values under the null hypothesis and allows for standard methods of correction for multiple

comparisons (Wiberg et al. 2017). Thus, we corrected for multiple comparisons using the Bonferroni method instead of the permutation method used in our CMH tests.

Lastly, we compared the 9 of the 10 SCO populations to 9 of the 10 C-type populations featured in Graves et al. (2017) using both the CMH test and quasibionomial GLM approach. Tests were performed using the same methods used for the pCB versus pSB comparisons with the exception of SNP identification criteria. In this data set, we used a minimum coverage requirement of 15X instead of 30X due to limited coverage in specific SCO and C-type populations (Table S4.2).

RESULTS

Heterozygosity and F_{ST}

As previously documented in Philips et al. (2016) and Graves et al. (2017), we have not found any evidence of large scale depletions in heterozygosity in the B populations (Figure 4.2). And this pattern is largely robust to reductions in window size (Figures S4.1-S4.2). In the pSB and pCB populations, we find more of these depressions and this pattern appears even more clearly at smaller window sizes (Figure 4.2, Figure S4.1, Figure S4.2). On average across the B populations, we find ~213 50kb windows with heterozygosity estimates < 0.1. By comparison, we find 336 and 264 such windows on average in the pSB and pCB populations respectively.

Mean genome-wide heterozygosity is 0.23, 0.22 and, 0.21 among the B, pSB, and pCB populations, respectively (Figure S4.3, Table S4.3). Based on a one-way ANOVA, there is significant difference in mean heterozygosity based on selection regime (p-value = 0.03). Based on a Tukey's range test, there is a significant difference between the pCB and B groups (p-value = 0.02), but no significant differences between the pSB and B groups (p-

value = 0.35) or between the pSB and pCB groups (p-value = 0.21). Lastly, based on the formula $Ht = Ho\ (1 - \frac{1}{2N})^t$, we would expect heterozygosity after 50 generations (t) to be 0.14 in the reduced N_e populations due to drift alone given the mean level of heterozygosity in the ancestral B populations (*Ho*) and our census population size (N). This estimate is lower than what we observe in both the pCB and pSB populations, even though our use of the census population size in this calculation is conservative.

Mean genome-wide F_{st} was 0.21 and 0.25 in the pSB and pCB populations respectively, and 0.11 in the B populations. This indicates that there is a greater level of similarity between the B population replicates, than between the pSB or pCB replicates, a clear signature of increased genetic drift in the "p" replicate populations.

SNP Differentiation

CMH tests comparing SNP frequencies between the pSB and pCB did not identify any significantly differentiated sites (Figure 4.3A). In contrast, when we compared SNP frequencies between 9 of the 10 SCO populations, and 9 of the 10 C-type populations featured in Kezos et al. (in prep.), we identified a total of 806 differentiated SNPs (Figure 4.3B). But, unlike the pSB and pCB populations, the SCO and C-type populations are maintained at much higher census sizes (~2000 individuals per population). Our findings using the quasibionomial GLM approach qualitatively paralleled these CMH results. We find no evidence of significant SNP differentiation between the pSB and pCB populations (Figure 4.4A), but we identified a total of 306 significantly differentiated sites between the SCO and C-type populations (Figure 4.4B). This pattern also holds true when we correct for multiple comparisons using the less stringent q-value method (Figure S4.4).

DISCUSSION

Our failure to find any significant differences in mean heterozygosity between the pSB and pCB populations suggest that patterns of genetic variation are being primarily shaped by genetic drift, not selection. However, we found more genetic variation in the pSB and pCB populations than we would expect based on conventional population genetic theory. This finding is similar to that previously found for the B populations in Phillips et al. (2016). This suggests that, even at low N_e, there may be mechanisms at play promoting the maintenance of genetic variation across the genome. A possible explanation for this comes from Michalak et al. (2017), who found that selection for increased lifespan resulted in increased nucleotide diversity in selected lines compared to controls. They argue that this apparent increase in balancing selection is consistent with the antagonistic pleiotropy theory of aging, but it is perhaps applicable to starvation selection as well.

Although at the genome-wide scale mean levels of genetic variation are comparable between the B populations and pCB/pSB populations, we do find more regions at or near fixation in the latter. So, while reductions in population size in the pSB and pCB population may not have dramatically affected levels of genetic variation at the genome-wide scale, it appears to have resulted in a greater number of regional depressions. This finding could explain why, unlike the bulk of E&R studies using *D. melanogaster*, Turner et al. (2011) found instances of fixation in response to selection on body size. In the study of Turner et al. (2011), for each generation the most extreme 160 males and 160 females in terms of body size were used to breed the next generation. Given such low census sizes, their hypothesis that they observed instances of selection driving specific alleles to fixation could be incorrect. Their cases of fixation could instead be a result of low N_e.

The idea that genetic drift is playing a large role in shaping patterns of variation in these small populations is also supported by our F_{ST} results. Past studies E&R studies using the moderately outbred *D. melanogaster* population maintained in the Rose lab have consistently found a high degree of similarity between replicate populations subjected to the same selection regime, as indicated by low mean genome-wide F_{ST} among replicate populations that share a common regime (Phillips et al. 2016, Graves et al. 2017, Phillips et al. in review, Kezos et al. in prep). In those studies, the high degree of similarity between replicates was attributed to highly parallel genome-wide responses to selection. In contrast to these findings, here we observe less similarity between replicates in the pSB and pCB populations as revealed by elevated levels of F_{ST}, suggesting that within a given replicate selection is not the predominant force shaping genetic variation.

