12-22-2009

The Evolutionary Genetics of Life History in Drosophila Melanogaster

Annalise B. Paaby
University of Pennsylvania, paaby@sas.upenn.edu

The Evolutionary Genetics of Life History in Drosophila Melanogaster

Abstract
Life history traits are critical components of fitness and frequently reflect adaptive responses to environmental pressures. Natural populations of Drosophila melanogaster exhibit patterns of lifespan, fecundity, development time, body size and stress resistance that vary predictably along environmental gradients. Artificial selection studies, genetic correlation analyses, and quantitative trait mapping efforts have demonstrated a genetic basis for the observed phenotypic variation, but few genes have been identified that contribute to natural life history variation. This work employs a candidate gene approach to discover genes and specific polymorphisms that contribute to genetic variance for D. melanogaster life history. Three aging genes, which have been characterized to mediate longevity, reproduction and stress tolerance, have been evaluated for natural genetic variation from samples derived from the wild. Allelic variation at one gene, methuselah (mth), shows functional effects on lifespan, lifetime fecundity and resistance to oxidative stress. A polymorphism in the mth promoter has been identified which may contribute to variation in these traits by affecting levels of gene expression. Natural genetic variation at two genes in the insulin signaling pathway reveals a history of positive selection at the Insulin-like Receptor (InR), but evidence of neutral evolution at the InR substrate, chico. Furthermore, an indel polymorphism in the first exon of InR shows striking, nonrandom distributions on two continents, a sign that it may contribute to the observed patterns of phenotypic variation across these same habitats. Functional evaluation of alternate InR alleles demonstrates predictable effects on phenotype and levels of insulin signaling, which implicates this polymorphism in the adaptive evolution of wild D. melanogaster populations. These findings provide novel examples of how allelic variation underlies adaptive changes in life history evolution, and contribute complementary characterization of genetic function to the biology of aging.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Biology

First Advisor
Paul S. Schmidt

Keywords
life history, lifespan, genetic variation, natural populations, Drosophila

Subject Categories
Evolution

THE EVOLUTIONARY GENETICS OF LIFE HISTORY IN *DROSOPHILA MELANOGASTER*

Annalise B. Paaby

A DISSERTATION

in

Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2009

Supervisor of Dissertation:

_____

Paul S. Schmidt
Associate Professor of Biology

Graduate Group Chairperson:

_____

Paul D. Sniegowski
Associate Professor of Biology

Dissertation Committee:

Nancy M. Bonini, Professor of Biology, University of Pennsylvania
Arthur E. Dunham, Professor of Biology, University of Pennsylvania
Warren J. Ewens, Professor of Biology, University of Pennsylvania
Joshua B. Plotkin, Assistant Professor of Biology, University of Pennsylvania
Paul S. Schmidt, Associate Professor of Biology, University of Pennsylvania
Paul D. Sniegowski, Associate Professor of Biology, University of Pennsylvania
Marc Tatar, Professor of Ecology and Evolutionary Biology, Brown University

The evolutionary genetics of life history in *Drosophila melanogaster*

ACKNOWLEDGEMENTS

It is a pleasure to thank my thesis advisor, Paul Schmidt, whose generosity in both practical affairs and enthusiasm provided a bedrock foundation for this research, and who has been an inspirational mentor to me. I thank the members of my committee, Nancy Bonini, Art Dunham, Warren Ewens, Josh Plotkin, Paul Sniegowski and Marc Tatar, for their service; I am honored to have shared my work with them. Thanks to past and present members of the Schmidt lab for good-natured technical assistance and good-humored company, and thanks to many members of the Biology Department for support, both academic and social, which is remembered with affection. I am especially grateful to Peter Petraitis and Warren Ewens, who are peerless role models in the academia that I have come to love, and whose friendliness and encouragement is cherished. The research was improved by funding from the American Federation for Aging Research and the Binns-Williams Fund, and by collaboration with Australian colleagues Mark Blacket and Ary Hoffmann, whose excellent data were contributed with expedience and good spirit. It is much appreciated. Finally, I am indebted to my parents, ever supportive, and Andrew, ever patient, who continually reminds me of the rewards of intellectual curiosity and the pleasures of exploring it together.

ABSTRACT

THE EVOLUTIONARY GENETICS OF LIFE HISTORY IN

DROSOPHILA MELANOGASTER

Annalise B. Paaby

Paul S. Schmidt

Life history traits are critical components of fitness and frequently reflect adaptive responses to environmental pressures. Natural populations of *Drosophila melanogaster* exhibit patterns of lifespan, fecundity, development time, body size and stress resistance that vary predictably along environmental gradients. Artificial selection studies, genetic correlation analyses, and quantitative trait mapping efforts have demonstrated a genetic basis for the observed phenotypic variation, but few genes have been identified that contribute to natural life history variation. This work employs a candidate gene approach to discover genes and specific polymorphisms that contribute to genetic variance for *D. melanogaster* life history. Three aging genes, which have been characterized to mediate longevity, reproduction and stress tolerance, have been evaluated for natural genetic variation from samples derived from the wild. Allelic variation at one gene, *methuselah* (*mth*), shows functional effects on lifespan, lifetime fecundity and resistance to oxidative stress. A polymorphism in the *mth* promoter has been identified which may contribute to variation in these traits by affecting levels of gene expression. Natural genetic variation at two genes in the insulin signaling

pathway reveals a history of positive selection at the *Insulin-like Receptor* (*InR*), but evidence of neutral evolution at the *InR* substrate, *chico*. Furthermore, an indel polymorphism in the first exon of *InR* shows striking, nonrandom distributions on two continents, a sign that it may contribute to the observed patterns of phenotypic variation across these same habitats. Functional evaluation of alternate *InR* alleles demonstrates predictable effects on phenotype and levels of insulin signaling, which implicates this polymorphism in the adaptive evolution of wild *D. melanogaster* populations. These findings provide novel examples of how allelic variation underlies adaptive changes in life history evolution, and contribute complementary characterization of genetic function to the biology of aging.

TABLE OF CONTENTS

LIST OF TABLES

Chapter Two

Chapter Four

Chapter Five

Chapter Six

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Empirical exploration of life history evolution. Understanding how natural selection shapes organismal traits requires teasing apart the mechanisms and dynamics governing the complex interactions between genotypes, phenotypes and the environment. Nonrandom patterns of phenotypic variation across environmental heterogeneity can indicate adaptive responses to selection, and evaluating these patterns can yield insight into evolutionary dynamics. For example, variation in coat pigmentation in the oldfield mouse *Peromyscus polionotus* appears to be an adaptation to changes in substrate reflectance (Mullen & Hoekstra 2008), and flower color in the desert plant *Linanthus parryae* likely represents adaptation to different soil environments (Schemske & Bierzychudek 2007). In particular, clinal variation in phenotypes across environmental gradients is often indicative of natural selection, and suggestive of mechanisms responsible for the maintenance of functional variation at the nucleotide level.

In *Drosophila melanogaster*, the characterization of nucleotide targets of selection is facilitated by the distribution of natural populations across latitudinal gradients. *D. melanogaster* originated in tropical Africa, and has colonized temperate regions in the North American and Australian continents within the last few hundred years (David & Capy 1988). In this system, surveys at multiple loci have revealed patterns of variation along latitudinal clines in which the frequencies of the derived alleles increase with latitude, suggesting

adaptation to novel, temperate habitats (Sezgin et al. 2004; Hoffmann & Weeks 2007). In addition to clinal variation at the genotypic level, natural populations of *D. melanogaster* also exhibit clinal variation in phenotypes, including life history traits: high latitude populations exhibit longer lifespan, lower fecundity, higher incidence of reproductive diapause, larger body size, and higher cold stress tolerance than low latitude populations (Capy et al. 1993; Mitrovski & Hoffmann 2001; De Jong & Bochdanovits 2003; Schmidt et al. 2005a; Trotta et al. 2006). Organismal life history is closely tied to fitness, and these phenotypic patterns are probably indicative of adaptive responses to the local environment. But life history traits are also highly quantitative, so discovering the genetic basis of adaptive variation in life history phenotypes requires careful parsing of complex and correlated phenomena.

Natural populations of *D. melanogaster* have been shown to harbor substantial allelic variation for life history traits, through artificial selection experiments (e.g. Rose & Charlesworth 1981; Promislow et al. 1996), genetic correlation analyses (e.g. Tatar et al. 1996; Schmidt et al. 2005b), and identification of aging genes or chromosomal regions through QTL analysis (Nuzhdin et al. 1997; Leips & Mackay 2000; Mackay 2002; De Luca et al. 2003; Geiger-Thornsberry & Mackay 2004; Pasyukova et al. 2004; Wilson et al. 2006). However, despite well-characterized life history variation with a clear genetic basis, few genes have been shown to contribute to these observed phenotypic patterns (De Luca et al. 2003; Carbone et al. 2006; Paaby & Schmidt 2008; Schmidt et al. 2008).

Longevity genes are good candidates in the search for genic elements that determine life history variation in the wild. Genes identified to regulate aging have routinely showed pleiotropic effects on traits that exhibit phenotypic correlations in natural populations (e.g.

Lin et al. 1998; Clancy et al. 2001; Tatar et al. 2001). *D. melanogaster* in particular has been an important model for studying the evolutionary genetics of lifespan (Paaby & Schmidt 2009). Three of the best characterized aging genes in *D. melanogaster* include *methuselah* (*mth*), which encodes a G-protein coupled receptor involved in neuroendocrine signaling, and the *Insulin-like Receptor* (*InR*) and its substrate, *chico*, two genes in the insulin signaling pathway. All three genes confer lifespan extensions when mutated, but also increases in stress tolerance and compromises to reproductive success (Lin et al. 1998; Tatar et al. 2001; Clancy et al. 2001; Mockett & Sohal 2006).

Genetic variance for and genetic correlations among these traits in natural populations indicate that selection in the local environment may act on certain phenotypes but drive expression of others through tradeoffs (David 1975; Anderson et al. 2003; De Jong & Bochdanovits 2003; Schmidt et al. 2005b; Rako et al. 2007; Schmidt & Paaby 2008). Variation in stress tolerance has been especially well described, and may be a key component in life history evolution (Hoffmann & Harshman 1999; Hoffmann et al. 2001; Hoffmann et al. 2005; Hoffmann et al. 2007). This framework suggests a hypothetical selection regime: high latitude, seasonally cold climates impose seasonal stress and favor genotypes that confer stress tolerance; correlated traits, which may evolve as co-adapted responses to the same selection regime or by indirect selection via pleiotropy, are characterized by better overwintering ability, larger body size, longer lifespan, and lower fecundity relative to low latitude populations (Paaby & Schmidt 2009).

This body of work evaluates natural genetic variation at three *D. melanogaster* aging genes, *mth*, *InR*, and *chico*, in an effort to identify the specific genic elements that contribute

to life history evolution. In addition to elucidating genetic mechanisms of adaptation, this research also addresses several major questions in the fields of life history genetics and evolution. For example, it is largely unknown whether the pronounced tradeoffs in organismal life history are governed by pleiotropy of individual elements, or if whole genes and pathways, which otherwise determine multiple phenotypes, may be parsed to reveal polymorphisms that act independently. One study has shown that single nucleotide polymorphisms in the pleiotropic gene *Catsup* affect traits singly (Carbone et al. 2006), but my work at *InR* suggests that this gene may harbor a polymorphism that directly mediates levels of insulin signaling and affects multiple traits. Furthermore, it is unresolved whether the genetic determination of life history is principally governed by many genes of small effect, or by a few genes of large effect; nor do we know the degree to which aging genes discovered by mutational analyses are the same genes that contribute to differences life history variation in the wild. Finally, evaluating natural genetic variation can elucidate subtleties in genetic functional constraint, characterize pathway dynamics, and identify regions of genic function that other methods may overlook, providing a complementary approach to the study of genetic determination of phenotypes.

My thesis. In the following chapters, I present the results of my evaluation of natural genetic variation at three longevity genes. The starting place for the data that follow is the natural habitats of wild *D. melanogaster* populations. The flies used in my experiments were collected, by myself and others, primarily from farms and orchards (Figure 1) where *D. melanogaster* populations persist over changes in seasonal climate and variations in the type and abundance of food resources. Observing *D. melanogaster* or even identifying their

FIGURE 1. Many of the field samples used in the work presented here were collected from Terhune Orchards, a family-run farm in Lawrenceville, NJ. Peaches are an especially reliable crop for collecting flies. Photos by A.B. Paaby.

refugia is nearly impossible in the temperate winter and spring, but in the summer and fall, when temperatures are warm and ripening fruit is abundant, we have observed explosions in *D. melanogaster* numbers (Figure 2). These fluctuations in environmental conditions, which have such strenuous effects on life histories and population dynamics, provide the foundation upon which this work is predicated. We assume that variation in the the climatic environment across geography imposes shifts in natural selection pressures, which are in turn reflected in the nucleotide variation of populations along these gradients. The grand naturalistic experiment provided by the distribution of *D. melanogaster* populations across latitude enables us to bring samples back into the laboratory and employ molecular and computational techniques to uncover the history of selection and identify functional genic elements responsible for phenotypic adaptations.

The following five chapters are comprised of four experimental summaries and one review, each written as stand-alone manuscripts. Chapter Two describes the first evidence that a well-characterized aging gene, *mth*, contributes to natural genetic variance for lifespan

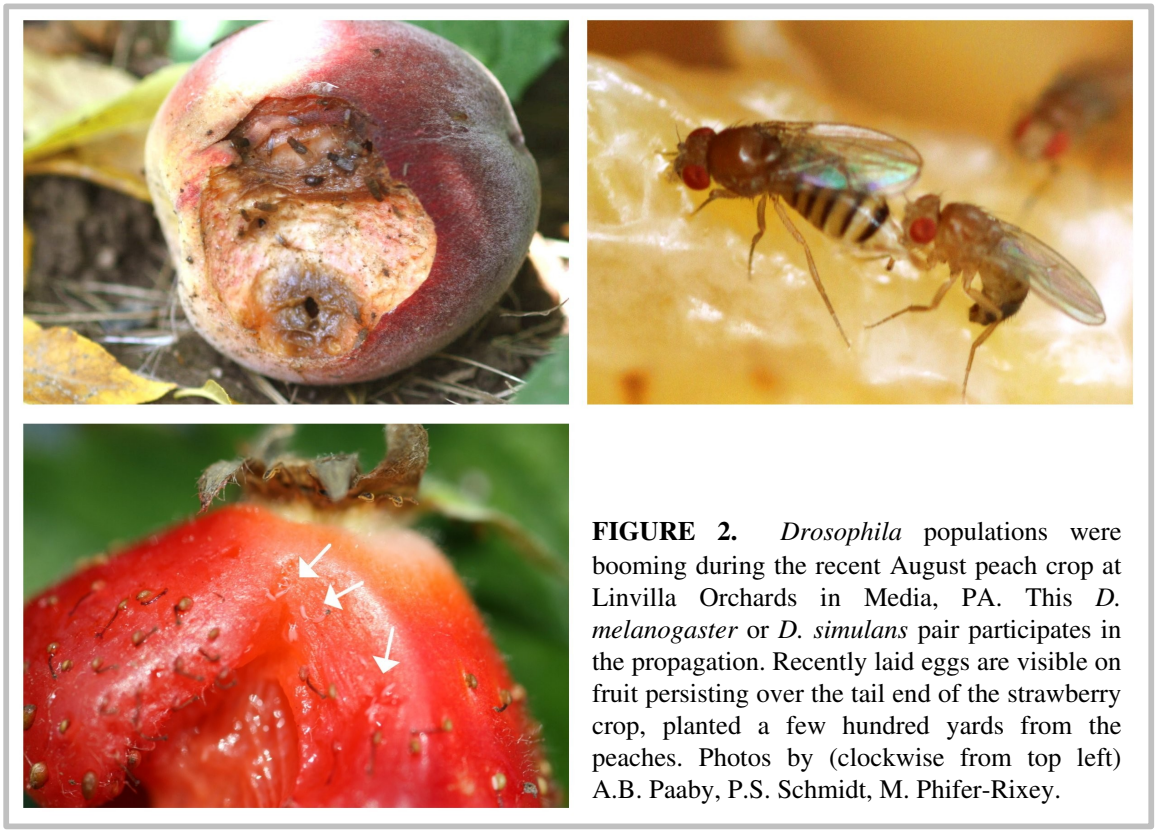


FIGURE 2. *Drosophila* populations were booming during the recent August peach crop at Linvilla Orchards in Media, PA. This *D. melanogaster* or *D. simulans* pair participates in the propagation. Recently laid eggs are visible on fruit persisting over the tail end of the strawberry crop, planted a few hundred yards from the peaches. Photos by (clockwise from top left) A.B. Paaby, P.S. Schmidt, M. Phifer-Rixey.

in the wild. It demonstrates that identification of longevity genes by classical genetics techniques, such as mutation screens, can be useful in the search for the functional genes and polymorphisms that underlie life history evolution in natural populations. Chapter Three follows up on this work by exploring genetic variation in noncoding regions of *mth*, in an effort to identify the functionally significant elements that produce phenotypic variation among different *mth* alleles; the evidence so far suggests that a functional polymorphism resides in the 5' promoter and mediates phenotype determination through differential gene expression. Chapter Four presents evidence that a polymorphism at at different aging gene, *InR*, likely contributes to genetic variance for *D. melanogaster* life history via changes in the amino acid sequence. I show that alternate alleles of an indel polymorphism in the first exon

exhibit reciprocal clines along the latitudinal gradient of the U.S. east coast, and with contributions from collaborators in Australia, that these clines are perfectly replicated on that continent as well. These striking patterns are strongly suggestive of contemporaneous responses to environmentally-imposed natural selection, probably for traits like increased stress tolerance in high latitude populations, where seasonal fluctuations are severe, and early reproductive success in low latitude populations, where intraspecific competition may be important. I also show that despite similar functional characterization by mutation genetics, another aging gene, the *InR* substrate, *chico*, shows neutral patterns of evolution; an example of how natural genetic variation can elucidate biologically relevant constraints that laboratory-based methods may miss. The functional significance of the *InR* indel polymorphism is described in Chapter Five, where I report that the *InR* allele common at high latitudes is associated with better stress tolerance, the allele at low latitudes is associated with faster development time and higher fecundity, and that these responses appear to be mediated by different levels of insulin signaling. These phenotypic responses accord with observed trait variation among natural populations, which suggests that this polymorphism contributes to *D. melanogaster* life history evolution. Finally, Chapter Six reviews the genetics of longevity in *D. melanogaster*, including the evidence for the hypothetical selection regime that may be driving the evolution of life history in the species.