Santos (2018) found that while the pSB populations show a reduced phenotypic response to starvation selection, compared to starvation-selected populations maintained at a higher N_e, they do still respond to selection (Figure 4.1). But despite this observed phenotypic response to selection, our analysis did not identify any SNPs that were significantly differentiated between the pSB and pCB populations. This is also in contrast to the findings of Kezos et al. (in prep) where comparisons between the starvation selected SCO populations and their C-type controls resulted in the identification of hundreds of differentiated candidate SNPs. Unlike the pSB and pCB populations, the SCO and C-type populations featured in Kezos et al. (in prep.) were maintained at census sites in excess of 2000 individuals per replicate. It is also worth noting that, despite featuring over twice the level of replication seen in many published E&R studies that have identified dozens to hundreds of candidate SNPs, we nonetheless failed to identify any consistently

differentiated sites between the pSB and pCB populations (Burke et al. 2010; Orzoco-terWengel et al. 2012; Tobler et al. 2014, Franssen et al. 2014, Huang et al 2014). Taking all these results together, *we conclude that N_e is indeed a major determinant of statistical power in E&R studies, corroborating our central hypothesis*.

To summarize, the present findings support our hypothesis that E&R studies featuring populations with low N_e will have limited power to detect genetic variants underlying any observed phenotypic response to selection. We find evidence that at low N_e, genetic drift plays a larger role in shaping patterns of genetic variation than seen in our previous work. The reduced parallelism across replicates that results then in turn reduces our power to identify candidate SNPs, despite high replication and a clear phenotypic response to selection. Much like replication, we conclude that N_e is an important experimental parameter to maximize in E&R studies aimed at deciphering the genetic architecture of complex phenotypes.

REFERENCES

Allendorf FW, Luikart G. Conservation and the genetics of populations. Oxford: Blackwell Publishing; 2009.

Baldwin-Brown JG, Long AD, Thornton KR. The power to detect quantitative trait loci using resequenced, experimentally evolved populations of diploid, sexual organisms. Mol Biol Evol. 2014;31:1040–1055.

Barnett DW, Garrison EK, Quinlan AR, Stromberg MP, Marth GT BamTools: a C ++ API and toolkit for analyzing and managing BAM files. Bioinformatics. 2011;27:1691–1692.

Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. Nature. 2010;467:587–590.

Burke MK, Liti G, Long AD. Long standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. Mol Biol Evol. 2014;32:3228–3239.

Feder AF, Petrov DA, Bergland AO. LDx: estimation of linkage disequilibrium from high-trhoughout pooled resequncing data. PLos OONE. 2012;doi:10.1371/journal.pone.0048588

Frankham R. Stress and adaption in conservation genetics. J Evol Biol. 2005;18:750-755

Franssen SU, Volte N, Tobler R, Schlötterer C. Patterns of linkage disequilibrium and long range hitchhiking in evolving experimental *Drosophila melanogaster* populations. Mol Biol Evol. 2015;32:495–509.

Graves JL, Hertweck KL, Phillips MA, Han MV, Cabral LG, Barter TT, et al. Genomics of parallel experimental evolution in Drosophila. Mol Biol Evol. 2017;34:831-842

Hedrick PW. Genetics of populations. Massachusetts: Jones & Bartlett Learning Press; 2009.

Huang Y, Wright SI, Agrawal AF. Genome-wide patterns of genetic variation within and among alternative selective regimes. PLoS Genet. 2014;10, e1004527.

Kezos JK, Phillips MP, Barter TT, Santos MA, Brandon DW, Arnold KJ, et al. The effects of intense selection for starvation resistance on Drosophila physiology. In prep.

Kofler R, Pandey RV, Schlötterer C. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics. 2011;27:3435–3436.

Kofler R, Schlötterer C. A guide for the design of evolve and resequencing studies. Mol Biol Evol. 2014;31:474–483.

Kolaczkowski B, Kern AD, Holloway AK, Begun DJ. Genomic differentiation between temerpate and tropical Australian population of *Drosophila melanogaster*. Genetics. 2011;187:245-260.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–2079.

Long AD, Liti G, Luptak A, Tenaillon O. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. Nat Rev Genet. 2015;16:567–582.

Madalena FE, Robertson A. Population structure in artificial selection: studies with *Drosophila melanogaster*. Genetical Research. 1975;24:113-126.

Marden JH, Wolf MR, Weber KE. Aerial performance of *Drosophila melanogaster* from populations selected for upwind flight ability. J Exp Biol. 1997;200:2747-2755.

Michalak P, Kang L, Sarup PM, Schou MF, Loeschcke V. Nucleotide diversity inflation as a genome-wide response to experimental lifespan extension in Drosophila melanogaster. BMC Genomics. 2017;18:84.

Mueller LD, Joshi A, Santos M, Rose RM. Effective population size and evolutionary dynamics in outbred laboratory populations of Drosophila. J Genet. 2013;92:349-361.

Orozco-ter Wengel P, Kapun M, Nolte V, Kofler R, Flatt T, Schlötterer C. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected traits. Mol Ecol. 2012;21:4931–4941.

Phelan JP, Archer MA, Beckman KA, Chippindale AK, Nusbaum TJ, Rose MR. Breakdown in correlations during laboratory evolution. I. Comparative analyses of *Drosophila* populations. Evolution. 2003;57: 527-535.

Phillips MA, Long AD, Greenspan ZS, Greer LF, Burke MK, Bryant V, et al. Genome-wide analysis of long-term evolutionary domestication in Drosophila melanogaster. Sci. Rep. 2016. doi:10.1038/srep39281

Roberstson A. A theory of limits in aritifical selection. Proc R Soc Lond B Biol Sci. 1960;153:234-249.

Rose MR. Laboratory evolution of postponed senescence in *Drosophila melanogaster*. Evolution. 1984;38:1004–1010.

Santos, M. Evolutionary lessons from *Drosophila melanogaster* during colonization. How do history, selection, and effective size shape evolution? (Doctoral dissertation) 2018. Retrived from Repository of NOVA University (RUN).

Schlötterer C, Kofler R, Versace E, Tobler R, Franseen SU. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. Heredity 2015;114:331–440.

Tobler R, Franssen SU, Kofler R, Orozco-ter Wengel P, Nolte V, Hermisson J, Schlötterer C. Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. Mol Biol Evol. 2014;31:364–375.

Teotónio H, Chelo IM, Bradiá M, Rose MR, Long AD. Experimental evolution reveals natural selection on standing genetic variation. Nat Genet. 2009;41:251–257.

Turner TL, Steward AD, Fields AT, Rice WR, Tarone AM. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. PLoS Genet.2011; 7:e10001336.

Weber KE, Diggins LT. Increased selection response in larger populations. II. Selection for ethanol vapor resistance in *Drosophila melanogaster* at two population sizes. Genetics. 1990;125:585-597.

Wiberg RAW, Gaggiotti OE, Morrissey MB, Ritchie MG, Johnson L. Identifying consistent allele frequency differences in studies of stratified populations. Methods Ecol Evol. 2017;8:1899-1909.

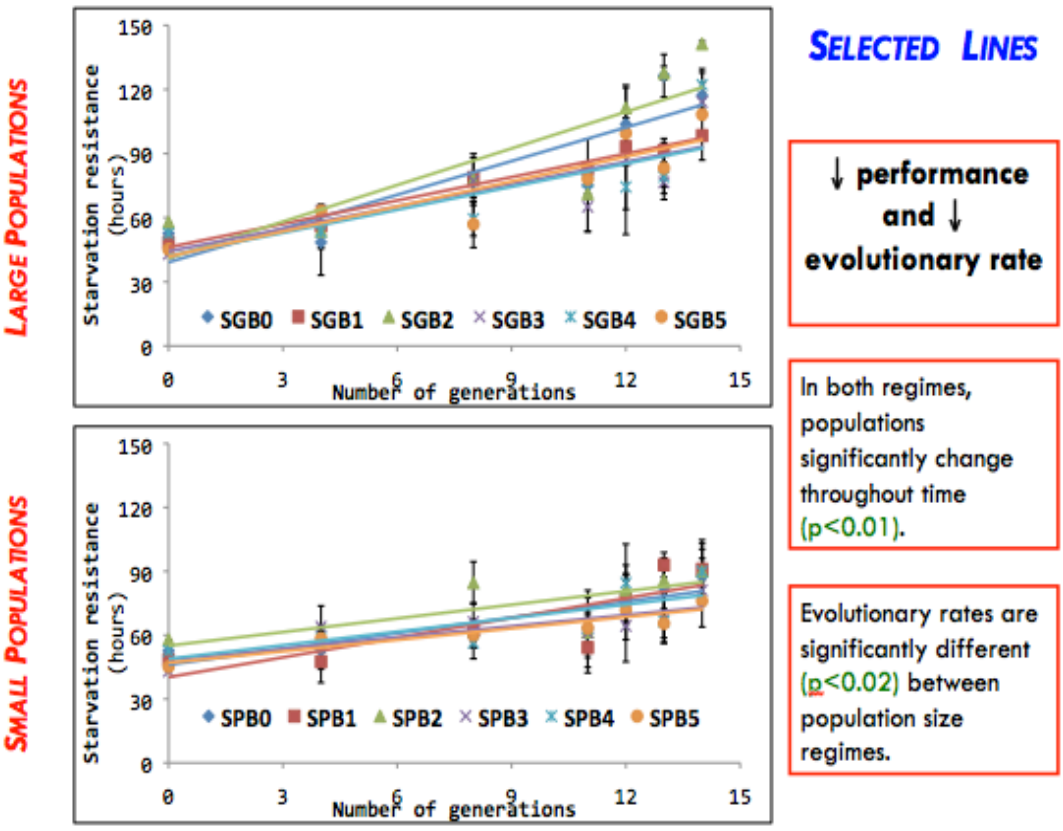

Figure 4.1. Starvation selection results. Results from Santos (2018) contrasting rates of initial response to starvation selection in 12 large N_e (SGB) compared to the response to selection in the pSB (termed SPB in their paper).