CHAPTER TWO

FUNCTIONAL SIGNIFICANCE OF ALLELIC VARIATION AT *METHUSELAH*, AN AGING GENE IN *DROSOPHILA*[†]

ABSTRACT

Background. Longevity and age-specific patterns of mortality are complex traits that vary within and among taxa. Multiple candidate genes for aging have been identified in model systems by extended longevity mutant phenotypes, including the G-protein coupled receptor *methuselah* (*mth*) in *D. melanogaster*. These genes offer important insights into the mechanisms of lifespan determination and have been major targets of interest in the biology of aging. However, it is largely unknown whether these genes contribute to genetic variance for lifespan in natural populations, and consequently contribute to lifespan evolution.

Methodology/Principle Findings. For a gene to contribute to genetic variance for a particular trait, it must meet two criteria: natural allelic variation and functional differences among variants. Previous work showed that *mth* varies significantly among wild populations; here we assess the functional significance of wild-derived *mth* alleles on lifespan, fecundity and stress resistance using a quantitative complementation scheme. Our results demonstrate that *mth* alleles segregating in nature have a functional effect on all three

† This chapter has been published as a journal article: Paaby AB, Schmidt PS. 2008. Functional significance of allelic variation at *methuselah*, an aging gene in *Drosophila*. *PLoS ONE* 3(4): e1987.

traits.

Conclusions/Significance. These results suggest that allelic variation at *mth* contributes to observed differences in lifespan and correlated phenotypes in natural populations, and that evaluation of genetic diversity at candidate genes for aging can be a fruitful approach to identifying loci contributing to lifespan evolution.

INTRODUCTION

Lifespan and age specific mortality rates are primary life history components and vary significantly among natural populations (Promislow et al. 1996). QTL and artificial selection experiments have demonstrated a highly complex genetic architecture, with many genetic correlations among lifespan and associated life history traits (e.g. Rose & Charlesworth 1981; Mackay 2002). Nonetheless, single gene manipulations have identified candidate genes for aging by extended longevity phenotypes, and in the model system *Drosophila melanogaster* such genes include the *Insulin-like Receptor* (Tatar et al. 2001), *chico* (Clancy et al. 2001), *dFOXO* (Hwangbo et al. 2004), *Indy* (Rogina et al. 2000) and *methuselah* (Lin et al. 1998).

Williams' (1957) theory of antagonistic pleiotropy describes how pleiotropic alleles that increase fitness early in life may experience positive selection even though they incur a fitness cost later in life. Identified aging genes have consistently shown costs to lifespan extension, particularly in reproduction. Although there is some evidence that lifespan and reproductive success can be decoupled (Partridge et al. 2005), multiple analyses have

revealed previously undetected tradeoffs under specific conditions (e.g. Marden et al. 2003; Jenkins et al. 2004; Mockett & Sohal 2006). In addition to demonstrating negative effects on reproduction, longevity mutations are positively correlated with stress resistance (Vermeulen & Loeschcke 2006). Such correlations may explain aspects of lifespan evolution, and why loss-of-function mutants can result in lifespan extension. However, it remains unclear whether identified aging genes are major contributing factors to the genetic variance for longevity that is routinely observed in populations (e.g. Carbone et al. 2006). The role of aging genes, such as the components of the insulin (IGF-1/IIS) pathway, are highly conserved in metazoans (Tatar et al. 2003). Such genes may experience strong selection constraints and consequently exhibit little variation at the nucleotide level. Furthermore, a gene may be polymorphic, but the variation may be functionally neutral. In order to comprehensively examine the contribution of candidate genes for aging to the genetic variance for longevity and correlated traits, the functional significance of allelic variation must be assessed.

The G-protein coupled receptor *methuselah* (*mth*) is a promising candidate for such analyses. This gene, discovered by its extended longevity mutant phenotype, was the first candidate gene for aging identified in *D. melanogaster*. Individuals homozygous for a P-element disruption at *mth* lived an average of 35% longer than the parental strain and showed significant resistance to oxidative stress, starvation and heat stress (Lin et al. 1998). Mutants also show a tradeoff in lifetime reproductive success under heat stress conditions (Mockett & Sohal 2006). The gene encodes a G-protein coupled receptor, showing seven hydrophobic regions suggestive of transmembrane domains (Lin et al. 1998) and an ectodomain

containing a ligand binding site (West et al. 2001). Disruption of *mth* ligand activity also promotes lifespan: mutation at *stunted*, a gene that produces two *mth* peptide ligands, and constitutive expression of antagonist peptide ligands both produce extensions in longevity (Cvejic et al. 2004; Ja et al. 2007). Mutants show a reduction in excitatory neurosecretion, and *mth* appears to modulate synaptic strength in neurons by regulating vesicle trafficking (Song et al. 2002); this may be important in sensorimotor ability, as *mth* mutants also show enhanced visuomotor synchronization and phototaxis (Petrosyan et al. 2007). Despite these implications for neuroendocrine signaling, the mechanism of lifespan regulation by *mth* is not well understood. The normal reduction in germline stem cell division is not exhibited in aging *mth* mutants (Wallenfang et al. 2006), revealing a potentially different process by which *mth* affects physiology. It may be that *mth* is pleiotropic both in the traits it determines and the mechanisms by which determination happens.

In addition to results demonstrating its function as a pleiotropic aging gene, data from the wild suggest that *mth* may be an important component in lifespan evolution. In a comparison of expression levels for *D. melanogaster*, *D. simulans* and their F1 hybrids for 31 genes, *mth* was shown to have one of the strongest patterns for compensatory *cis-trans* regulatory evolution (Landry et al. 2005), suggesting that *mth* has experienced strong selection on expression level since *D. melanogaster* and *D. simulans* shared a common ancestor approximately two million years ago. Over this same timescale, *mth* also shows a very high rate of protein evolution (Schmidt et al. 2000). Furthermore, wild populations of *D. melanogaster* show a cline in the frequency of the most common *mth* haplotype along the latitudinal gradient of the U.S. east coast (Schmidt et al. 2000), a pattern that decays with

decreasing linkage disequilibrium both 5' and 3' of the *mth* locus (Duvernell et al. 2003). This pattern covaries with clines in longevity, fecundity, stress resistance and other life history traits in these populations (Schmidt et al. 2005). Genetic variance for and genetic correlations among these traits underlie predictable life history variation that reflects distinct selection pressures that vary spatially and temporally (Schmidt et al. 2005; Schmidt & Conde 2006; Schmidt & Paaby 2008). Together, these results imply that *mth* has experienced directional selection pressures over short and long timescales, and that *mth* may be an important target in the selection regime driving the observed patterns of life history variation in natural populations. However, it is unknown whether the observed allelic variation at *mth* is of functional significance, and contributes to the genetic variance for lifespan in natural populations.

Here we present results from a modified quantitative complementation scheme (Pasyukova et al. 2000), in which we tested whether a set of wild-derived *mth* alleles show differences in lifespan, fecundity and resistance to oxidative stress.

MATERIALS AND METHODS

Quantitative complementation scheme. We used a modified quantitative complementation scheme (Pasyukova et al. 2000) to test the contribution to phenotype of eight wild *mth* alleles still embedded in their natural genomes. Each wild line was crossed to one of two mutant *mth* alleles, and the phenotype of this genetic construct was compared to the phenotype of the same wild line over a functional, wild-type *mth* allele. The magnitude

of difference among line pairs was then used to test for differences among lines. The use of two mutant *mth* alleles generated two independent complementation tests for each line in each of three trait assays.

Each of the eight wild-derived lines was crossed to the three lab-derived *mth* alleles, mth^{1}, mth^{R3} and $mth^{\Delta 6}$. All parental lines were maintained at low density on standard cornmeal-molasses medium in bottle cultures to limit confounding environmental variation. In the crosses with $mth^{\Delta 6}$, which is homozygous lethal, F1 pupae not exhibiting the balancer phenotype were selected prior to eclosion. The F1 progeny used for the assays were of the following genotypes: $+_i/mth^{\Delta 6}$, $+_i/mth^{1}$ and $+_i/mth^{R3}$, where $+_i$ is one of the eight wild-derived lines. These three genotypic categories allowed us to compare the phenotype of each $+_i$ line over a *mth* mutant (either $mth^{\Delta 6}$ or mth^{1}) to the phenotype of that same line over a functional copy of *mth* (mth^{R3}). In the statistical analyses, there were two criteria for establishing differences among the wild *mth* alleles. First, significant interaction terms between the wild lines and the lab-derived *mth* alleles were used to indicate a failure to complement, or significant functional variation in the wild lines. Second, an F-statistic was used to ensure that the variance in the wild-type background was not greater than variance in the mutant backgrounds; greater variance in the wild-type background would suggest that epistatic interaction between the wild lines and the wild-type mth^{R3} allele, rather than allelic variation, produced the failure to complement (Geiger-Thornsberry & Mackay 2004).

Flies. Approximately 40 isofemale lines were established from wild populations in Maine, New Jersey, Pennsylvania and Florida (described in Schmidt & Paaby 2008). Flies were made isogenic at the third chromosome using balancers and a ~1 kb region of *mth* was

sequenced to characterize each line by haplotype as determined by SNP identity (Schmidt et al. 2000; Duvernell et al. 2003). The two remaining chromosomes were not background-replaced or otherwise made isogenic within or between lines. We confirmed that the frequency of the common *mth* haplotype in our three geographic regions was consistent with the cline in frequency demonstrated previously (Schmidt et al. 2000). Eight lines (referenced here as *BF51*, *BF54*, *RR5*, *S108*, *S97*, *SL5135*, *T28* and *T50*) were selected for the functional tests, based on this haplotype identity as well as polymorphism at the remaining sites within the sequenced ~1 kb region. All the polymorphisms we observed in this region were consistent with the polymorphisms reported earlier (see Schmidt et al. 2000), with the exception of four additional singleton SNPs. Consequently the eight wild lines selected were unique according to their sequence identity within this region.

The lab-derived *mth* alleles include three *mth* genotypes in a standardized background (Lin et al. 1998; Song et al. 2002): a P-element insertion hypomorph (mth^{1}), a *mth* null ($mth^{\Delta 6}$), and a wild-type *mth* allele (mth^{R3}). The mth^{1} hypomorph was generated by insertion of the standard P{lacW} element in the *w1118* background (Lin et al. 1998). Both the null and the wild-type alleles were created by excision of the P-element in the original mth^{1} line. The $mth^{\Delta 6}$ null allele resulted from an imprecise excision that removed the C-terminal end of exon 3, the N-terminal end of exon 4, and the intron between exons 3 and 4 (Song et al. 2002). The mth^{R3} wild-type revertant allele resulted from a precise excision of the P-element, which restored wild-type function (Lin et al. 1998; Song et al. 2002). The $mth^{\Delta 6}$ allele is homozygous lethal, and was maintained over a TM6 balancer.

Functional assays. For the lifespan and fecundity assays, flies were collected,

freshly eclosed, over three days. Flies experienced a short, single exposure to CO_2 during virgin/male sorting. For each of the 24 genotypic combinations, demography cages were initiated by putting 40 males and 40 females of a single 24-hour cohort into perforated 6-oz polypropylene bottles. Each combination of wild line by *mth* allele was replicated three times. Cornmeal-molasses food plates were changed every day for the first 16 days and every other day afterwards. Eggs laid and dead flies were scored at every plate change. Cages were kept at room temperature. For the oxidative stress assay, flies were collected, freshly eclosed, over an 8-hour window, using a short, single exposure to CO_2. For each genotypic combination, five males and five females were put in vials with standard cornmeal-molasses media and aged for four days. Each vial was replicated four times. Flies were then transferred into media-free vials with cotton saturated with 5 mL of 30 mM methyl viologen (paraquat) in 5% sucrose solution. After 19 hours of continuous exposure, patterns of mortality were determined for all replicates.

Statistics. Our analyses tested the effects of line, allele, sex and the interactions among these effects. Here, as elsewhere, "line" refers to the eight wild-derived variants and "allele" refers to the three lab-derived *mth* alleles. Lifespan data for each sex were also analyzed separately and there was no sex term in the fecundity analysis. For the lifespan data, a proportional hazards model was used to test for main and interaction effects and to estimate risk ratios. Total fecundity was analyzed with ANOVA, treating the line and line by allele interaction as random effects. The oxidative stress data were analyzed with nominal logistic regression, modeling the log odds (mortality/survivorship). Because the wild lines varied at multiple loci, epistasis between these other genes derived from natural populations

and the lab-derived wild-type *mth* allele may have contributed confounding effects. To test whether epistatic interactions between genes on the wild-derived chromosomes and the mth^{R3} background were responsible for the failures to complement, we used an F-statistic to examine the equivalency of variances in both the *mth* wild-type and mutant backgrounds (Geiger-Thorsberry & Mackay 2004). Variance components were computed as sums of squares for each genotype (mth^{R3}, $mth^{\Delta 6}$ and mth^{1}) for each of the functional assays. For each comparison, F = Var($mth^{R3}/+_i$)/Var($mth^{\Delta 6}/+_i$ or $mth^{1}/+_i$). F values lower than the critical value for a type I error of 0.05 allowed us to accept the null hypothesis, that variance in the mth^{R3} background was not significantly greater than variance in the mutant background and consequently not responsible for observed differences among lines. All statistical procedures were performed using JMPv5 (SAS Institute, Cary, NC).

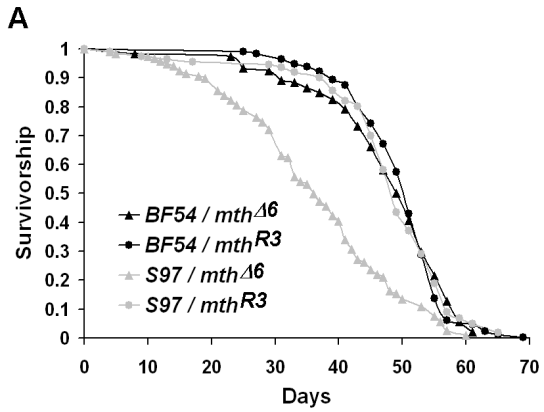
RESULTS

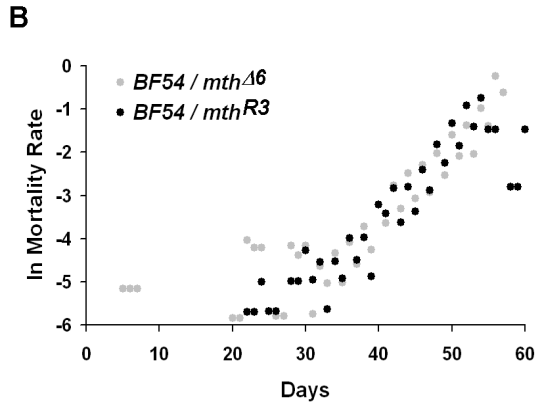
The wild *mth* lines exhibited functional differences in lifespan, fecundity and resistance to oxidative stress. These differences were demonstrated in both of our independent complementation schemes, which each used a different mutant *mth* allele to measure the wild *mth* variants' contribution to phenotype. Differences in function among wild *mth* lines were seen in the magnitudes of difference between lines over the mutant *mth* alleles ($mth^{\Delta 6}$, the deletion, or mth^{1}, the P-element hypomorph) and these same lines over a functional *mth* allele (mth^{R3}, the wild-type revertant), or by a significant line by allele interaction term in the statistical analysis. Qualitative comparisons between the two

independent tests showed similarities in which lines fail to complement, the directionality of complementation failure, and the significance of effect and interaction terms. For each functional assay, the variance in the mth^{R3} wild-type background was not significantly greater than the variance in either the $mth^{\Delta 6}$ or mth^{1} mutant backgrounds, which supports allelism over epistasis as the genetic mechanism generating the differences in phenotype among lines.

Lifespan. The eight wild *mth* lines showed phenotypic variation in lifespan in both the $mth^{\Delta 6}$ and the mth^{1} complementation tests (Table 1). Additionally, females and males exhibited sex-specific differences in *mth* allelic contribution to lifespan. As expected, the three lab-derived *mth* alleles, $mth^{\Delta 6}$, mth^{1} and mth^{R3} showed variation in longevity as well. Since reduced *mth* expression is associated with lifespan extension (Lin et al. 1998), we predicted that the $mth^{\Delta 6}$ and mth^{1} mutants would exhibit increased longevity relative to the mth^{R3} wild-type allele. However, this was not observed: some lines exhibited greater longevity when tested over mth^{R3} (Table S1). The significant line by allele interaction terms, which indicate that wild *mth* variants produce functional differences in longevity, were consistent across the two complementation tests (Table 1). This suggests that the observed results were not due to idiosyncratic effects of the various *mth* mutants or interactions with the wild *mth* lines. Furthermore, the phenotypic differences among wild *mth* lines did not result from epistatic effects between the functional mth^{R3} allele and the wild-derived lines, because the variation in the wild-type background was not significantly greater than the variation in the mutant backgrounds (Table 1).