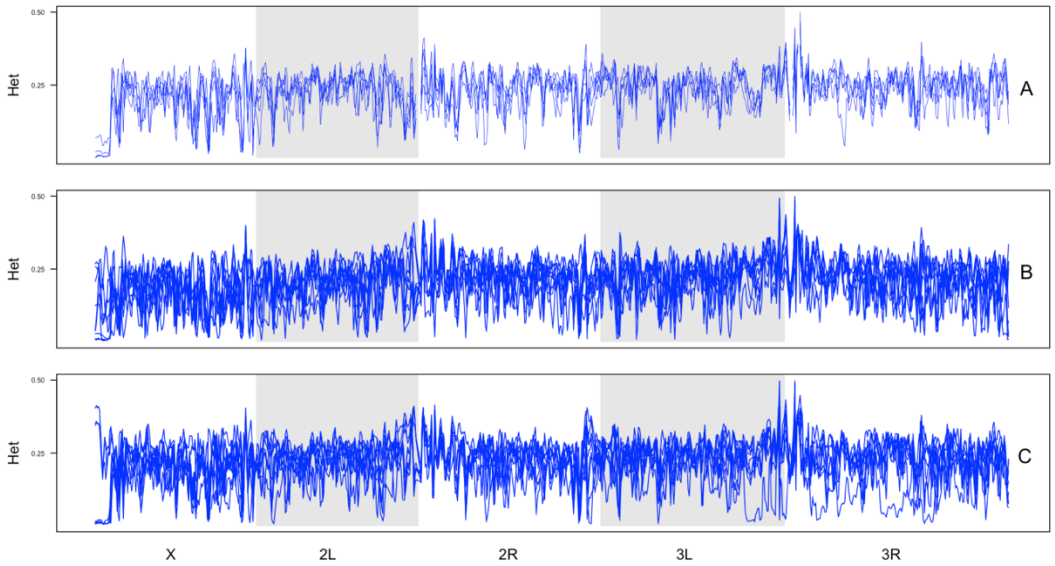

Figure 4.2. Genome-wide heterozygosity 150-kb windows. Heterozygosity in the B populations (A), pCB populations (B), and pSB (C) populations plotted over 150-kb windows across all major chromosome arms. All replicates are shown for each population.

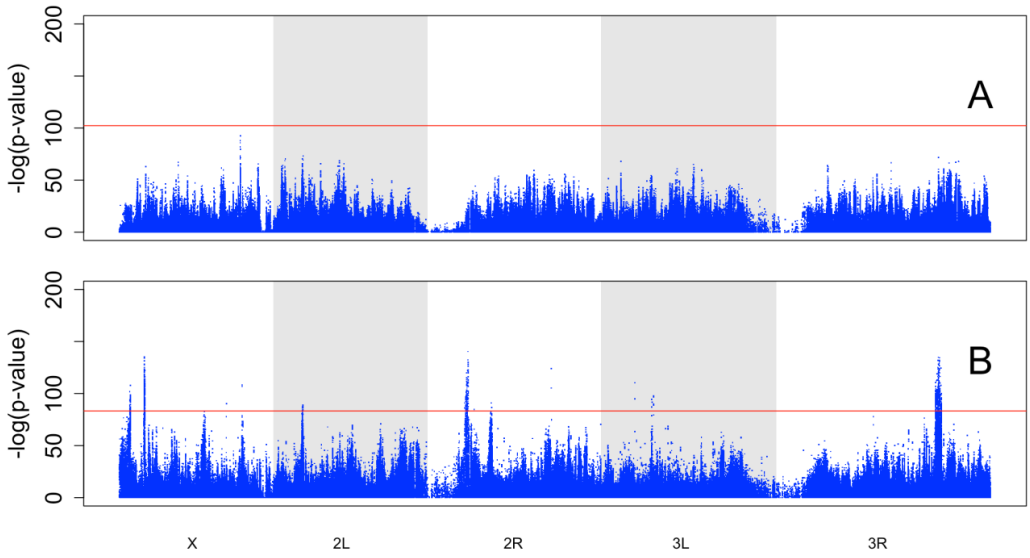

Figure 4.3. CMH test results. Results from CMH tests comparing SNP frequencies in the pSB and pCB populations (A), and the starvation resistant populations (SCO) and control populations (C-type) featured in Kezos et al. (in prep). To be consistent, only 9 of the 10 SCO and 9 of the 10 C-type populations were used in the latter comparison. Permutation derived significance thresholds are indicated by the red lines.

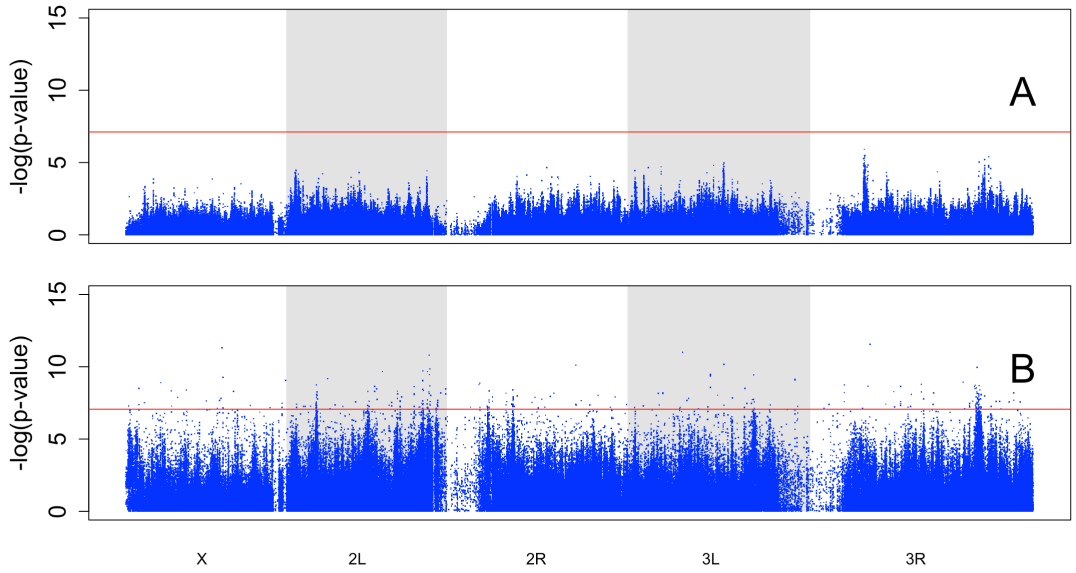

Figure 4.4. Quasibinomial GLM results. Results from quasibinomial GLM approach comparing SNP frequencies in the pSB and pCB populations (A), and the starvation resistant populations (SCO) and control populations (C-type) featured in Kezos et al. (in prep) (B). To be consistent, only 9 of the 10 SCO and 9 of the 10 C-type populations were used in the latter comparison. Bonferroni corrected significance thresholds are shown in red.

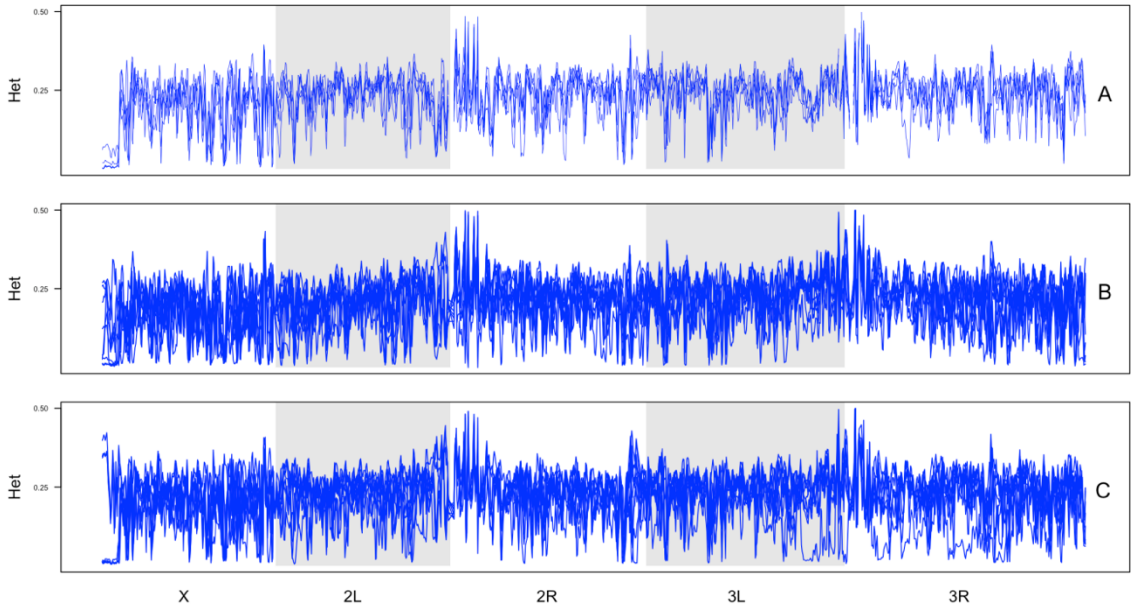

Figure S4.1. Genome-wide heterozygosity 100-kb windows. Heterozygosity in the B populations (A), pCB populations (B), and pSB (C) populations plotted over 100-kb windows across all major chromosome arms. All replicates are shown for each population.

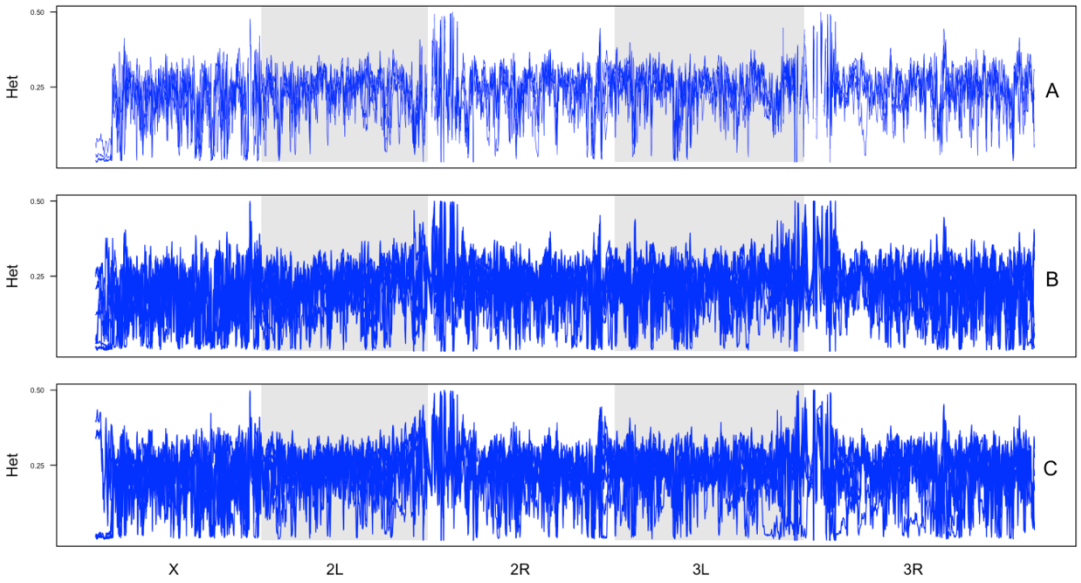

Figure S4.2. Genome-wide heterozygosity 50-kb windows. Heterozygosity in the B populations **(A)**, pCB populations **(B)**, and pSB (**C)** populations plotted over 50-kb windows across all major chromosome arms. All replicates are shown for each population.