Variation in longevity among the wild lines can be seen in a sample comparison of paired survivorship curves and mortality rates (Figure 1). For example, the difference in





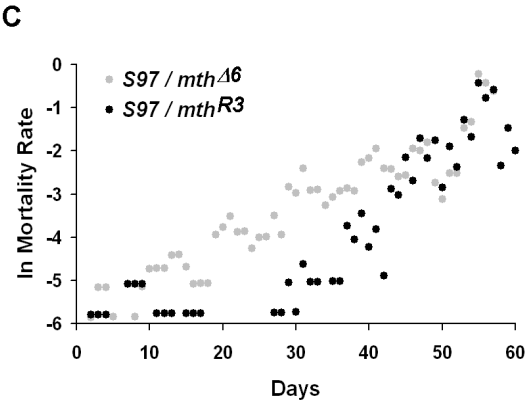


FIGURE 1. A sample comparison of paired survivorship curves and paired mortality rates for two wild lines. Males from lines *BF54* and *S97* show differences in survivorship, as *BF54* over $mth^{\Delta 6}$ survived the same as *BF54* over mth^{R3}, while *S97* over $mth^{\Delta 6}$ survived more poorly than *S97* over mth^{R3} (A). These lines also exhibit differences in mortality rate: *BF54* shows the same mortality rate when paired with $mth^{\Delta 6}$ as when paired with mth^{R3} (B), while *S97* shows a higher mortality rate when paired with $mth^{\Delta 6}$ than when paired with mth^{R3} (C).

survivorship between line *S97* over $mth^{\Delta 6}$ and over mth^{R3} is greater than the difference in survivorship between line *BF54* over $mth^{\Delta 6}$ and over mth^{R3} (Figure 1A). In this case, the single copy of *mth* in line *S97* failed to complement, shown by the difference in longevity between this line over the $mth^{\Delta 6}$ deletion and the mth^{R3} wild-type allele; line *BF54* complemented, with no difference in longevity when paired with $mth^{\Delta 6}$ than when paired with mth^{R3}. The *S97* and *BF54* lines also aged differently. Line *S97* showed no difference in mortality rate when over the $mth^{\Delta 6}$ allele than when over mth^{R3} (Figure 1B), but line *BF54* showed a higher mortality rate when paired with $mth^{\Delta 6}$ than when paired with mth^{R3} (Figure 1C). Risk ratios generated from the proportional hazards analysis illustrate the heterogeneity

in longevity among all the natural *mth* lines (Figure 2). The risk ratios show the risk of death for each wild line over a *mth* mutant ($mth^{\Delta 6}$, Figure 2A and 2C, or mth^{1}, Figure 2B and 2D) relative to that same line over the functional *mth* allele (mth^{R3}). Some lines survived longer in the wild-type background, while others survived longer in the mutant background; this is shown by positive ratios for some lines and negative ratios for others (e.g. lines *BF51* and *S108*, Figure 2A and 2B). Females (Figure 2A and 2B) and males (Figure 2C and 2D) are

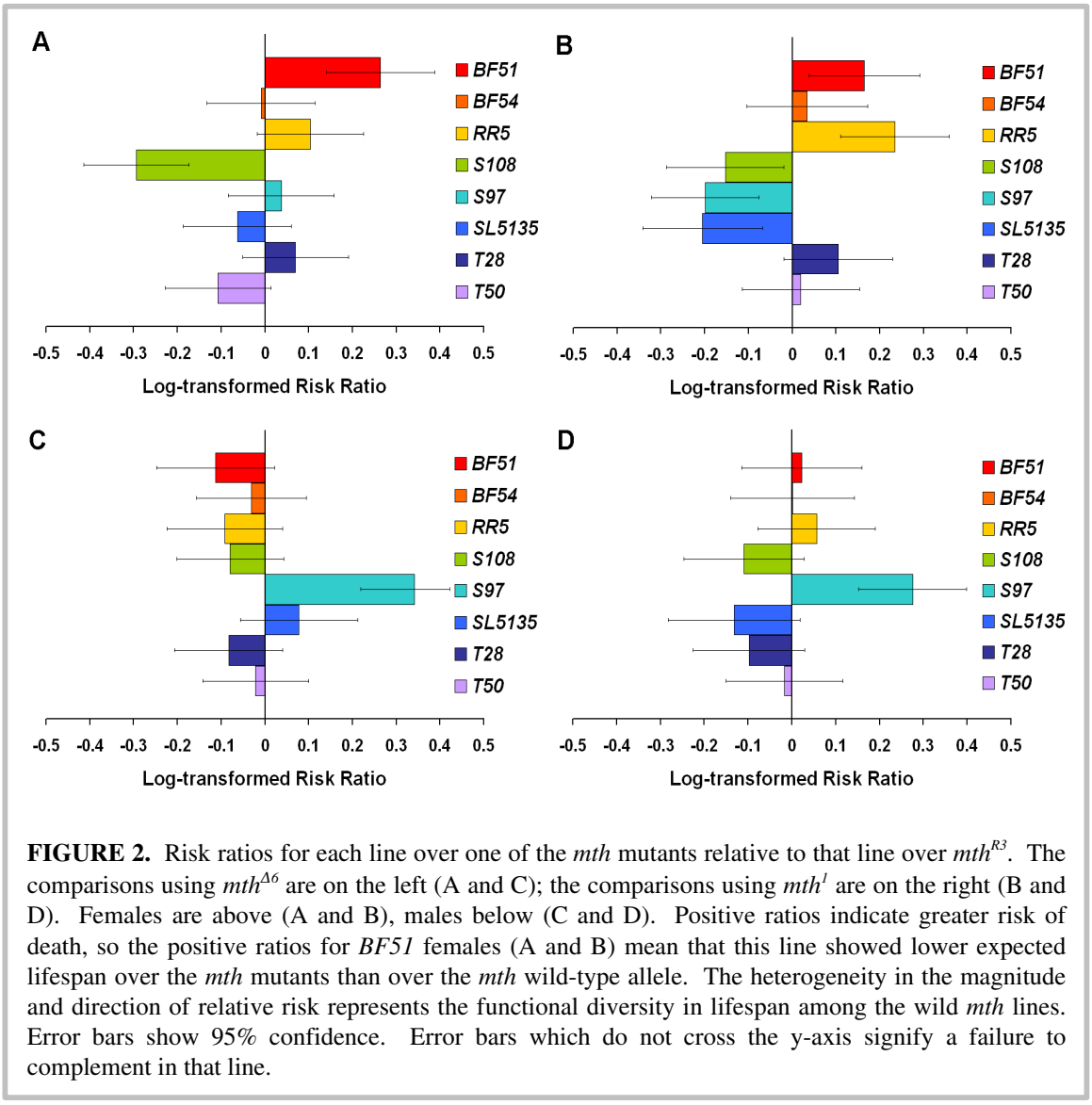


FIGURE 2. Risk ratios for each line over one of the *mth* mutants relative to that line over mth^{R3}. The comparisons using $mth^{\Delta 6}$ are on the left (A and C); the comparisons using mth^{1} are on the right (B and D). Females are above (A and B), males below (C and D). Positive ratios indicate greater risk of death, so the positive ratios for *BF51* females (A and B) mean that this line showed lower expected lifespan over the *mth* mutants than over the *mth* wild-type allele. The heterogeneity in the magnitude and direction of relative risk represents the functional diversity in lifespan among the wild *mth* lines. Error bars show 95% confidence. Error bars which do not cross the y-axis signify a failure to complement in that line.

shown separately, as females and males were affected differently by allelic contribution to longevity (Table 1).

Fecundity. The wild *mth* lines showed functional differences in total lifetime fecundity for both complementation tests (Table 2), and an analysis of fecundity per female per day yielded quantitatively identical results (data not shown). The fecundity assays also demonstrated, predictably, an effect by *mth* allele (although the effect was not statistically significant in the in the test using mth^{1}); unpredictably, all lines but one showed lowest total fecundity when paired with the functional *mth* allele (Table 2, Figure 3, Table S1). Comparison of variation among *mth* allele backgrounds supports allelism, as variation in the wild-type background was not greater than variation in the mutant backgrounds (Table 2).

Stress resistance. The wild *mth* lines also demonstrated differences in response to oxidative stress in both complementation tests (Table 3). Odds ratios illustrate the variation in resistance to oxidative stress among lines (Figure 4). Analogous to the risk ratios in the lifespan analyses, these ratios show the odds of death for each line over one of the *mth* mutants ($mth^{\Delta 6}$, Figure 4A; mth^{1}, Figure 4B) relative to that same line over the functional *mth*

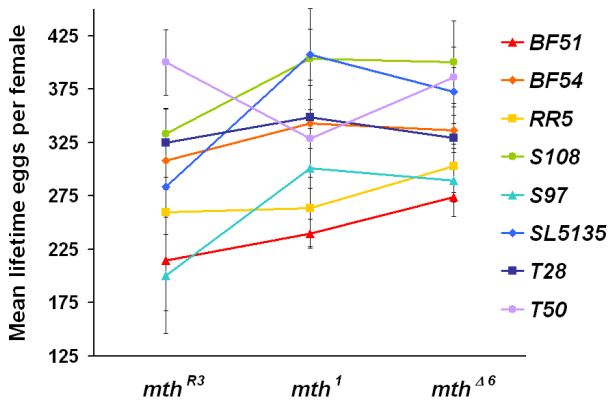


FIGURE 3. Average lifetime fecundity per line for all genotypic combinations. Non-parallel slopes between lines in the three lab-derived *mth* backgrounds indicate the functional variation in lifetime fecundity among the wild lines. The general increase in eggs laid across the x-axis represents a correlation of increasing fecundity with decreasing background *mth* expression; all lines but *T50* showed lowest fecundity when paired with the wild-type mth^{R3} allele. Error bars show 95% confidence.

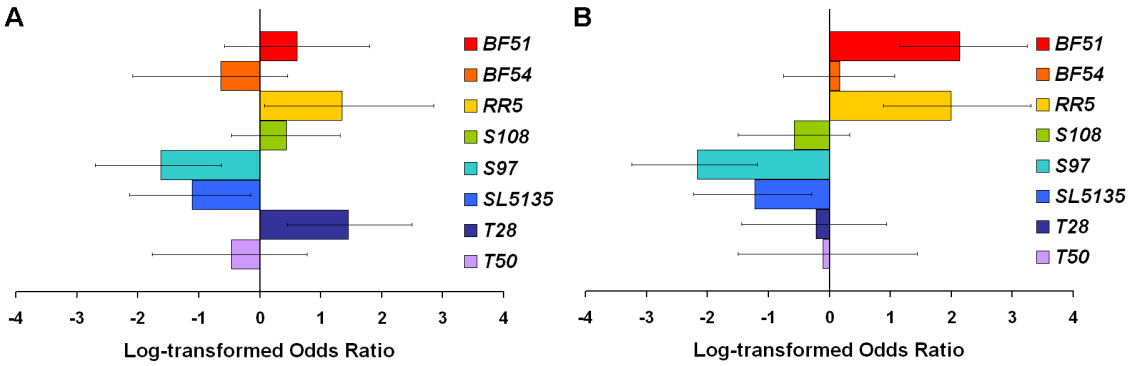


FIGURE 4. Odds ratios for each line over one of the *mth* mutants ($mth^{\Delta 6}$, A; mth^{1}, B) relative to the line over mth^{R3}. Positive ratios indicate greater risk of death. The heterogeneity in the magnitude and direction of relative odds represents the functional diversity in stress resistance among the wild *mth* lines. Error bars show 95% confidence. Error bars which do not cross the y-axis signify a failure to complement in that line.

allele (mth^{R3}). Data were pooled across sexes, as we observed no effect of sex on allele function in this assay (Table 3). Like the lifespan risk ratios, the oxidative stress odds ratios demonstrate extensive heterogeneity among the wild lines: some lines showed better performance in the wild-type background (a positive odds ratio), whereas some lines showed better performance in the mutant background (a negative odds ratio). For example, line *RR5* was more resistant to oxidative stress when paired with the functional mth^{R3} allele than with either of the mutant alleles, $mth^{\Delta 6}$ or mth^{1}; however, lines *S97* and *SL5135* were both more resistant to oxidative stress when paired with $mth^{\Delta 6}$ and mth^{1} than with mth^{R3} (Figure 4A and 4B). Comparison across the lifespan and oxidative stress assays suggests a positive correlation between these traits among females. Four lines show significant differences in both assays: *BF51* and *RR5* show increased stress resistance and longer lifespan in the wild-type background; *S97* and *SL5135* show increased stress resistance and longer lifespan in the

mutant backgrounds; no line showed increased stress resistance but shorter lifespan (or vice versa) in the same background. Like the lifespan and fecundity assays, the differences among lines in resistance to oxidative stress suggests allelism, as variation in the wild-type background was not greater than variation in the mutant backgrounds (Table 3).

TABLE 1. Statistical results for the lifespan assay.

Proportional hazards model effect likelihood ratio tests

test using $mth^{\Delta 6}$ and mth^{R3}				test using mth^{1} and mth^{R3}			
Source	DF	χ^2	p	Source	DF	χ^2	p
line	1	213.966	<0.0001	line	1	143.172	<0.0001
allele	7	73.708	<0.0001	allele	7	69.211	<0.0001
sex	1	6.609	0.0101	sex	1	10.956	0.0009
line x allele	7	36.517	<0.0001	line x allele	7	27.352	0.0003
line x sex	1	157.598	<0.0001	line x sex	1	211.332	<0.0001
allele x sex	7	8.604	0.0034	allele x sex	7	7.601	0.0058
line x allele x sex	7	35.610	<0.0001	line x allele x sex	7	34.726	<0.0001

F-statistic testing variances for age at death among alleles

test using $mth^{\Delta 6}$ and mth^{R3}				test using mth^{1} and mth^{R3}			
Ratio	DF	F	critical F	Ratio	DF	F	critical F
$\sigma^{2}_{mthR3/+i}/\sigma^{2}_{mth\Delta 6/+i}$	7, 7	1.2511 (f) 0.1375 (m)	3.787	$\sigma^{2}_{mthR3/+i}/\sigma^{2}_{mth1/+i}$	7, 7	1.5810 (f) 0.1887 (m)	3.787

TABLE 2. Statistical results for the fecundity assay.

ANOVA of lifetime fecundity

test using $mth^{\Delta 6}$ and mth^{R3}					test using mth^{1} and mth^{R3}				
Source	DF	SS	F	p	Source	DF	SS	F	p
line	7	1.935e8	8.100	0.0066	line	7	1.874e8	3.586	0.0569
allele	1	4.051e7	11.874	0.0108	allele	1	2.601e7	3.515	0.1028
line x allele	7	2.388e7	2.900	0.0182	line x allele	7	5.227e7	6.197	0.0002

F-statistic testing variances for total eggs laid among alleles

test using $mth^{\Delta 6}$ and mth^{R3}				test using mth^{1} and mth^{R3}			
Ratio	DF	F	critical F	Ratio	DF	F	critical F
$\sigma^{2}_{mthR3/+i}/\sigma^{2}_{mth\Delta 6/+i}$	7, 7	2.0146	3.787	$\sigma^{2}_{mthR3/+i}/\sigma^{2}_{mth1/+i}$	7, 7	1.4202	3.787

TABLE 3. Statistical results for the oxidative stress resistance assay.

Nominal logistic model effect Wald tests

test using $mth^{\Delta 6}$ and mth^{R3}				test using mth^{1} and mth^{R3}			
Source	DF	χ^2	p	Source	DF	χ^2	p
line	1	95.899	<0.0001	line	1	87.944	<0.0001
allele	7	11.812	0.0006	allele	7	0.828	0.3629
sex	1	0.015	0.9039	sex	1	0.899	0.3429
line x allele	7	26.799	0.0004	line x allele	7	46.216	<0.0001
line x sex	1	18.390	0.0103	line x sex	1	8.980	0.2541
allele x sex	7	1.189	0.2755	allele x sex	7	3.761	0.0525
line x allele x sex	7	7.695	0.3603	line x allele x sex	7	10.096	0.1832

F-statistic testing variances for flies dead at hour 19 among alleles

test using $mth^{\Delta 6}$ and mth^{R3}				test using mth^{1} and mth^{R3}			
Ratio	DF	F	critical F	Ratio	DF	F	critical F
$\sigma^{2}_{mthR3/+i}/\sigma^{2}_{mth\Delta 6/+i}$	7, 7	3.711	3.787	$\sigma^{2}_{mthR3/+i}/\sigma^{2}_{mth1/+i}$	7, 7	2.148	3.787

DISCUSSION

Our results demonstrate that natural variants of *mth*, a gene for aging in *Drosophila*, contribute to significant functional variation in lifespan, fecundity and resistance to oxidative damage. In the lifespan analyses, the significant line by allele by sex interaction term also demonstrates that variation at *mth* affects males and females differently. This observation is consistent with the sex-specific effects of identified QTLs for lifespan (Nuzhdin et al. 1997) and complementation data for other candidate genes for aging (Pasyukova et al. 2004). Furthermore, the observed data do not result from idiosyncrasies associated with a particular *mth* mutant allele, as the patterns were qualitatively identical between the two complementation tests. Nor does the failure to complement appear to be determined by epistatic interaction between the functional lab-derived wild-type allele and the wild-derived lines, as the F ratio tests were nonsignificant for all comparisons. Other studies have demonstrated the significance of genetic background and epistatic interactions on patterns of longevity (Spencer et al. 2003; Pasyukova et al. 2004; Spencer et al. 2005). In this study, the standardization of the genetic background reduces the likelihood of identifying interesting epistatic interactions among genes affecting longevity, but may increase the power to detect small differences in function among naturally-occurring wild-type variants.

While our data support the hypothesis that allelic variation at *mth* contributes to the genetic variance for longevity in natural populations, the nature of the complementation scheme precludes identifying which specific *mth* variants are associated with relative lifespan extension. This results from the fact that the assayed *mth* variants are embedded in their own

genetic backgrounds, and the tests of functional significance evaluate the heterogeneity among line by allele genotypic crosses. As such, these analyses do not test the hypothesis that wild *mth* alleles show a trend in *mth* function by geography, but that diverse *mth* alleles (derived from geographically diverse populations) are functionally distinct. However, other trends do emerge from the data. Interestingly, the *mth* mutants did not show consistently longer lifespan or lower fecundity than the wild-type genotypes, as the initial characterization of *mth* as a longevity gene may have predicted. (The $mth^{\Delta 6}$ allele did show greater resistance to oxidative stress than the wild-type mth^{R3} allele and the comparison using the mth^{1} allele was nonsignificant.) The original functional assays for *mth* showed a lifespan extension in flies homozygous for mth^{1} and an increased resistance to oxidative stress in flies heterozygous for mth^{1} and $mth^{\Delta 6}$ (Lin et al. 1998); a tradeoff with fecundity was later demonstrated at higher temperatures (Mockett et al. 2006). Consequently our results are consistent with the earlier observation that lifespan extension is not achieved by mutation at a single copy of *mth*, although it challenges the hypothesis that there is a straightforward relationship between *mth* expression and lifespan. For example, it is possible that the mutant mth^{1} and $mth^{\Delta 6}$ alleles showed shorter lifespan than the wild-type mth^{R3} allele (Table S1) because major reduction in *mth* expression extends lifespan, while minor reduction compromises overall fitness due to the pleiotropic nature of the gene. The fact that the $mth^{\Delta 6}$ allele showed predictably higher oxidative stress tolerance (Table S1) but unpredictably shorter lifespan also suggests that the highly quantitative determination of lifespan can complicate attempts to interpret mechanisms of genetic control.

Our results are consistent with data showing that lifespan extension by *mth* is limited

when males and females are permitted to mate (Baldal et al. 2006), since all our assays were conducted with mixed-sex replicates. Our results also support the conclusion that lifespan extension and life history phenotypes associated with longevity genes are highly dependent upon genetic and environmental context. The overexpression of *superoxide dismutase* (*SOD*) was originally shown to increase longevity and stress resistance (Parkes et al. 1998), but these phenotypes were later demonstrated to be genotype- and sex-specific when *SOD* was overexpressed in naturally long-lived genetic backgrounds (Spencer et al. 2003). The lifespan extension of *mth* mutants is also dependent upon genetic background and sex, and the reduced fecundity of *mth* mutants was only revealed under exposure to environmental stress (Mockett & Sohal 2006). These results are consistent with our data, which show a striking effect by sex in the lifespan assay.

The mechanisms by which variation at *mth* affects performance are unknown, but we hypothesize that differences in gene expression may drive the observed functional variation among *mth* alleles. Haplotypes at the *mth* locus demonstrate a significant latitudinal cline in frequency that mirrors differences among populations in expected lifespan, but none of the individual polymorphic sites at *mth* exhibit a clinal pattern (Schmidt et al. 2000). This cline decays in both directions away from the *mth* locus, indicating that the actual site(s) under selection reside not in the coding region, but possibly in promoter or regulatory regions (Duvernell et al. 2003). Expanded sequencing of previously characterized *mth* variants has revealed high polymorphism in the 5' and 3' *mth* UTR, and the potential functional impact of these polymorphisms on expression level is currently being explored. Alternatively, functional differences among wild *mth* variants may be caused by distinct properties of the

protein if the mechanism is more complex than a single amino acid polymorphism. We are currently evaluating this possibility by examining patterns of linkage disequilibrium and geographical distribution in our expanded *mth* sequence dataset.

In addition to identifying the pathways and genes that regulate aging, determining the underlying genetic basis for differences in lifespan among individuals within populations is critical for understanding how a quantitative trait evolves. The significant and predictable variation in life histories among *Drosophila* populations provides an opportunity to dissect these differences in nature while affording all the advantages of a model organism. Longevity is highly variable within *D. melanogaster* (Promislow et al. 1996), and wild populations show genetically correlated differences in lifespan and other life history traits by environment (Trotta et al. 2006; Schmidt & Paaby 2008). This standing genetic variance has enabled mapping of lifespan QTL (Mackay 2002; Wilson et al. 2006) and led to precise identification of additional candidate genes for aging (DeLuca et al. 2003; Pasyukova et al. 2004). Several studies have found functional significance of allelic variation at lifespan QTL (Geiger-Thornsberry & Mackay 2004) and candidate genes (DeLuca et al. 2003; Carbone et al. 2006), confirming that these loci may contribute to the observed variation in longevity phenotypes. However, in order to comprehensively describe the contribution of any one gene to lifespan evolution, knowledge about functional allelic variation must be integrated with patterns environmental heterogeneity and possible selection pressures in the wild.

By examining the functional significance of allelic variation at a single locus within the context of environmental heterogeneity, our study provides a complementary approach to the evaluation of a quantitative trait by QTL analysis. The significance of functional

variation at *mth* is reinforced by the adaptive pattern this allelic variation exhibits among natural populations, and suggests that *mth* may be an important component in lifespan evolution. By testing natural variation at this gene, we have also demonstrated the utility of forward genetics in identifying loci that contribute to the evolution of a complex quantitative trait. Ultimately, differences in expected lifespan among individual genotypes will be resolved by the joint processes of gene identification, characterization of molecular mechanisms, and associations between specific variants and phenotypes.

ACKNOWLEDGEMENTS

We thank Ravi Ranjan for providing the *mth* allelic series and Jayatri Das, Thomas Flatt and two anonymous reviewers for their helpful comments.

SUPPLEMENT

TABLE S1. Phenotype means for the three lab-derived *mth* alleles.

Age death (days)				
	Mean	St dev	St err mean	N
mth^{R3} females	46.89	13.73	0.46	901
mth^{R3} males	48.22	10.98	0.39	810
mth^{1} females	50.25	14.52	0.53	751
mth^{1} males	49.09	11.67	0.44	714
$mth^{\Delta 6}$ females	43.32	12.80	0.42	928
$mth^{\Delta 6}$ males	44.49	12.95	0.43	889

Lifetime eggs laid per female				
	Mean	St dev	St err mean	N
mth^{R3}	290.5	67.6	13.8	24
mth^{1}	321.4	60.3	13.5	20
$mth^{\Delta 6}$	336.5	48.6	9.9	24

Proportion of flies dead at hour 19 after exposure to paraquat				
	Mean	St dev	St err mean	N
mth^{R3}	0.4844	0.3139	0.1110	8
mth^{1}	0.4344	0.2142	0.0757	8
$mth^{\Delta 6}$	0.3313	0.1630	0.0567	8

CHAPTER THREE

NUCLEOTIDE VARIATION IN THE *METHUSELAH* PROMOTER MAY INCLUDE FUNCTIONAL POLYMORPHISMS AFFECTING *DROSOPHILA* LIFE HISTORY

ABSTRACT

The gene *methuselah* (*mth*) confers a 35% increase in longevity when mutated, and natural genetic variation at this locus has functional effects on lifespan, fecundity and stress tolerance. In the wild, the most common haplotype in the *mth* coding region shows a cline in frequency across latitude that coincides with phenotypic changes in life history traits. No functional polymorphisms have been identified at *mth*, but we hypothesize that one or more sites may contribute to phenotypic variation in natural populations as targets of selection in a regime maintaining alternate life histories across environmental heterogeneity. Several clues suggest that functional element(s) may lie outside the coding region and affect levels of *mth* expression. Here we provide characterization of nucleotide variation in *mth* noncoding regions. We observe several polymorphisms that fall within putative transcription factor binding sites, including an indel polymorphism that is associated with differential *mth* expression.

INTRODUCTION

The G-protein coupled receptor *methuselah* (*mth*) was the first aging gene discovered in *Drosophila*. A genetic screen for lifespan-extending mutations demonstrated that flies homozygous for a P-element insertion at *mth* live 35% longer than the parental strain (Lin et al. 1998). Consistent with well-characterized correlations among life history traits and pleiotropy at other life history genes (Partridge et al. 2005; Vermeulen & Loeschcke 2006), *mth* mutants also show a reduction in fecundity (Mockett & Sohal 2006) and an increased tolerance to oxidative stress, starvation and heat stress (Lin et al. 1998). Characterization of natural genetic variation at *mth* revealed that this gene appears to have undergone rapid adaptive evolution on both phylogenetic and contemporaneous timescales. The amino acid sequence shows a very high rate of divergence since *D. melanogaster* shared a common ancestor with *D. simulans* approximately two million years ago, particularly in the regions that modulate signal transduction. Furthermore, populations spanning the U.S. east coast demonstrate a cline in frequency of the most common haplotype within the *mth* coding region (Schmidt et al. 2000). This cline is unlikely to be an artifact of population structure, as variable sites just outside the *mth* locus show independence (Duvernell et al. 2003). Instead, it may reflect a recent history of adaptive responses, as this genotypic pattern co-occurs with latitudinal variation in life history phenotypes that include longevity, stress tolerance and associated traits that may be mediated by *mth* (Schmidt et al. 2005a).

The geographical variation in *Drosophila* life histories may be an adaptive response to environmental heterogeneity (Paaby & Schmidt 2009). High latitude climates likely

impose greater stress via seasonal fluctuations in temperature and resource availability; these environments may favor genotypes that confer stress tolerance, while low latitude environments favor genotypes with highest fecundity and fastest development time. Inherent genetic correlations among these traits produce tradeoffs, generating alternate life history regimes in which suites of associated traits evolve, directly or indirectly, in response to environmental pressure. This hypothetical framework might explain how lifespan, a trait unlikely to experience direct selection, may nonetheless evolve indirectly as natural selection targets pleiotropic genes, like *mth*, that affect multiple traits. Consequently, genetic variation at *mth* may contribute to the observed phenotypic differences in natural populations, but to date no putative functional polymorphisms have been identified at *mth*. Of the five SNPs that comprise the clinal haplotype within the *mth* coding region, none are likely to be functional, as none show a significant trend individually (Schmidt et al. 2000). However, the linkage disequilibrium that maintains the haplotype and the association with the functional site(s) decays rapidly both 5' and 3' of the *mth* locus (Duvernell et al. 2003), indicating that the site(s) under selection reside within or near *mth*.

Using a modified quantitative complementation scheme, we demonstrated previously that wild-derived *mth* alleles harbor functional polymorphisms that affect lifespan, fecundity and resistance to oxidative stress (Paaby & Schmidt 2008). Several lines of evidence suggest that the functional element(s) may affect levels of gene expression. The original characterization of *mth* by mutational genetics suggested that moderately reduced *mth* expression increased lifespan and stress tolerance, but severely reduced expression was lethal (Lin et al. 1998). Our evaluation of *mth* wild-derived alleles contradicted expectations about

phenotypic determination by *mth* expression level (Paaby & Schmidt 2008), but this may be due to the previously uncharacterized deletion of a neighboring gene in the original *mth* null allele (W. Ja, personal communication). Re-interpreting our results under the assumption that the null allele reduces overall fitness by activity outside of *mth*, we find that differences in *mth* expression level are consistent with our observed variations in phenotype. A reduction in *mth* expression was observed under low calorie conditions, demonstrating that regulation of *mth* expression may be a natural response in mediating stress tolerance (Pletcher et al. 2002). Furthermore, the failure of any single polymorphism within the *mth* coding region to show patterns of functionality implies that the functional genic elements either include complex interactions among multiple sites, or affect *mth* activity without changing the protein structure. Here, we determined sequences three kb upstream of the *mth* coding region for wild-derived lines. We observed multiple polymorphisms in the 5' promoter that appear to disrupt transcription factor binding and affect levels of *mth* expression.

MATERIALS AND METHODS

Population samples. *D. melanogaster* isofemale lines were established from North American populations in Damariscotta, ME (44.03°N latitude, 69.49°W longitude), Bowdoinham, ME (44.01°N latitude, 69.90°W longitude), New Hope, PA (40.36°N, 74.95°W), Lawrenceville, NJ (40.29°N, 74.73°W), Eutawville, SC (33.39°N, 80.34°W), Morven, GA (30.94°N, 83.50°W), Jasper, FL (30.54°N, 82.95°W), Orlando, FL (28.54°N,

81.38°W), and Homestead, FL (25.47°N, 80.48°W). Samples were also provided by W. Eanes, which included chromosome-extracted lines from Mount Sinai, NY (40.95°N, 72.84°W) and by T. Morgan, which included inbred lines from Raleigh, NC (35.77°N, 78.64°W).

Sequencing. Third chromosomes were extracted using the TM6B balancer (stock 279 from the Bloomington Stock Center), to permit sequencing at *mth*. Genomic DNA was extracted from whole flies using the Wizard SV Genomic DNA Purification System (Promega), overlapping fragments of the *mth* gene were amplified in sections of approximately 500 bp by polymerase chain reaction, and the products were sequenced in both directions by the University of Pennsylvania DNA Sequencing Facility using an ABI 3100 sequencer (Applied Biosystems). The fragments were assembled into gene sequences for each line using Sequencher 4.8 (Gene Codes Corporation). A total of 22 *mth* sequences were determined for the initial dataset exploring sequence variation from approximately 3.1 kb upstream from the start codon. Additional sequence fragments encompassing PE1 were later sequenced in more lines, and electronic sequences from the Drosophila Population Genomics Project were downloaded from the website (http://www.dpgp.org) with help from A. Bergland. Sequences were aligned using ClustalW. Analyses of linkage disequilibrium were performed with DnaSPv3.14 (Rozas & Rozas 1999).

Polymorphism screen. The two indel polymorphisms we identified near the transcription start site were screened at larger sample size (a total of 301 lines) from populations listed above. For the indel located 496 bp upstream from the transcription start site, we amplified an approximately 185 bp fragment by PCR (forward primer,

CTTGGCTTAATGTCCTTTG; reverse primer, TATCGGAACTGTTAACTCG) and noted the presence of the sequence exhibiting no gap in the alignment by digestion with the restriction enzyme Tsp509I (New England Biolabs) into two fragments of 152 and 35 bp. Undigested fragments carried the allele showing a gap in the alignment. For the indel located 356 bp downstream from the transcription start site, we amplified an approximately 500 bp fragment by PCR (forward primer, GACGGACTATGGTGATATCTTG; reverse primer, TGTGGATCTGGTGACTGAG) followed by digestion with the restriction enzyme PsiI (New England Biolabs). Fragments that were digested into two fragments of 333 and 164 bp carried the "A" allele at a SNP 31 bp upstream from the indel polymorphism, which showed perfect disequilibrium with the indel allele exhibiting a gap in the sequence alignment; fragments not digested carried a "T" allele at the SNP. Both PCR reactions required Mg^{++} concentrations of at least 2.5 to 5.0 mM. The relationships between allele frequency and latitude were analyzed by linear regression, using JMPv7 (SAS Institute, Cary, NC).

Quantitative PCR. Relative levels of *mth* gene expression were measured using quantitative (real-time) polymerase chain reaction (qPCR). The test samples, lines T-28 and T-50, were derived from the population in Lawrenceville, NJ. The lines carried extracted, wild-derived third chromosomes, which differed at all the polymorphisms identified in PE1, and background-replaced, isogenic X and second chromosomes. RNA was prepared from larvae and adults: from both lines, 30 third-instar larvae were collected over 1 hour; 30 virgin females were collected over 24 hours and aged for three days on standard media; larvae and adults were snap frozen and processed using RNeasy (Qiagen). RNA was reverse transcribed using iScript cDNA synthesis kit (Bio-Rad), following the manufacturer's protocol. Relative

abundance of transcript levels was determined using an ABI 7500 Fast Real-Time PCR machine and SYBR Green PCR Master Mix (Applied Biosystems) by the $\Delta\Delta C_T$ relative quantitation method. Four technical replicates were used for each sample and relative abundance was normalized using *GAPDH2* as an endogenous control. Primer sequences for the qPCR reaction were: *GAPDH2* forward, GCGGTAGAATGGGGTGAGAC; *GAPDH2* reverse, TGAAGAGCGAAAACAGTAGC; *mth* forward, AGGATGTGTGTGCAAGCTG; *mth* reverse, CCTCGTCGGACATGTTGTC.

RESULTS

Characterization of the *mth* 5' region. For 22 wild-derived lines, we determined approximately 3.1 kb of sequence from the start codon into the 5' region of the gene (Figure 1A). We observed substantial polymorphism, including two indels on either side of the putative transcription start site (Figure 1B). The indel polymorphism located upstream from

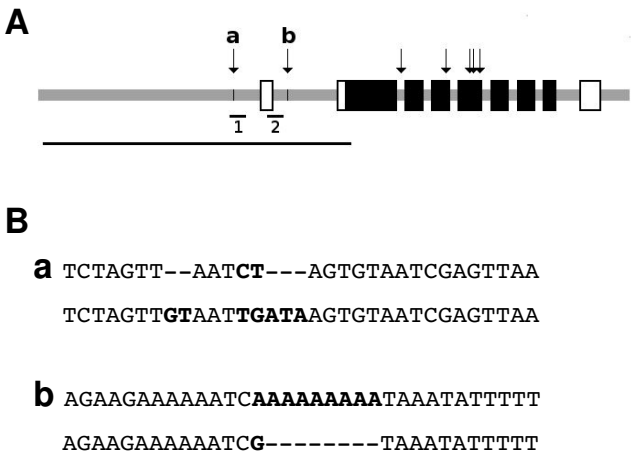


FIGURE 1. Cartoon of the *mth* gene: boxes indicate exons; open boxes indicate untranslated regions; unlabeled arrows show the locations of the five SNPs that comprise the previously identified clinal haplotype (Schmidt et al. 2000); the long horizontal line shows the region sequenced in this survey; the numbered, short horizontal lines indicate PE1 and 2 (Kim et al. 2006); labeled arrows indicate the locations of the two indel polymorphisms located just upstream (a) and downstream (b) of the transcription start site (A). Sequences of the indel polymorphisms located upstream (a) and downstream (b) of the transcription start site (B).

the start site is a composite of SNPs and missing/inserted nucleotides, affecting a total of seven bases (Figure 1B, alignment a); the indel polymorphism downstream from the start site is predominantly characterized by an insertion/deletion of eight adenine bases (Figure 1B, alignment b).

Variation across geography. If any of the polymorphisms we discovered affect *mth* activity and produce differences in phenotype, they are likely to vary in frequency across environmental gradients. To test whether either of the two indel polymorphisms we identified show clines in frequency across latitude, we screened samples derived from multiple populations along the U.S. east coast using differential restriction enzyme digestion. The rare allele of the indel located upstream from the start site showed an approximately 20% decrease in frequency across latitude (p=0.02) (Figure 2A), but the indel located downstream of the transcription start site exhibited no trend (p=0.59) (Figure 2B).

Polymorphism in the *mth* promoter. The sample from which we first identified the clinal indel polymorphism was the only sequence carrying this allele out of a total of 22

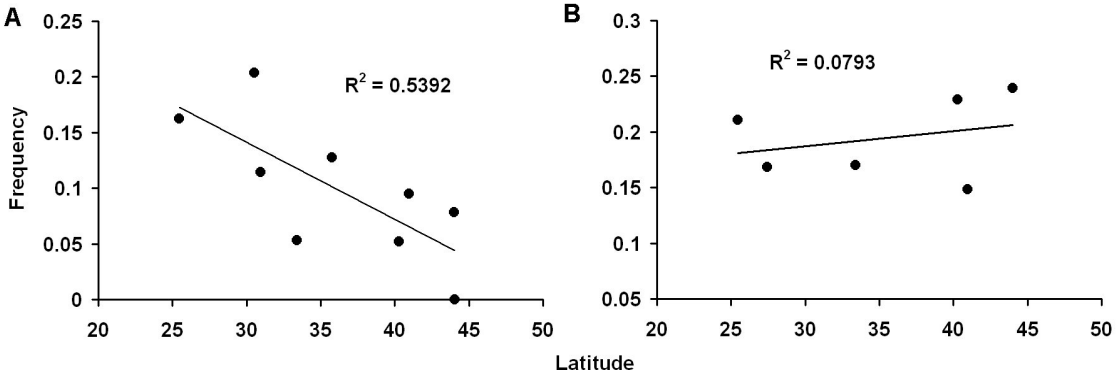


FIGURE 2. Pattern of allele frequency across latitude for two indel polymorphisms at *mth*, one 496 bp upstream from the transcription start site (A) and the other 365 bp downstream for the start site (B).

chromosomes. However, this same sequence also harbored five other single nucleotide polymorphisms (SNPs) within a 120 bp range. Earlier work has shown that these polymorphisms fall within a region identified as the first of two positive control elements (PE1) within the *mth* promoter (Kim et al. 2006). Interestingly, these polymorphisms also all fall within consensus transcription factor binding sites that were identified by Kim et al. (2006) (Figure 3). The composite indel polymorphism, which includes both SNPs and an indel of five nucleotides, affects a total of seven nucleotides that span a putative binding site in the AT2 core promoter. We observed no polymorphisms in the region identified as the second positive control element (PE2).