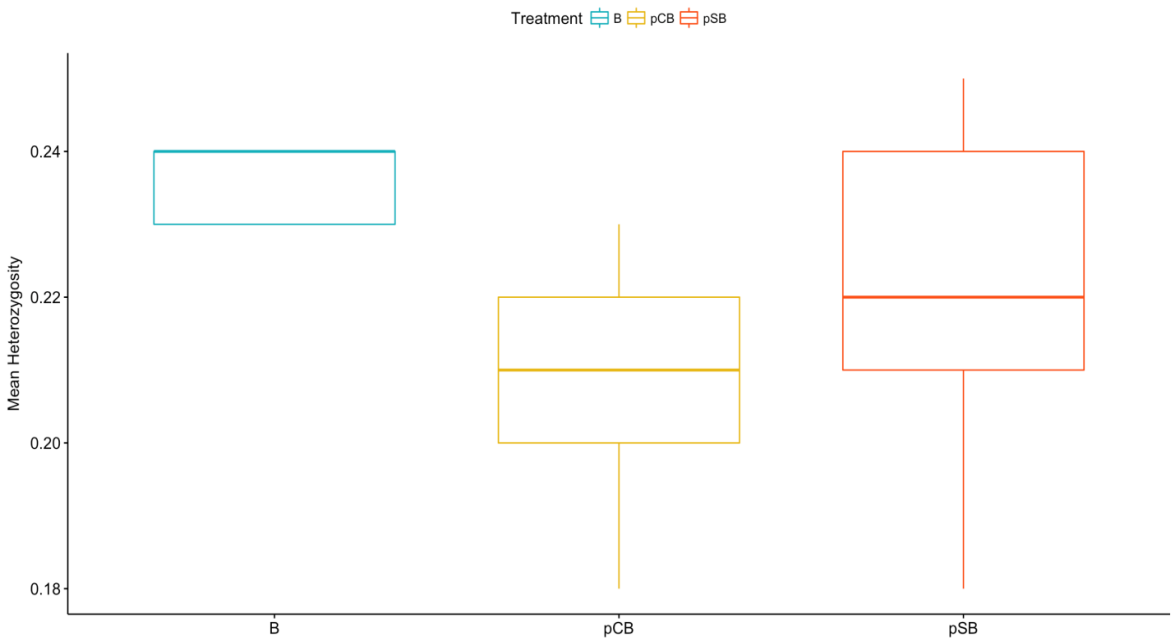

Figure S4.3. Mean genome-wide heterozygosity. Boxplot of mean genome-wide heterozygosity in the B, pCB, and pSB populations.

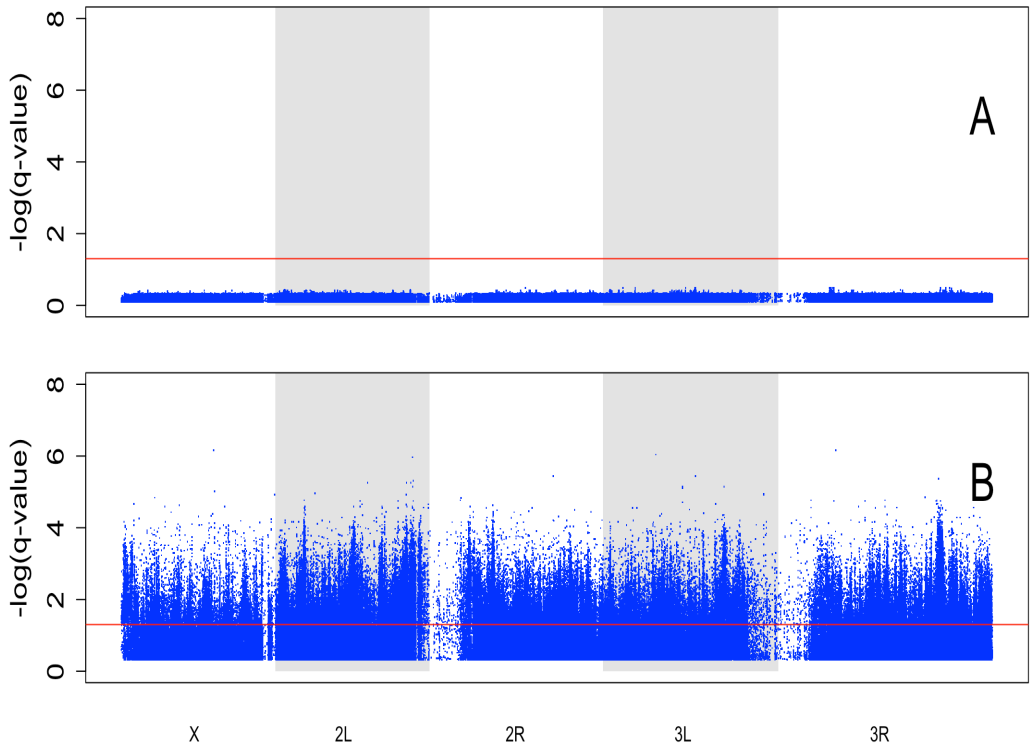

Figure S4.4. Additional Quasibinomial GLM results. Results from quasibinomial GLM approach comparing SNP frequencies in the pSB and pCB populations (A), and the starvation resistant populations (SCO) and control populations (C-type) featured in Kezos et al. (in prep) (B). To be consistent, only 9 of the 10 SCO and 9 of the 10 C-type populations were used in the latter comparison. Results are plotted as -log(q-values) and a 0.05 false discovery rate threshold was used, as shown in red.