Linkage disequilibrium across *mth*. We have shown that the indel polymorphism

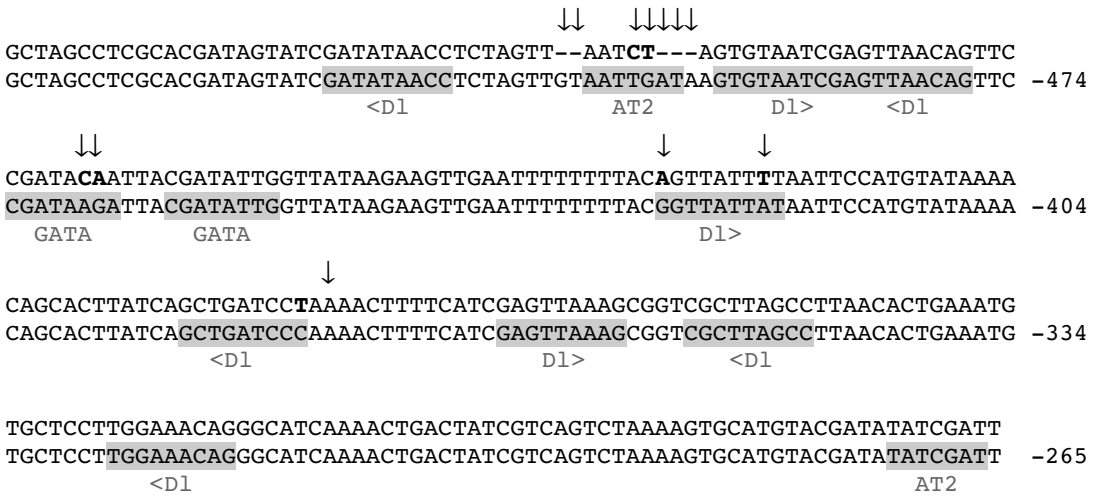


FIGURE 3. Nucleotide sequences of the positive element 1 (PE1) within the mth promoter that was identified by Kim et al. (2006). The bottom sequence represents the most common haplotype and that which was used in the identification of PE1; the top sequence shows the polymorphisms we discovered (bold, with arrows). All polymorphisms fall within or very near the consensus sites for transcription factor binding (grey shading). The putative transcription factors are labeled below the binding sites. Line numbers are relative to the putative transcription factor start site. The consensus binding sites used in the search that now show polymorphisms are: TATTGAT for the AT2 core; HGATAABV for GATA; GGGWWWHCB for Dl (Kim et al. 2006).

within PE1 is clinal, and previous work has shown that a haplotype comprised of five SNPs within the *mth* coding region is also clinal (Schmidt et al. 2000). It has been hypothesized that the clinal pattern of this coding region haplotype, "ATATC," is maintained by linkage disequilibrium with one or more other, unknown sites (Schmidt et al. 2000). To test whether our clinal indel polymorphism is in disequilibrium with the ATATC haplotype, and to survey the PE1 SNPs we observed in more samples, we sequenced or acquired sequences from an additional 12 lines. We observed disequilibrium between the clinal indel polymorphism and the ATATC haplotype: sequences carrying the common indel allele, which shows no sequence gap in the alignment, was more likely to be associated with the ATATC haplotype ($p=0.02$, Figure 4). The common indel allele and the ATATC haplotype both exhibit higher frequencies at higher latitudes. The linkage disequilibrium is not statistically significant after correcting for tests among all the polymorphisms in PE1 and the ATATC haplotype, but the indel polymorphism does show especially high linkage with the first two SNPs in the ATATC haplotype ($p=0.006$ for both), and none of the rare indel alleles co-occurred with the ATATC haplotype. Consequently, if the clinal indel is functional and under selection, its association with the ATATC haplotype should contribute to the clinal pattern at those sites. We also observed very strong (but not complete) disequilibrium between the clinal indel polymorphism and the other SNPs in PE1 (Figure 4).

Gene expression. We hypothesize that the rare PE1 haplotype, which contains the mutations described in Figure 3, may reduce *mth* expression relative to the more common haplotype. Of the 11 putative transcription factor binding sites in PE1 that were previously identified (Figure 3), four are disrupted by the mutations in the rare haplotype. All of these

	-496	-468	-467	-428	-421	-383	531	1048	1182	1188	1243
common haplotype	**no gap**	**A**	**G**	**G**	**A**	**C**	**A**	**T**	**A**	**T**	**C**
BF-54	-	-	-	-	-	-	-	-	-	-	-
S-108	-	-	-	-	-	-	-	-	-	-	-
RR-35	-	-	-	-	-	-	-	-	-	-	-
Ral-517	-	-	-	-	-	-	-	-	-	-	-
RR-3	-	-	-	-	-	-	-	-	-	-	-
RR-7	-	-	-	-	-	-	-	-	-	-	-
SL5-135	-	-	-	-	-	-	-	-	-	-	-
T-50	-	-	-	-	-	-	-	-	-	-	-
RR-40	-	-	-	-	-	-	-	-	C	A	-
SL5-96	-	-	-	-	-	-	-	-	C	A	-
BF-51	-	-	-	-	-	-	-	-	C	A	-
RR-42	-	-	-	-	-	-	-	-	C	A	A
RR-17	-	-	-	-	-	-	G	C	C	A	A
RR-21	-	-	-	-	-	-	G	C	C	A	A
S-97	-	-	-	-	-	-	G	C	C	A	A
LPS-5	gap	-	-	A	T	T	G	C	-	-	-
Ral-427	gap	-	-	A	T	T	G	C	-	-	-
Ral-514	gap	-	-	A	T	T	G	C	-	-	-
FSP08-2	gap	R	R	A	T	Y	G	C	-	-	-
Hin-37	gap	R	R	A	T	Y	G	C	-	-	-
DPF-18	gap	G	A	A	T	-	G	C	-	-	-
Ral-100	gap	G	A	A	T	-	G	C	-	-	-
DPF-44	gap	G	A	A	T	-	G	C	-	-	-
Ral-705	gap	G	A	A	T	-	G	C	-	-	-
Ral-379	gap	G	A	A	T	-	?	C	-	-	-
Ral-849	gap	G	A	A	T	-	-	-	C	A	-
T-28	gap	C	A	A	T	T	G	C	-	-	-

FIGURE 4. Variable site matrix for PE1 (-496 to -383) and the ATATC haplotype (531 to 1283). Nucleotide positions are relative to the putative transcription start site. Dashes are identical symbols relative to the common haplotype. Lines FSP08-2 and Hin-37 show heterozygosity: R indicates a sequence read of A or G; Y indicates a sequence read of C or T.

mutations are predicted to decrease binding efficiency relative to the common haplotype. These mutations include the composite mutation containing indels and SNPs that substantially disrupt a potential AT2 core binding site, two SNPs that disrupt a potential binding site for a GATA transcription factor, and SNPs that disrupt two potential binding

sites for the Dorsal (Dl) transcription factor (Figure 3).

For two samples that carried alternate alleles of the clinal indel polymorphism and the other SNPs in PE1 (T-28 and T-50, see Figure 4), we observed no difference in *mth* expression in three day old flies, but 1.26 times higher expression was associated with the common haplotype in third instar larvae ($p<0.05$) (Figure 5). These results support the prediction that the rare allele, which contains disruptions at putative transcription factor binding sites, confers lower levels of gene expression relative to the common allele. For the rare allele, transcript abundance was 5.12 times higher in larvae than in adults; for the common allele, abundance was 6.47 times higher in larvae than in adults. Little work has explored patterns of *mth* expression across tissues or developmental time, though it has been demonstrated that *mth* is required for neurotransmitter exocytosis in larval neuromuscular junctions (Song et al. 2002). Our results suggest that *mth* may be less critical for adult function, at least at young ages.

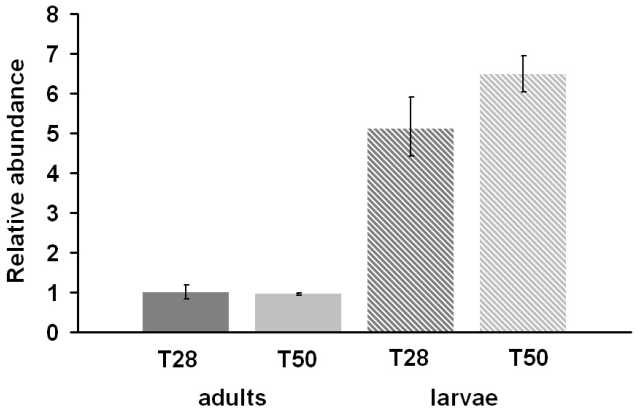


FIGURE 5. Relative *mth* transcript abundance for two alternate PE1 haplotypes, in young adults and larvae. The line T-28 carries the rare haplotype, which includes mutations that are predicted to disrupt putative transcription factor binding sites. Error bars show 95% confidence intervals.

DISCUSSION

We have provided characterization of noncoding sequence in the 5' region of the *mth* gene. In particular, we have identified a small (120 bp) region within a positive control element in the *mth* promoter that harbors substantial polymorphism, including a composite polymorphism affecting seven nucleotides within a putative transcription factor binding site for the AT2 core, and five additional SNPs that disrupt other putative transcription factor binding sites. The haplotype containing the mutations that should decrease transcription factor binding affinity shows a decreased level of *mth* transcript relative to the alternate haplotype, supporting the hypothesis that *mth* expression may be affected by these polymorphisms. Furthermore, the composite indel polymorphism within PE1 exhibits a nonrandom distribution across geography, which suggests that it may be functionally significant. However, while these results are promising, they are only preliminary. More work is required to fully evaluate whether these polymorphisms regulate *mth* activity, have functional affects on phenotype, and play a role in life history evolution.

Because *mth* mediates lifespan, stress tolerance and fecundity (Lin et al. 1998; Mockett & Sohal 2006; Paaby & Schmidt 2008), phenotypes that vary predictably across heterogeneous environments, and because we and others have demonstrated that genic elements within *mth* covary with these phenotypes across geography (Schmidt et al. 2000), we hypothesize that *mth* may contribute to adaptive responses in life history evolution. Although the clinal indel polymorphism in the *mth* promoter is associated with predictable differences in transcript abundance, the changes in allele frequency across latitude are inverse

to our expectations. Reduction in *mth* expression by mutational genetics increases longevity and tolerance to oxidative stress (Lin et al. 1998), and decreases fecundity (Mockett & Sohal 2006). These associated phenotypes are part of a life history strategy typically observed in temperate, high latitude environments, where seasonally imposed stresses are hypothesized to select for stress resistant genotypes and drive the evolution of compromised reproductive success through tradeoffs (Paaby & Schmidt 2009). Consequently, we expected that alleles conferring a reduction in *mth* expression would be found at highest frequency in high latitude populations. However, we observe that the *mth* haplotype that apparently reduces *mth* expression is at highest frequency at low latitudes (Figure 2).

This contradiction may be explained in several ways. First, *mth* encodes four distinct transcripts, and our gene expression assay did not distinguish among them; different transcripts may be differentially regulated, and determine different phenotypes. Second, although resistance to desiccation and cold stress are highest in temperate, high latitude populations (Hoffmann & Parsons 1989; Hoffmann et al. 2005), tolerance to heat stress and starvation show reverse clines (Hoffmann et al. 2002; Hoffmann et al. 2005). We have also observed that an otherwise stress intolerant allele of the *Insulin-like Receptor*, which is at high frequency at low latitudes, is more resistant to heat stress than the alternate, high latitude allele (Paaby & Schmidt in preparation). Therefore, our observation of the putatively low-expression allele at higher frequency in the more tropical environment may be a response to selection for heat or starvation resistance. This would also be consistent with the observed reduction in *mth* expression in wild type flies under caloric restriction (Pletcher et al. 2002). If true, other pleiotropic effects of reduced *mth* expression, like increased longevity and

decreased fecundity, may experience indirect positive selection. Such a dynamic would be an intriguing example in the maintenance of optimum life history strategies in the face of conflicting, pleiotropic genetic determinants. Finer evaluation of the *mth* regulatory mechanism is necessary, to discern the functional significance of individual polymorphisms, their effects on the different transcripts, and the role of the transcripts in mediating phenotypes. It may be that *mth* exhibits pleiotropy at the whole-gene level, but that individual polymorphisms within *mth* affect traits singly (Carbone et al. 2006).

Analysis of *mth* expression over different developmental stages and under different conditions should also yield valuable insight. Since it is an longevity gene, the expression of *mth* at late ages may affect lifespan. We observe differences in *mth* expression among larvae carrying different PE1 alleles, but no differences among young adults, when expression is lower; perhaps the polymorphisms differentially affect expression at old ages to mediate longevity. Likewise, the observation that wild type flies show lower *mth* expression under caloric restriction (Pletcher et al. 2002) prompts an evaluation of gene expression under different stress conditions. Finally, the observation that the clinal PE1 polymorphisms are in linkage disequilibrium with the previously characterized clinal polymorphisms in the coding region (Figure 4) (Schmidt et al. 2000), but presumably not with the nonclinal indel polymorphism that lies between (Figure 1), suggests that polymorphic influences on *mth* function may be complex. An unknown mechanism may be maintaining linkage disequilibrium between the two clinal haplotypes, and selection on promoter alleles may be driving the distribution of the coding region haplotype across geography. However, it is also possible that more than one genic region is targeted by clinal selection, including selection

for epistasis across multiple sites (Berry & Kreitman 1993). In addition to testing expression at different ages and conditions, future investigations should attempt to isolate the individual polymorphisms and replicate the test alleles across multiple lines.

ACKNOWLEDGEMENTS

This work was completed with kind assistance from several people. We thank N. Bonini and the members of her lab for providing gracious access to their qPCR machine, and A. Bergland for downloading sequences from the Drosophila Population Genomics Project. We also thank W. Eanes and T. Morgan for providing fly samples.

CHAPTER FOUR

IDENTIFICATION OF A CANDIDATE ADAPTIVE POLYMORPHISM FOR *DROSOPHILA* LIFE HISTORY BY PARALLEL INDEPENDENT CLINES ON TWO CONTINENTS[†]

ABSTRACT

Life history traits are critical components of fitness and frequently reflect adaptive responses to environmental pressures. However, few genes that contribute to natural life history variation have been identified. Insulin signaling mediates the determination of life history traits in many organisms, and single gene manipulation in *Drosophila melanogaster* suggests that individual genes in the pathway have the potential to produce major effects on these quantitative traits. We evaluated allelic variation at two insulin signaling genes, the *Insulin-like Receptor* (*InR*) and its substrate, *chico*, in natural populations of *D. melanogaster*. We found different patterns of variation: *InR* shows evidence of positive selection and clines in allele frequency across latitude; *chico* exhibits neutral patterns of evolution. The clinal patterns at *InR* are replicated between North America and Australia, showing striking similarity in the distribution of specific alleles and the rate at which allele frequencies change across latitude. Moreover, we identified a polymorphism at InR that appears to be

† This chapter has been accepted for publication: Paaby AB, Blacket MJ, Hoffmann AA, Schmidt PS. In Press. Identification of a candidate adaptive polymorphism for *Drosophila* life history by parallel independent clines on two continents. *Molecular Ecology.*

functionally significant and consistent with hypothetical patterns of selection across geography. This polymorphism provides new characterization of genic regions of functionality within InR, and is likely a component in a suite of genes and traits that respond adaptively to climatic variation.

INTRODUCTION

Life history traits include aspects of an organism's biology that directly affect reproduction and survival (Roff 1992; Stearns 1992). These traits are closely connected with fitness, and life history strategies have provided valuable examples of adaptive responses to natural selection (e.g. Endler 1986; Reznick et al. 1990; Martin 1995). Many phenotypic and genetic components contribute to life history, and complex and correlated suites of traits have evolved to maximize the fitness of life history regimes. But despite the vast diversity of life history strategies across taxa, empirical studies have demonstrated widely shared correlations among life history traits. Two major correlations underlying life history determination include the negative correlation between reproduction and survival, and the positive correlation between longevity and stress tolerance (Reznick 1985; Stearns 1991; Partridge et al. 2005; Vermeulen & Loeschcke 2006; Harshman & Zera 2007; Toivonen & Partridge 2009). These relationships between traits are mediated by pleiotropic genic elements, or by genic elements that affect single traits but co-occur through linkage disequilibrium and act as pleiotropic alleles. Such pleiotropic effects of individual genes are routinely observed in mutation genetics experiments, and indicate that genetic determinants that affect one

component of life history are also likely to affect others (Partridge et al. 2005; Paaby & Schmidt 2009).

Patterns of phenotypic variation across environmental heterogeneity can indicate adaptive responses to selection, and evaluating these patterns can yield insight into the evolutionary dynamics of natural populations. For example, variation in coat pigmentation in the oldfield mouse *Peromyscus polionotus* suggests adaptation to changes in substrate reflectance (Mullen & Hoekstra 2008), and flower color in the desert plant *Linanthus parryae* likely represents adaptation to different soil environments (Schemske & Bierzychudek 2007). Examination of life history variation across climatic gradients may provide similar understanding into how this complex suite of traits evolves. Natural populations of *Drosophila melanogaster* exhibit variation in many traits across latitudinal clines, including life history phenotypes: high latitude populations exhibit longer lifespan, lower fecundity, higher incidence of reproductive diapause, larger body size, and higher cold stress tolerance than low latitude populations (Capy et al. 1993; Mitrovski & Hoffmann 2001; De Jong & Bochdanovits 2003; Schmidt et al. 2005a; Trotta et al. 2006). Genetic variance for and genetic correlations among these traits indicate that selection in the local environment may act on certain phenotypes but drive expression of others through tradeoffs (David 1975; Anderson et al. 2003; De Jong & Bochdanovits 2003; Schmidt et al. 2005b; Rako et al. 2007; Schmidt & Paaby 2008). Variation in stress tolerance has been especially well described, and may be a key component in life history evolution (Hoffmann & Harshman 1999; Hoffmann et al. 2001; Hoffmann et al. 2005; Hoffmann et al. 2007). This framework suggests a hypothetical selection regime: high latitude, seasonally cold climates impose seasonal stress

and favor genotypes that confer stress tolerance; correlated traits, which may evolve as co-adapted responses to the same selection regime or by indirect selection via pleiotropy, are characterized by better overwintering ability, larger body size, longer lifespan, slower development and lower fecundity relative to low latitude populations (Paaby & Schmidt 2009).

Despite the well-characterized life history variation in natural populations of *D. melanogaster*, few genes have been shown to contribute to these observed phenotypic patterns (De Luca et al. 2003; Carbone et al. 2006; Paaby & Schmidt 2008; Schmidt et al. 2008). However, likely candidate genes include those characterized by extended longevity mutant or overexpression phenotypes in model systems (e.g. Kenyon et al. 1993; Lin et al. 1998; Rogina et al. 2000; Tatar et al. 2001; Clancy et al. 2001; Hwangbo et al. 2004), which routinely show correlated responses in other life history traits (Partridge et al. 2005; Paaby & Schmidt 2009). In particular, the insulin signaling pathway and its pleiotropic determination of a suite of correlated traits has been well characterized. Insulin signaling is conserved across taxa and appears universally relevant to the determination of life history phenotypes in metazoans: mutations that reduce insulin signaling in *D. melanogaster*, mutations in the dauer pathway in *Caenorhabditis elegans* and reduction of IGF-I in mice produce correlated phenotypic responses that include increases in lifespan and stress tolerance and decreases in reproductive success (Partridge & Gems 2002; Tatar et al. 2003; Giannakou & Partridge 2007). A reduction in insulin signaling by manipulation at almost any point in the pathway can induce these pleiotropic effects in either *D. melanogaster* or *C. elegans*, as demonstrated by genetic manipulations of multiple elements in both systems (Giannakou & Partridge 2007). Whether

loci within this pathway harbor significant natural genetic variation may be a function of position: across 12 *Drosophila* species, downstream elements show evidence of stronger purifying selection than upstream elements (Alvarez-Ponce et al. 2009). Two upstream pathway members include the *Insulin-like Receptor* (*InR*) and the receptor substrate, *chico*. Disruption at either gene reduces insulin signaling and produces similar mutant phenotypes, including longevity extension, increased stress tolerance, decreased growth and development and compromised reproduction (Tatar et al. 2001; Clancy et al. 2001).

The characterization of these genes in mediating organism life history offers an opportunity to identify the loci that are involved in life history evolution. However, aging genes characterized by mutation genetics do not necessarily contribute to genetic variance for traits in the wild (Curtsinger 2003; Flatt 2004). For example, genes under strong selective constraints may vary little in natural populations, or genes may harbor only neutral polymorphisms. But populations do harbor substantial allelic variation for lifespan: artificial selection experiments (e.g. Rose & Charlesworth 1981; Promislow et al. 1996), genetic correlation analyses (e.g. Tatar et al. 1996; Schmidt et al. 2005a), and identification of aging genes or chromosomal regions through QTL analysis (Nuzhdin et al. 1997; Leips & Mackay 2000; Mackay 2002; De Luca et al. 2003; Geiger-Thornsberry & Mackay 2004; Pasyukova et al. 2004; Wilson et al. 2006) have demonstrated significant genetic variance for this trait in natural populations of *D. melanogaster*. Consequently, the evaluation of natural genetic variation at candidate genes has the potential to identify functionally significant polymorphisms that affect lifespan and other life history phenotypes, and provide insight into genic function that is complementary to investigations using mutation genetics.

In *D. melanogaster*, characterization of nucleotide targets of selection is facilitated by the distribution of natural populations across latitudinal gradients. *D. melanogaster* originated in tropical Africa, and has colonized temperate regions in the North American and Australian continents within the last few hundred years (David & Capy 1988). Surveys at multiple loci have revealed patterns of variation along latitudinal clines in which the frequencies of the derived alleles increase with latitude, suggesting adaptation to novel, temperate habitats (Sezgin et al. 2004; Hoffmann & Weeks 2007). While patterns across geography can result from stochastic processes, independent clinal patterns replicated on multiple continents imply selection. For example, clinal variation at the alcohol dehydrogenase locus has been documented on several continents, implicating thermal regimes in the maintenance of allelic variation (Oakeshott et al. 1982). Furthermore, analysis of genome-wide tiling arrays generated from temperate and subtropical populations has revealed multiple regions of differentiation and parallel responses between North America and Australia (Turner et al. 2008). Consistency between continents likely reflects responses to independent but similar selection pressures, as North American and Australian populations were founded at different times and from different source populations (Bock & Parsons 1981; David & Capy 1988). Clines have also been observed in the distribution of multiple chromosomal inversions, which are likely maintained by selection on genes within or near the inversions (Hoffmann et al. 2004).

The genetic correlations between life history traits in natural populations (Schmidt et al. 2005b) and the pleiotropic expression of these same traits by mutation analyses in insulin signaling genes (Tatar et al. 2001; Clancy et al. 2001) suggest that *InR* and *chico* might

contribute to the observed genetic variance for these phenotypes in the wild. Here, we evaluated allelic variation at these loci, derived from populations spanning latitudinal gradients on two continents, for evidence of historical and contemporaneous selection. We observed substantial differences in the molecular evolution of *InR* and *chico*, suggesting that different members of the pathway have responded differently to selection. Furthermore, we identified a polymorphism in the first exon of *InR* that shows a striking nonrandom distribution across both continents and functional effects on phenotype. These results suggest that this polymorphism may play a role in the determination of adaptive life history phenotypes, and contributes new characterization of genic regions of functionality within the *Drosophila* insulin receptor.

MATERIALS AND METHODS

Population samples. To sequence *InR* and *chico*, *D. melanogaster* isofemale lines were established from North American populations in Bowdoinham, ME (44.01°N latitude, 69.90°W longitude), New Hope, PA (40.36°N, 74.95°W), Lawrenceville, NJ (40.29°N, 74.73°W), Orlando, FL (28.54°N, 81.38°W) and Homestead, FL (25.47°N, 80.48°W). Second and third chromosomes were extracted using the CyO (stock 5439 from the Bloomington Stock Center) and TM6B (stock 279 from the Bloomington Stock Center) balancers, to permit sequencing at *chico* (second chromosome) and *InR* (third chromosome). Putatively clinal *InR* polymorphisms identified by sequencing were then screened at larger sample sizes, using genomic preparations from isofemale lines or wild-caught males from

additional collections from North American populations (encompassing a total of nine populations across 18.6° latitude) and from wild-caught females (or single F1 individuals from wild-caught females) from populations in eastern Australia (encompassing a total of 17 populations across 27.3° latitude). The North American samples were collected from Bowdoinham, ME, Lawrenceville, NJ, Eutawville, SC (33.39°N, 80.34°W), Morven, GA (30.94°N, 83.50°W), Jasper, FL (30.54°N, 82.95°W), Fort Pierce, FL (27.45°N, 80.33°W) and Homestead, FL. North American lines were also provided by W. Eanes, which included chromosome-extracted lines from Mount Sinai, NY (40.95°N, 72.84°W) and by T. Morgan, which included inbred lines from Raleigh, NC (35.77°N, 78.64°W). A total of 537 North American lines were used in the polymorphism screen. The Australian samples were collected from Cooktown (15.47°S, 145.25°E), Cape Tribulation (16.02°S, 145.48°E), Innisfail (17.52°S, 146.03°E), Cardwell (18.25°S, 146.03°E), Mackay (21.13°S, 149.18°E), Gladstone (23.83°S, 151.25°E), Maryborough (25.53°S, 152.93°E), Brisbane (27.47°S, 153.02°E), Kingscliff (28.25°S, 153.57°E), Coffs Harbour (30.30°S, 153.13°E), Port Maquarie (31.42°S, 152.92°E), Sydney (33.87°S, 151.22°E), Nowra (34.87°S, 150.60°E), Bega (36.67°S, 149.83°E), Dromana (38.33°S, 144.97°E), Spreyton (41.22°S, 146.25°E), and Sorrell (42.78°S, 147.42°E). A total of 384 Australian lines were used in the polymorphism screen. Both alleles were counted in all samples except those from Mount Sinai and Raleigh, which had isogenic chromosomes. Lines were screened for the *In(3R)Payne* inversion following Matzkin et al. (2005) for the North American samples and Anderson et al. (2005) for the Australian samples. The association between the *InR* indel polymorphism and *In(3R)Payne* was evaluated by nominal logistic regression, modeling the

log odds (standard/inverted), using JMPv5 (SAS Institute, Cary, NC).

Sequencing. To sequence North American samples, genomic DNA was extracted from whole flies using the Wizard SV Genomic DNA Purification System (Promega), the *InR* and *chico* genes were amplified in sections of approximately 500 bp by polymerase chain reaction, and the products were sequenced in both directions by the University of Pennsylvania DNA Sequencing Facility using an ABI 3100 sequencer (Applied Biosystems). Overlapping sequence fragments were assembled into gene sequences for each line using Sequencher 4.8 (Gene Codes Corporation). A total of 27 *chico* and 52 *InR* sequences were determined. Sequences have been deposited in GenBank under accession nos. GQ927177-GQ927244. To characterize the glutamine-histidine indel length polymorphism in the Australian populations, an approximately 240 bp fragment in the first exon at *InR* was sequenced from 28 Australian samples. *InR* and *chico* sequences from *D. simulans* and *D. yakuba* were acquired electronically through the UCSC Genome Bioinformatics website, http://genome.ucsc.edu/. Comparison of the *InR* indel sequence region across 12 *Drosophila* species was performed with the Vista Genome Browser, http://pipeline.lbl.gov/.

Sequence analysis. Sequences were aligned using ClustalW. Analyses of polymorphism and linkage disequilibrium, estimates of nucleotide diversity, and tests for neutrality and divergence were performed with DnaSPv3.14 (Rozas & Rozas 1999). Linkage disequilibrium between the indel polymorphism in the first *InR* exon and the 177 observed SNPs at that locus was evaluated by computing Fisher Exact Test p-values for three by two contingency tables in the R programming language. For expediency, only sequences with the three most common indel alleles (representing 87% of the total alleles) were used in the

analysis (seven sequences were excluded). The selection coefficient was estimated using unfolded and folded configurations of silent and replacement changes by applying the Poisson Random Field method to the *InR* and *chico* sequence datasets, following Hartl et al. (1994). The frequency distributions at sites were polarized using the *D. simulans* sequence as an outgroup.

Survey for polymorphism. Three SNPs were screened at larger sample sizes, described above, using differential restriction enzyme digestion after PCR amplification of specific fragments. BtgI cut the SNP at position 1468, MboI cut the SNP at position 3052 and BsgI cut the SNP at position 4531 (positions relative to GenBank accession no. GQ927244). Primer pairs were as follows: for SNP 1468, forward AACCCAACTGGTGGTG CTG, reverse GCAGAGTTTGCTGTTCCAG; for SNP 3052, forward GGTCATGGTATTAA GCAATTTG, reverse AGGCTTATCTAGGTAGCTCC; for SNP 4531, forward TCAATAGC GGGATACGGC, reverse GAGCCAACTGAATGATGTTC. Fragments were amplified in 30 cycles using 1.5 mM Mg^{++}; annealing temperatures were 55°C (SNP 1468) and 52°C (SNPs 3052 and 4531). The *InR* indel polymorphism was screened at larger sample sizes after fragment amplification with a fluorescent-tagged primer, using either an Applied Biosystems 3100 capillary sequencer or using 6% acrylamide gels on a Licor system (IR^2, BioSciences). Primer pairs for amplification of this fragment were forward CAATATCTTTAGCAACTGT CAC and reverse TTTAGGGCTTAAACTCAGTC. This fragment was amplified in 30 cycles using 2.5 mM Mg^{++} with an annealing temperature of 51.1°C. Digestion conditions followed protocols provided by New England Biolabs. Purification by drop dialysis (Millipore 0.025 μm VSPW filter) of the PCR product containing SNP 4531 was usually required for clean

digestion by BsgI. The indel identities named in this paper refer to fragment lengths generated by these primers. The relationships between allele frequency and latitude were analyzed by linear regression, using JMPv5 (SAS Institute, Cary, NC).

Phenotype assays. The effects of the 248 and 254 *InR* alleles on phenotype were tested in three assays, using lines developed to minimize confounding background effects. All lines were maintained on standard cornmeal-molasses media at room temperature and subject to ambient light cycles. First, a third chromosome carrying the 248 allele and a third chromosome carrying the 254 allele, derived from the Mount Sinai population, were extracted using balancers and put in a background with isogenic X (w* from stock 2475 from the Bloomington Stock Center) and second (from stock 6326) chromosomes. A stock with fused second and third chromosomes (stock 2475) was used to facilitate the selection of +/CyO;+/TM3 progeny in the F1 generation. The two extracted lines carrying the 248 and 254 chromosomes were crossed, and the offspring permitted to recombine for four generations. Individual third chromosomes in the F4 generation were again extracted with balancers, and each line was genotyped for the *InR* indel polymorphism and for two of the clinal SNPs, 3052 and 4531 (SNP 1468 was identical between the parentals). Four 248 lines and four 254 lines, for which genetic variation was randomly distributed on the third chromosome (outside of the interval containing *InR*) across all lines, and for which the X and second chromosomes were isogenic across all lines, were used in each assay. Bottle cultures were reared at low density to minimize confounding environmental effects. Flies were collected, freshly eclosed, over 24 hours, and sorted into vials of five males and five females each. To measure tolerance to oxidative stress, flies in three replicate vials of each of the

eight lines were aged at 25°C for five days, then transferred into media-free vials with cotton saturated with 1 mL of 30 mM methyl viologen (paraquat) in 5% sucrose solution. After 48 hours of continuous exposure, patterns of mortality were determined for females in all replicates. The oxidative stress data were analyzed with nominal logistic regression, modeling the log odds (mortality/survivorship). To measure recovery from chill coma, flies in five replicate vials of each of the eight lines were aged at 25°C for four days, then females were removed (with minimal CO_2 exposure) and aged another 24 hours in vials with fresh media. To induce cold stress, vials were completely covered in ice and placed at 4°C for three hours, then restored to room temperature. Time to recovery (transition to the upright position) was recorded using a video camera and analyzed by ANOVA. To measure fecundity, flies were transferred to fresh vials (without topical yeast) and eggs were counted every day for two weeks. Three replicates of each of the eight lines were used in this assay. Cumulative fecundity was analyzed by ANOVA. Line nested within allele was treated as a random effect in all statistical analyses, which were performed using JMPv5 (SAS Institute, Cary, NC).

RESULTS

Polymorphism. Sequencing revealed a trend in allele frequency across latitude for several *InR* polymorphisms, including an amino acid insertion-deletion (indel) polymorphism in the first exon. This indel polymorphism disrupts repeats of glutamine and histidine, and the same six sequence variants were observed in both the North American and Australian

sequence datasets (Figure 1). The remainder of the *InR* locus exhibited substantial polymorphism in the North American sequence data: 88 synonymous single nucleotide polymorphisms (SNPs), 15 nonsynonymous SNPs, and 14 other indels (all but one of which are intronic) were observed. Of the observed polymorphisms at *InR*, 13 showed trends in frequency across latitude, including the indel polymorphism in the first exon and 12 SNPs well-distributed across the locus (Figure 2). Considerable polymorphism was also found at *chico*: 25 synonymous SNPs, 12 nonsynonymous SNPs, and one indel were observed. However, no trend in allele frequency across latitude was observed at *chico*.

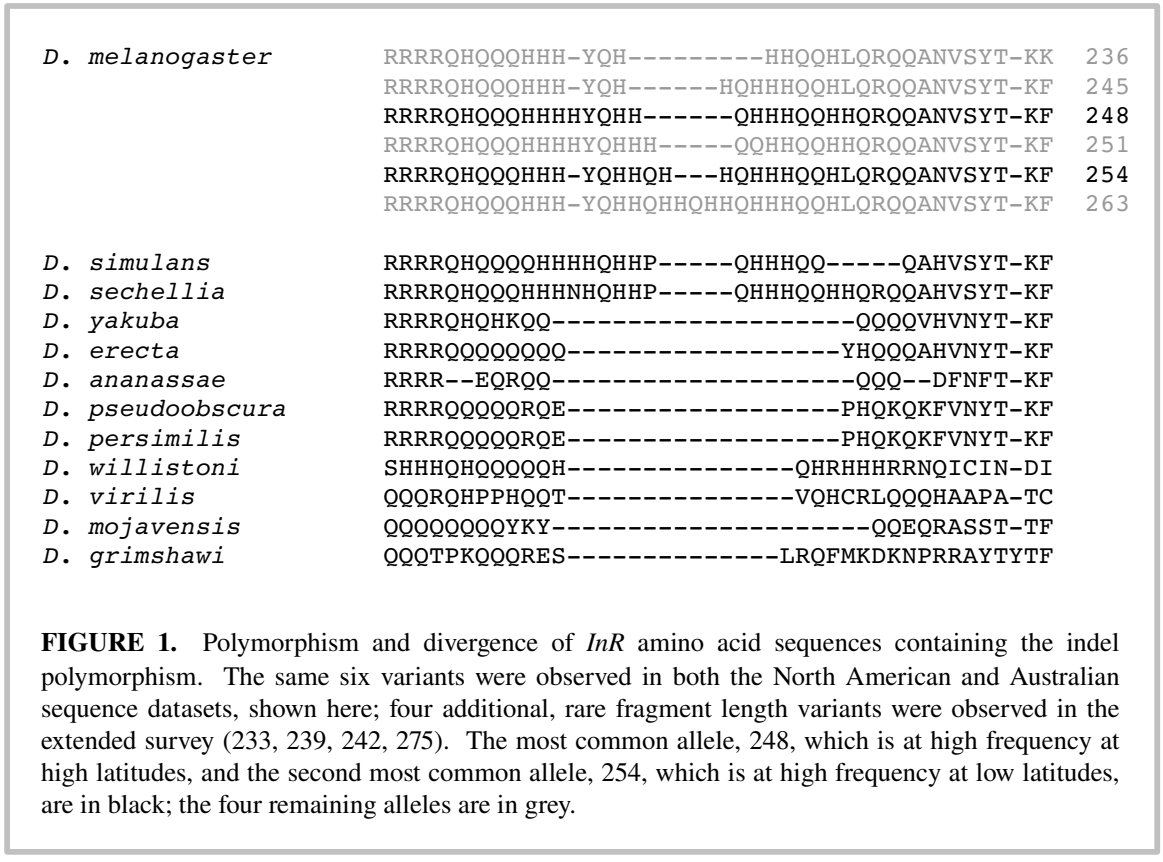


FIGURE 1. Polymorphism and divergence of *InR* amino acid sequences containing the indel polymorphism. The same six variants were observed in both the North American and Australian sequence datasets, shown here; four additional, rare fragment length variants were observed in the extended survey (233, 239, 242, 275). The most common allele, 248, which is at high frequency at high latitudes, and the second most common allele, 254, which is at high frequency at low latitudes, are in black; the four remaining alleles are in grey.