Table S4.1. Average read coverage across the genome for all populations used in this study

Treatment	Replicate	Mean Coverage
B	1	58
	2	62
	3	55
	4	61
	5	67
pCB	1b	57
	1e	50
	2b	57
	2e	61
	3e	59
	4b	60
	4e	63
	5b	56
	5e	48
pSB	1b	58
	1e	56
	2b	52
	2e	69
	3e	55
	4b	53
	4e	53
	5b	56
	5e	57

Table S4.2. Average read coverage across the genome for all populations used in this study

Population	Replicate	Average Read Coverage
SCOa	1	46
	2	43
	3	29
	4	72
	5	62
SCOb	1	49
	2	45
	3	54
	4	60
	5	26
CO	1	53
	2	104
	3	120
	4	83
	5	28
nCO	1	109
	2	65
	3	66
	4	103
	5	56

Table S4.3. Mean genome-wide heterozygosity calculated from SNP data for B, pCB, and pSB populations

Treatment	Replicate	Mean Heterozygosity
B	1	0.23
	2	0.24
	3	0.24
	4	0.24
	5	0.23
pCB	1b	0.21
	1e	0.18
	2b	0.20
	2e	0.23
	3e	0.20
	4b	0.22
	4e	0.20
	5b	0.21
	5e	0.22
pSB	1b	0.24
	1e	0.21
	2b	0.24
	2e	0.25
	3e	0.20
	4b	0.22
	4e	0.18
	5b	0.24
	5e	0.22

CONCLUSIONS

Early evolve and resequence ("E&R") studies featuring sexually reproducing populations were typically limited in both duration and replication. As such, while pioneering studies using the approach suggested that there were fundamental differences in adaptation between sexual and asexual populations, many of these differences could have been accounted for by experimental limitations. For instance, early studies indicated that adaptation in asexual populations was driven by beneficial *de novo* mutations and characterized by hard sweeps, while adaptation in sexual populations was primarily fueled by standing genetic variation and therefore characterized as based on "soft sweep" events (Burke 2012). However, it is possible such discrepancies were due to the problem that sexual E&R studies are often limited in duration, replication, and population size (e.g. Burke et al. 2010; Turner et al. 2011; Orozco-terWengel et al. 2012, Tobler et al. 2014, Huang et al. 2014; Franssen et al. 2015) compared to asexual E&R studies (Barrick et al. 2009; Tenaillon et al. 2012; Maddamsetti et al. 2015). As a whole, this dissertation sought to address some of these concerns. In addition, it offers new insights into adaptation in sexual populations using the dozens of experimentally evolved populations *Drosophila melanogaster* maintained in the Rose Lab at the University of California, Irvine.

Chapter 1 of my thesis involved a group of population subjected to the same laboratory domestication regime for nearly ~1000 generations. As such, it provides a reasonable opportunity to characterize the long-response to selection in sexual populations. I found that even after 1000 generations of selection in moderately outbred populations, there was no evidence of hard sweeps involving d*e novo* and associated haplotypes fixing in the populations. In fact, population genetic simulations indicated there

that there was more genetic variation being maintained in these populations than predicted by conventional genetic theory. This led to two conclusions. First that, even on longer time scale, adaptation in sexual E&R studies is still primarily driven by standing genetic variation. Second, there are mechanisms in play favoring the genome-wide maintenance of genetic variation. The former conclusion is strongly supported by work in outcrossing yeast populations, where even at very at much higher population sizes and hundreds of generations of selection, adaptation is still primarily fueled by standing genetic variation (Burke et al. 2014). The latter is somewhat supported by work from Michalak et al. (2017), where it was reported that selection for increased longevity resulted in increased nucleotide diversity. The authors argue that this is consistent with the antagonistic pleiotropy theory of aging, an idea that is also potentially applicable in other contexts.

Chapter 2 of my thesis studied newly-derived (dozens of generations under selection) and long-standing (hundreds of generations under selection) populations subjected to the same selection regimes, in order to compare long versus short-term responses to selection. It took as little as ~40 generations for newly derived populations to converge on long standing populations phenotypically (see Burke et al. 2016 for a more detailed phenotypic characterization than presented in Chapter 2). Similarly, I found similarly rapid convergence genomically, with respect to SNPs. (See Graves et al. 2017 for transposable element and structural variant analyses.) These findings led to the following conclusions regarding evolution in outbred sexual populations: (i) adaptation can be fast and repeatable at both phenotypic and genetic levels due to standing genetic variation; (ii) adaptation does not wait for new functional mutations and subsequent hard sweeps; (iii)

E&R studies using sexual outbred populations do not need to run for hundreds of generations to provide useful insights.