Our sequencing results suggested that there might be four regions of functional interest at *InR*. Linkage disequilibrium analysis showed that the 12 clinal SNPs at *InR* cluster tightly into three regions of disequilibrium (Figure 2). Within each cluster, the SNPs are in perfect or near-perfect disequilibrium; among clusters, the SNPs are independent. These three SNP clusters are also independent of the indel polymorphism in the first exon. Consequently, we hypothesized that the indel polymorphism and any one (or several) of the SNPs within each cluster might be functionally important, affecting phenotype and experiencing differential selection across latitude. But because only three latitudinal regions from five populations were represented in our sequence dataset, and because clinal patterns can occur randomly, we expanded our polymorphism survey with larger sample sizes, replicated on two continents.

Geographical patterns. Screened at larger sample sizes, the indel polymorphism in the first exon at *InR* exhibits a significant cline in frequency across latitude in both North

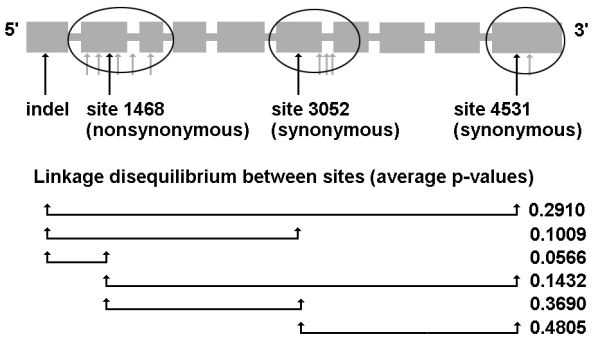


FIGURE 2. Diagram of the *InR* gene showing clinally varying polymorphisms. Grey boxes are exons, interior grey lines are introns. The indel polymorphism and the 12 SNPs exhibiting clines in frequency in the sequence data are indicated by arrows. Black arrows indicate the four polymorphisms that were screened at larger sample sizes. From left to right, the positions of the 13 polymorphisms, beginning with the indel and with reference to GenBank accession no. GQ927244, are: 430, 1277, 1466, 1468, 1493, 1566, 1753, 3052, 3093, 3122, 3124, 4531, 4534. Circles indicate which SNPs cluster together by linkage disequilibrium, and the horizontal lines show the average disequilibrium between clusters. Linkage disequilibrium was estimated by Fisher Exact Test; the numbers report the average p-values for associations across clusters. All disequilibrium p-values for sites within clusters were <0.0001. The diagram is not accurate to scale.

America and Australia (Figure 3, Table 1). The two most common alleles, which comprise 65.1% and 66.1% of the total alleles in North America and Australia, respectively, vary significantly with latitude. The 248 allele (named for its PCR fragment length) is rare at low latitudes and increases in frequency with latitude ($p=0.0018$ in North America; $p<0.0001$ in Australia); the 254 allele is common at low latitudes and decreases in frequency with latitude ($p=0.0134$ in North America; $p=0.0006$ in Australia). The reciprocal frequencies of the 248 and 254 alleles and the redundant clines on both continents strongly suggest nonrandom distribution by selection. Moreover, the rates at which the allele frequencies change across latitude are very similar between continents (Figure 3). Comparisons of the regression coefficients for both alleles show no significant differences between the continents (for 248, $F_s=0.4930$; for 254, $F_s=3.6908$; df=1,22; critical F=4.301), and on both continents the two common alleles are present at equal frequencies at approximately 25° latitude (Figure 3). In

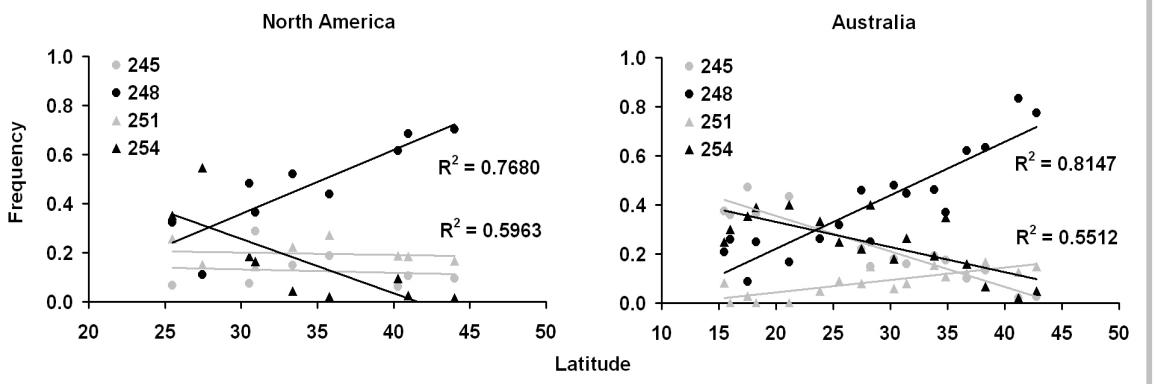


FIGURE 3. Allele frequencies across latitude for the four most common variants of the *InR* indel polymorphism. The two most common variants, 248 and 254, show significant clines in frequency on both continents (see Table 1) and are shown in black. The slopes of the regression lines also show similarity between continents (North America: slope 248=0.0260, SE=0.0053, slope 254=-0.0221, SE=0.0067; Australia: slope 248=0.0218, SE=0.0027, slope 254=-0.0102, SE=0.0024; see text for statistical comparisons).

addition to the six fragment length variants observed in the sequence data, this survey revealed another four (rare) alleles.

None of the other *InR* polymorphisms screened at large sample size demonstrates a significant trend across latitude on both continents. Representative SNPs from each of the three clusters showing clinal patterns in the sequence dataset were screened, but only the Australian populations showed a significant pattern of allele frequency across latitude (Figure 4, Table 1). However, the regression coefficients for each SNP are not significantly different between continents (for SNP 1468, F_s=0.1384; for SNP 3052, F_s=0.2274; for SNP 4531, F_s=0.6512; df=1,22; critical F=4.301). This may reflect similar responses across the continents but insufficient power for detection in North America, where fewer populations were screened over a narrower range of latitude. Of the 12 SNPs comprising these clusters, one affects the amino acid sequence (Figure 2); the other 11 are synonymous or intronic. While any of these SNPs could be under direct selection, the robust clinal patterns of the indel polymorphism make it the strongest candidate functional polymorphism.

The clinal pattern of the *InR* indel polymorphism does not appear to be caused by linkage with another site on the third chromosome. The *In(3R)Payne* inversion exhibits a cline in frequency across latitude in both North America and Australia (Knibb 1982; Anderson et al. 2005), and worldwide there is a significant association between the indel polymorphism and the chromosomal arrangement: the 248 allele is 3.29 times more likely to be associated with the standard chromosome than is the 254 allele (odds ratio=25.74; $p<0.0001$). Such an association would be expected between any two clinal elements. Furthermore, two pieces of evidence suggest that the observed changes in the indel allele

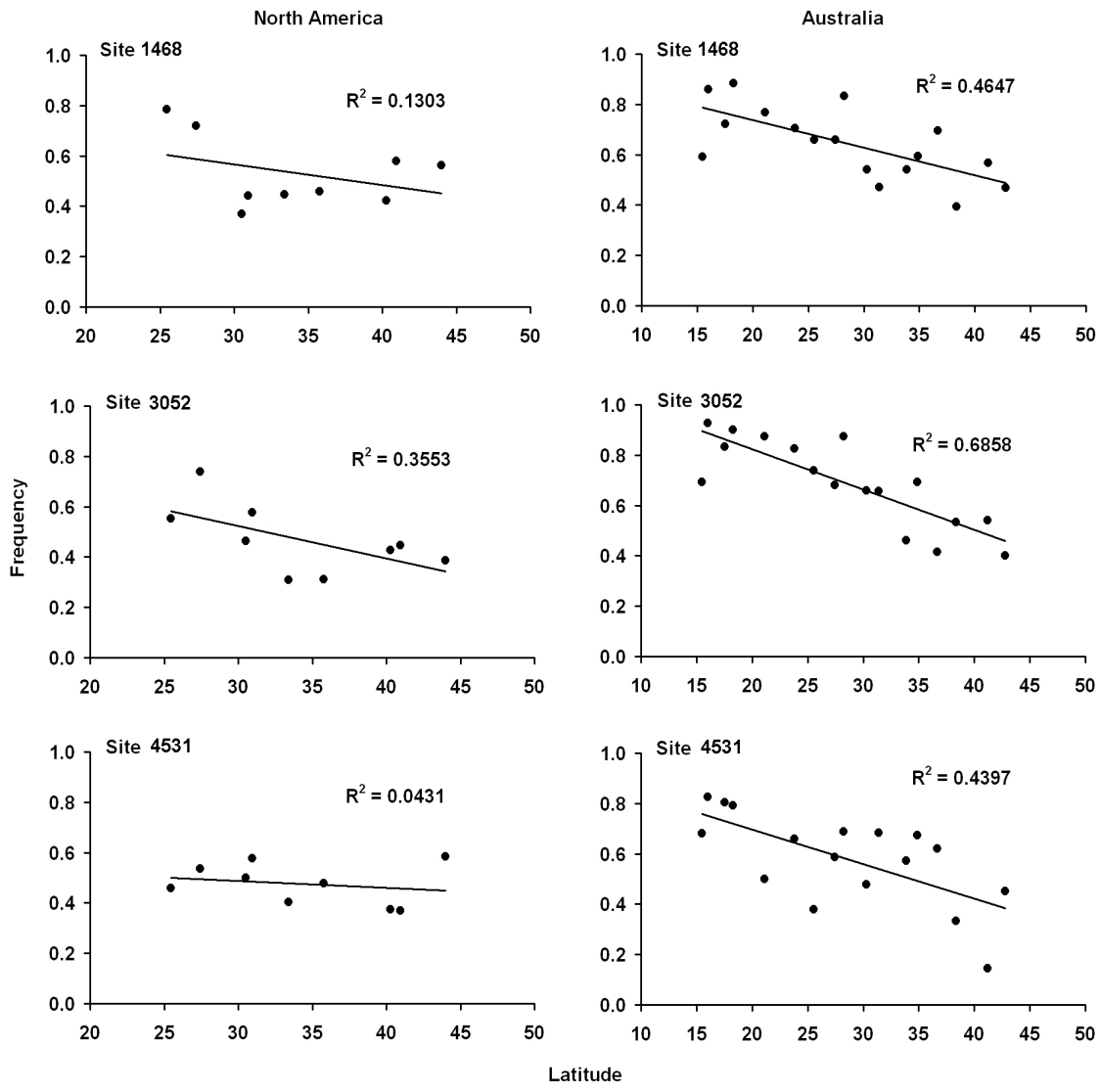


FIGURE 4. Allele frequencies across latitude for three *InR* SNPs which showed clinal trends in the sequence data. Each SNP represents a cluster of SNPs exhibiting high linkage disequilibrium. Only the Australian populations showed statistically significant changes in allele frequency across latitude (see Table 1), but the slopes of the regression lines are similar between continents (North America: slope 1468=-0.0082, SE=0.0078, slope 3052=-0.0129, SE=0.0065; slope 4531=-0.0028, SE=0.0047; Australia: slope 1468=-0.0110, SE=0.0030, slope 3052=-0.0161, SE=0.0028; slope 4531=-0.0136, SE=0.0040; see text for statistical comparisons).

frequency are not driven by association with this inversion. First, in North America, the magnitudes of the indel clines are stronger than those reported for the inversion (Knibb 1982).

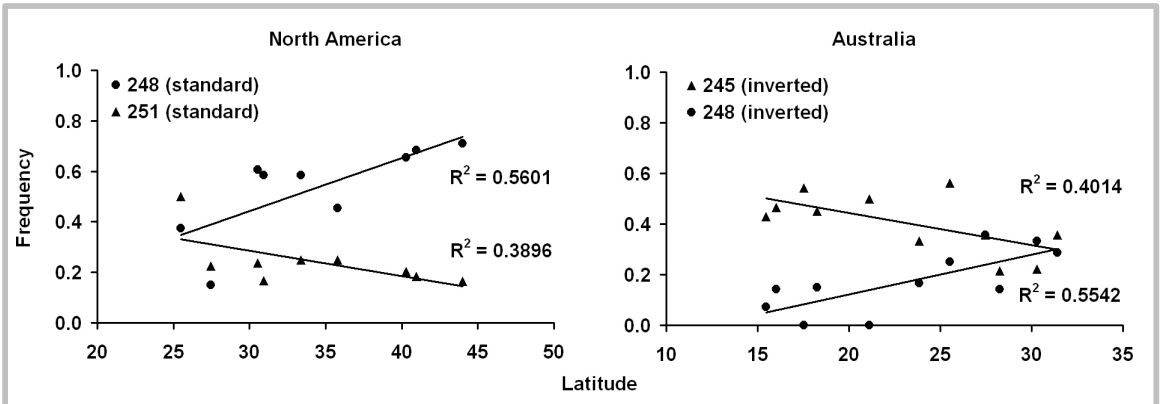


FIGURE 5. Allele frequencies across latitude for *InR* indel alleles on either standard or *In3RPayne* inverted chromosomes. The alleles shown here exhibit statistically significant or near-significant changes across latitude (see Table 1).

Second, indel clines generated using only standard chromosomes in North America and only inverted chromosomes in Australia still show significance (Figure 5, Table 1); low sample size of specific alleles at the ends of the clines (i.e. 248 at low latitudes, 254 at high latitudes) probably constrained detection of significance in additional comparisons. It is also possible that the indel's geographical patterns are driven by disequilibrium with another nucleotide polymorphism. However, of the 177 SNPs we observed within the *InR* locus, only 11 are in significant disequilibrium with the indel polymorphism (Figure 6). Nine of these are

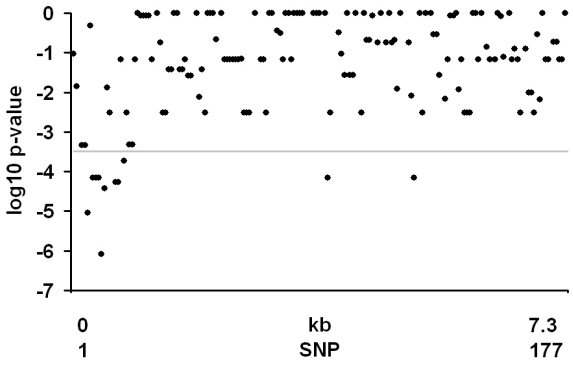


FIGURE 6. Linkage disequilibrium estimates between the *InR* indel polymorphism and the 177 SNPs observed across the *InR* locus. Disequilibrium was estimated by Fisher Exact Test; p-values below the horizontal line indicate associations that are statistically significant at the 0.05 level, after correction for multiple tests. Most of the SNPs in significant linkage disequilibrium reside with a few hundred bases of the indel polymorphism, and decay of disequilibrium is demonstrated in both the 5' and 3' directions.

close neighbors; the remaining two encode a silent third-position site and an intronic site. Hence, it is unlikely that the indel polymorphism is linked to a site within *InR* that is responsible for the cline. Furthermore, despite occurring in the 5' region of the sequence, we observe significant decay of linkage disequilibrium both 5' and 3' of the indel polymorphism (Figure 6). While we cannot rule out the possibility that this polymorphism is linked to a site outside the locus we sequenced, all our evidence suggests that the indel polymorphism is an independent site under direct selection.