Chapter 3 of my thesis used data from populations previously subjected to intense selection for desiccation resistance in order to study the role of evolutionary history in sexual E&R studies. Here I tested the following hypotheses: Hypothesis 1: evolutionary history matters in populations where past generations were subjected to very intense selective pressures; Hypothesis 2: previous intense selection does not produce historical effects, in the absence of other factors such as inbreeding or chromosomal rearrangement. My analyses found no evidence of genetic fixation in intensely selected populations, and very limited genetic differentiation between these populations and their long-standing controls after long-term relaxation of selection. The lack of present differentiation between previously desiccation-selected populations and their controls was also reflected across an array of phenotypic characters (Phillips et al. in review). Taken together, these findings support Hypothesis 2. I conclude extreme selection does not have major long-lasting impacts on genomic or phenotypic differentiation in Drosophila experimental evolution. This conclusion is also supported by the findings of Burke et al. (2016) and Graves et al. (2017), where most recent selection regime was found the primary determinant of patterns of genetic variation and phenotypic measure.

Lastly, Chapter 4 of my thesis sought to empirically test how population size impacts the power to detect causal variants in sexual E&R studies using findings from two starvation selection experiments: one with moderately outbred populations, and the other with populations where population size was deliberately compressed. Specifically, I test the hypothesis that due to the increased prominence of genetic drift, E&R studies featuring

populations with low effective population sizes (N_e) will have limited power to detect casual variants, even if there is clear phenotypic differentiation, as well as sufficient replication and generations under selection. Here I find that while more genetic variation is still being maintained in population where N_e was deliberately compressed than population genetic theory would predict, patterns of differentiation between replicate populations as indicated by F_{st} estimates indicate drift is playing a more prominent role when compared to studies using populations with larger N_e (Phillips et al. 2016; Graves et al. 2017). I also found that SNP frequency comparisons between starvation-selected populations and controls maintained at a larger N_e resulted in the identification of hundreds of candidate sites, comparisons between selected and control population maintained at low N_e failed to detect any candidate sites. Given that both experiments ran for dozens of generations and featured twice the level of replication seen in sexual E&R studies that successfully identified candidate sites (Burke et al. 2010; Orzoco-terWengel et al. 2012; Tobler et al. 2014, Franssen et al. 2014, Huang et al 2014), these findings support our hypothesis that E&R studies featuring populations with low *Ne* will have limited power to detect genetic variants underlying phenotypic responses to selection. As such, I conclude, much like replication, N_e is an important experimental parameter to maximize in E&R studies aimed at deciphering the genetic architecture of complex phenotypes.

References

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, et al. Genome evolution and adaptation in long-term experiment with *Escherichia coli*. Nature. 2009;461:1243-1274.

Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. Nature. 2010;467:587–590.

Burke MK. How does adaptation sweep through the genome? Insights from long-term selection experiments. Proc R Soc. 2012;279:5029-5038.

Burke MK, Liti G, Long AD. Long standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. Mol Biol Evol. 2014;32:3228–3239.

Burke MK, Barter TT, Cabral LG, Kezos JN, Phillips MA, Rutledge GA, et al. Rapid convergence and divergence of life-history in experimentally evolved *Drosophila Melanogaster*. Evolution 2016;70:2085–2098.

Franssen SU, Volte N, Tobler R, Schlötterer C. Patterns of linkage disequilibrium and long range hitchhiking in evolving experimental *Drosophila* melanogaster populations. Mol Biol Evol. 2015;32:495–509.

Graves JL, Hertweck KL, Phillips MA, Han MV, Cabral LG, Barter TT, et al. Genomics of parallel experimental evolution in Drosophila. Mol Biol Evol. 2017;34:831-842.

Huang Y, Wright SI, Agrawal AF. Genome-wide patterns of genetic variation within and among alternative selective regimes. PLoS Genet. 2014;10, e1004527.

Maddamsetti R, Lenski RE, Barrick JE. Adaptation, clonal interference, and frequency-dependent interactions in a long-term evolution experiment with *Escherichia coli.* Genetics 2015;200:619-631.

Michalak P, Kang L, Sarup PM, Schou MF, Loeschcke V. Nucleotide diversity inflation as a genome-wide response to experimental lifespan extension in Drosophila melanogaster. BMC Genomics. 2017;18:84.

Orozco-ter Wengel P, Kapun M, Nolte V, Kofler R, Flatt T, Schlötterer C. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected traits. Mol Ecol. 2012;21:4931–4941.

Phillips MA, Long AD, Greenspan ZS, Greer LF, Burke MK, Bryant V, et al. Genome-wide analysis of long-term evolutionary domestication in *Drosophila melanogaster.* Sci. Rep. 2016. doi:10.1038/srep39281

Phillips MA, Rutledge GA, Kezos JN, Talbott A, Matty S, Arian H, et al. Effects of evolutionary history on genome wide and phenotypic convergence in Drosophila populations. In review BMC Genomics.

Tenaillon O, Rodriguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, et al. The molecular diversity of adaptive convergence. Science. 2012;335:457-461.

Tobler R, Franssen SU, Kofler R, Orozco-ter Wengel P, Nolte V, Hermisson J, Schlötterer C. Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. Mol Biol Evol. 2014;31:364–375.

Turner TL, Steward AD, Fields AT, Rice WR, Tarone AM. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. PLoS Genet.2011; 7:e10001336.