Phenotype assays. In assays testing the effect of the *InR* polymorphism on stress tolerance and fecundity, lines carrying the 248 and 254 alleles showed significantly different contributions to phenotype, and in the direction predicted by the allele frequencies at low and high latitudes (Figure 7, Table 2). Flies homozygous for the 248 allele, which is at high frequency at high latitudes, tolerated oxidative stress better than flies homozygous for the 254 allele ($p=0.0007$), and recovered faster from chill coma ($p=0.0305$). Alternately, flies with the 254 allele laid more eggs than those with the 248 allele ($p=0.0030$). These effects are accordant with phenotypic effects of laboratory-induced mutations at *InR* and *chico*, which show increased stress tolerance and reduced fecundity when insulin signaling is reduced (Tatar *et al*. 2001; Clancy *et al*. 2001). Consequently, the differences in performance between lines carrying the naturally derived *InR* alleles imply that the 248 allele reduces insulin signaling relative to the 254 allele. The results are also consistent with the hypothesis that the nonrandom allele distributions across latitude are driven by alternate selection regimes imposed by environments with varying degrees of seasonal stress. The effect of line (nested within allele) was significant or marginally significant for the two stress assays, but there was

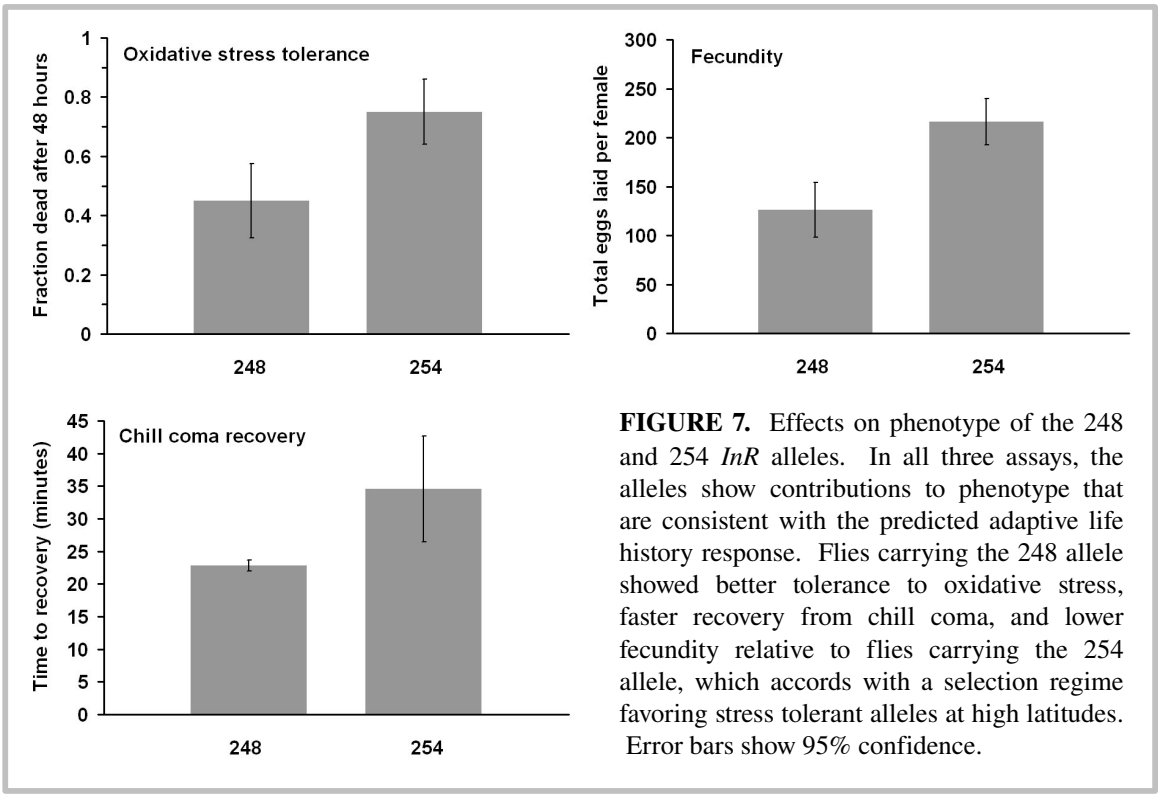


FIGURE 7. Effects on phenotype of the 248 and 254 *InR* alleles. In all three assays, the alleles show contributions to phenotype that are consistent with the predicted adaptive life history response. Flies carrying the 248 allele showed better tolerance to oxidative stress, faster recovery from chill coma, and lower fecundity relative to flies carrying the 254 allele, which accords with a selection regime favoring stress tolerant alleles at high latitudes. Error bars show 95% confidence.

no line effect at all in the fecundity assay (Table 2). These results do not provide direct evidence that the 248 and 254 alleles confer differences in stress tolerance or fecundity, as the indel polymorphism was not perfectly isolated from other allelic variation in the recombinant strains used in the tests. In these strains, the two clinal SNPs at sites 3093 and 4574 retained the linked allele associations present in the original parental lines (the clinal SNP at site 1508 was indentical between the parentals). Consequently, although these results demonstrate a significant association between the indel alleles and predicted phenotypes, they cannot distinguish between contributions from the indel and any other linked polymorphisms within the interval containing *InR*.

Tests for neutrality and selection. Classical tests of molecular evolution revealed evidence of selection at *InR*, but not at *chico*. *InR* showed evidence of adaptive protein evolution over the approximately two million years since *D. melanogaster* shared a common ancestor with *D. simulans*: the McDonald-Kreitman test (McDonald & Kreitman 1991) for divergence demonstrated an excess of fixed replacement changes between the species (Table 3). These differences between species were evenly distributed across both lineages, according to polarization by an outgroup *D. yakuba* sequence (data not shown), and consistent with other evidence for amino acid evolution at *InR* along the *D. melanogaster* lineage (Guirao-Rico & Aguadé 2009). Our results are qualitatively identical using sequences derived from just a single population (from Bowdoinham, ME, N=20) and gene trees showed no evidence of sequence similarity by geography, indicating that the test is unlikely to be biased by population structure within our sample set (data not shown). The polymorphisms we screened at larger sample size showed that each allele state was present in each chromosomal arrangment, indicating that there is significant genetic exchange even at sites relatively close to the *In(3R)P* inversion break points. Consequently, we included sequences derived from both chromosomal arrangements in these analyses. However, the McDonald-Kreitman test at *InR* does show different results when only standard and only inverted chromosomes are evaluated: samples with standard chromosomes lose statistical significance ($p=0.31$), largely due to a decrease in the number of synonymous polymorphisms, but samples with inverted chromosomes gain significance, following decreases in both synonymous and replacement polymorphisms ($p<0.01$) (Table 3). The McDonald-Kreitman test showed no significant results at *chico*, and the null hypothesis of neutral protein evolution could not be rejected

(Table 3). The McDonald-Kreitman test evaluates evolution over a relatively long timescale, while the clinal polymorphism screen assumes contemporaneous evolution. Since the analyses show adaptive patterns at *InR* and neutral patterns at *chico* over both macro- and microevolutionary time, these results may reflect differences in functional constraint between the genes that is both ancient and and modern. *InR* also shows evidence of selection according to estimates of the selection coefficient. These estimates were determined by applying the Poisson Random Field (PRF) method to the frequency distributions of both silent and replacement SNPs (Hartl et al. 1994). Although the unfolded frequency distribution at replacement sites yielded a statistically neutral estimate of the selection coefficient (γ_{est}=0.42, CI=-2.18 to 2.63), both the unfolded and folded distributions of silent SNPs at *InR* showed significantly positive estimates (γ_{est}=4.09, CI=1.85 to 11.82; γ_{est}=1.98, CI=0.24 to 18.71). We interpret these results as evidence of selection on replacement sites linked to silent sites. The PRF method assumes independence between sites, but this is not biologically realistic: an analysis of all polymorphisms across *InR* showed patterns of linkage disequilibrium that are typical for *D. melanogaster* (Miyashita et al. 1993; Long et al. 1998; Langley et al. 2000), in which disequilibrium extends for 1-2 kb (data not shown). Estimates for *chico* by both silent and replacement unfolded sites were not significantly different from zero (γ_{est}=-0.23, CI=-1.65 to 2.20; γ_{est}=0.97, CI=-1.30 to 17.80). In our study, the estimation of the selection coefficient may not be an especially valuable measure because of the assumption of site independence in the face of known linkage disequilibrium. However, these estimates are consistent with our other results, which show evidence of selection at *InR* but neutrality at *chico*. For both genes, nucleotide diversity was estimated at levels typical for *D.*

melanogaster (*InR* π_{est}=0.0049; *chico* π_{est}=0.0033). Tests for neutrality, including the Tajima (1989) test and the Fu and Li (1993) test, did not reveal any significant departures from the null hypothesis (for *InR*, Tajima's D=-0.16, p>0.10, Fu and Li's D=-1.63, p>0.10; for *chico*, Tajima's D=-0.21, p>0.10, Fu and Li's D=0.26, p>0.10). While the presence of the indel polymorphism at reciprocal frequencies across latitude is evidence of balancing selection, the absence of significantly positive D values for the neutrality tests at *InR* may be a function of an overrepresentation of high latitude sequences in our dataset, coupled with relatively few polymorphisms contributing to a signal for overdominance by linkage with the indel (Figure 6).

TABLE 1. Results from the linear regression analyses evaluating *InR* allele frequencies across latitude.

North America					Australia				
Indel polymorphism (All samples)									
Allele	F	df	R^2	p	Allele	F	df	R^2	p
245	0.1183	1,7	0.0166	0.7410	245	104.0094	1,15	0.8740	<0.0001***
248	23.6822	1,7	0.7680	0.0018**	248	65.9628	1,15	0.8147	<0.0001***
251	0.1612	1,7	0.0225	0.7001	251	26.5220	1,15	0.6387	0.0001***
254	10.7729	1,7	0.5963	0.0134*	254	18.4251	1,15	0.5512	0.0006***
SNPs representative of the three linked clusters									
Site	F	df	R^2	p	Site	F	df	R^2	p
1508	1.1085	1,7	0.1303	0.3274	1508	13.0197	1,15	0.4647	0.0026**
3039	3.9521	1,7	0.3553	0.0871	3039	32.7394	1,15	0.6858	<0.0001***
4574	0.3462	1,7	0.0431	0.5747	4574	11.7721	1,15	0.4397	0.0037**
Indel polymorphism (Standard chromosomes only)[a]									
Allele	F	df	R^2	p	Allele	F	df	R^2	p
245	1.3779	1,7	0.1645	0.2788	245	1.2729	1,5	0.2030	0.3104
248	8.9143	1,7	0.5601	0.0204*	248	2.3958	1,5	0.3242	0.1823
251	4.4678	1,7	0.3896	0.0724	251	0.7031	1,5	0.1231	0.4400
254	1.1042	1,7	0.1363	0.3283	254	5.4709	1,5	0.5226	0.0665
Indel polymorphism (Inverted chromosomes only)[a,b]									
					Allele	F	df	R^2	p
					245	6.0345	1,9	0.4014	0.0364*
					248	11.1870	1,9	0.5542	0.0086**
					251	2.9945	1,9	0.2497	0.1176
					254	1.1762	1,9	0.1156	0.3063

[a]Only populations with seven or more of the indicated chromosomal states were used in the analysis.
[b]The difference between heterozygous and homozygous inverted chromosomal arrangements could not be determined in the North American samples, so results from inverted chromosomes from these populations are not reported.

TABLE 2. Results of the assays testing for functional differences between *InR* alleles.

Oxidative stress assay: Nominal logistic model effect Wald tests				
Source	DF		χ^2	p
allele	1		11.4927	0.0007***
line[allele]	6		11.5941	0.0717
Chill coma assay: ANOVA				
Source	DF	MS	F	p
allele	1	2.477×10^7	7.9284	0.0305*
line[allele]	6	3124151	7.7423	<0.0001***
error	16	403518		
Fecundity assay: ANOVA				
Source	DF	MS	F	p
allele	1	1213650	23.0509	0.0030**
line[allele]	6	52651	0.9345	0.4973
error	16	56341		

TABLE 3. Results of the McDonald-Kreitman test for divergence at *InR* and *chico*.

	Fixed differences		Polymorphisms		
	S	NS	S	NS	p[a]
InR	101	37	88	15	0.03
standard	109	38	59	14	0.31
In(3R)P	106	37	62	5	<0.01
chico	56	27	25	12	1.0

[a]P-values were computed by Fisher's Exact Test using a two-tailed distribution.

DISCUSSION

By examining patterns of nucleotide variation across latitude, we have identified a putative functional polymorphism in the first exon at *InR*. The distribution of the 248 allele at high latitudes and the 254 allele at low latitudes, replicated across two continents, is strongly suggestive of selection. Moreover, the functional tests suggest that the alleles confer effects that are consistent with adaptation to the environment in which they are most common: lines carrying the 248 allele, which is common in cold climates, show enhanced stress tolerance; lines carrying the 254 allele, which is common in warm climates, show higher fecundity. We hypothesize that the 248 allele reduces insulin signaling relative to the 254 allele, and that seasonally imposed stress maintains high frequencies of the 248 allele at high latitudes, which would otherwise experience negative selection due to its pleiotropic cost to reproductive success. We are currently measuring insulin signaling using biochemical and gene expression assays, to determine if signaling levels differ between the alleles. This mechanism would be consistent with the role of insulin signaling in mediating stress tolerance, body size, lipid content, longevity and reproduction (reviewed in Giannakou & Partridge 2007), observed clines in these traits across latitude (e.g. David 1975; Karan et al. 1998; Hoffmann et al. 2005; Schmidt et al. 2005a, Trotta et al. 2006; Schmidt & Paaby 2008), and the observed genetic correlations between these traits among natural strains (Hoffmann et al. 2001). Consequently, the identification of this polymorphism might be an example of how a specific, pleiotropic genic element may contribute to the life history tradeoffs between reproduction, stress tolerance, and very likely, longevity.

However, the phenotype assays do not provide direct evidence that the *InR* indel alleles confer differences in stress tolerance or fecundity, as the polymorphism was not completely isolated from other genetic variation within the recombinant test lines. Rather, the assays provide compelling preliminary results in support of our adaptive hypothesis, by demonstrating a significant association between chromosomal fragments harboring distinct alleles and the expected phenotypic outcome. We have not characterized the recombination breakpoints in the test lines, but it is likely that the intervals containing the distinct *InR* indel alleles span many genes. Current work is underway to test the functional significance of this amino acid polymorphism more robustly, using higher replication and additional alleles from other source populations, and to evaluate how the polymorphism affects a broad suite of traits, including lifespan and other phenotypes, levels of insulin signaling, and other dynamics within the pathway.

Natural genetic variation at other sites may contribute to the patterns we observe. However, clinal patterns are weaker at the other three polymorphisms we screened, showing significance only in Australia. We conclude that the indel polymorphism is the most likely functional candidate polymorphism at *InR*. However, associations between phenotypes and cosmopolitan inversions have been observed in natural populations of *D. melanogaster*, and are likely due to selection at loci within or near the inversions (Hoffmann et al. 2004). The association of multiple insulin signaling loci with inversions, along with observed clines in body size, development, and inversion frequencies, have been proposed to be a part of the same adaptive strategy (De Jong & Bochdanovits 2003). *InR* is associated with the inversion *In(3R)Payne*, which contains alleles under selection (Kennington et al. 2006), is associated

with clinal variation in body size (Weeks et al. 2002, Rako et al. 2006), and contributes to additive genetic variance for other phenotypes that vary clinally (Kennington et al. 2007). Consequently, it may be difficult to distinguish between selection at *InR* and selection on *In(3R)Payne*. For example, if *InR* experiences direct selection and contributes to differences in phenotypes mediated by insulin signaling, its proximity to *In(3R)Payne* may also contribute to the observed inversion clines; alternately, selection at other loci associated with the inversion may be driving the observed distribution of *InR* alleles. However, the persistence of clinal variation among *InR* indel alleles in standard and inverted lines, the relatively stronger cline at *InR* compared to the inversion in North America, and the decay of linkage disequilibrium both 5' and 3' of the indel polymorphism suggest a good measure of independence between *InR* and *In(3R)Payne*.

Our data demonstrate that for two genes in the insulin signaling pathway, only *InR* exhibits significant patterns of polymorphism and divergence that are suggestive of adaptive evolution. The clinal pattern of the indel polymorphism and the evidence for protein evolution and positive selection at *InR* suggest that this locus has undergone adaptive evolution on both short and long timescales; the evidence for protein evolution by the McDonald-Kreitman method is signficant considering the potential for this test to underestimate adaptive responses (Charlesworth & Eyre-Walker 2008). We do not assume that the hypothetical selection pressures that may be influencing life history evolution in contemporaneous populations are necessarily relevant to or predictive of pressures that drove the amino acid evolution of *InR* over phylogenetic time. However, in the face of similar functional characterization between *InR* and *chico* by mutational genetics (Tatar et al. 2001,

Clancy et al. 2001), our consistent results for adaptive evolution at *InR* and neutral evolution at *chico* suggest that these genes may be independently constrained. These results are consistent with the hypothesis that although disruption of multiple genes within the insulin signaling pathway can lead to a reduction in signaling, the receptor molecule may be one of only a subset responsive to selection. This situation may be analogous to metabolic control theory, in which only some points can control flux through the pathway (Kacser & Burns 1973).

Characterization of other insulin signaling genes in other systems suggests that this pathway may be a fruitful research target in the identification of additional loci that contribute to additive genetic variance for life history traits. The transcription factor *dFOXO* is a promising candidate: overexpression of *dFOXO* extends lifespan (Hwangbo et al. 2004), upregulation of *dFOXO* is associated with lifespan extension by other genes in the pathway, and the *C. elegans* ortholog *daf-16* is essential for lifespan extension by insulin signaling (reviewed in Giannakou & Partridge 2007). Natural variation at the human ortholog of *dFOXO*, *FOXO3A*, contributes significantly to differences in longevity in human populations (Willcox et al. 2008, Flachsbart et al. 2008), but variation in other natural populations remains to be investigated. Polymorphism in the human *insulin-like growth factor I receptor* (*IGFIR*) is also associated with variation in human longevity (Suh et al. 2008). In teleost fishes, the *insulin-like growth factor II* (*IGF2*) shows evidence of selection responses that are coincident with the evolution of placentation (O'Neill et al. 2007). Transgenic experiments have shown that *Dp110*, which encodes the insulin-regulated phosphatidylinositol 3-kinase, affects the expression of reproductive diapause, a trait important in overwintering (Williams

et al. 2006). Diapause is variable in natural populations, shows genetic correlations to multiple life history traits (Schmidt et al. 2005a), and shows strong selective responses over both spatial (Schmidt et al. 2005b) and temporal (Schmidt & Conde 2006) scales. However, there is an absence of pronounced polymorphism between two natural *Dp110* alleles that differentially affect diapause: of 20 polymorphisms detected, none affect the amino acid sequence, and no differences in RNA levels have been detected (Williams et al. 2006). Future characterization of natural populations may yield further insight into whether this gene, or others in the insulin signaling pathway, vary significantly in the wild. Such investigations of natural genetic variation at target genes can identify genic elements that contribute to phenotype evolution, as well as elucidate important dynamics within pathways and characterize genic functionality on a fine scale. Genetic manipulations in the laboratory typically demonstrate functions of whole genes; the function of smaller regions and even specific nucleotides can be resolved by evaluating natural alleles if there is sufficient variation in the wild, where natural selection can impose subtle pressure over many generations. In our study, the putatively functional *InR* allele identifies a mutation that likely contributes to genetic variance for lifespan along with associated life history phenotypes, and provides fresh insight into the functional genetics of insulin signaling.

ACKNOWLEDGEMENTS

We thank J. Plotkin for generous assistance in analyzing linkage disequilibrium and estimating the selection coefficient. We also thank W. Eanes and T. Morgan for providing *D.*

melanogaster samples. This work was supported by an American Federation for Aging Research/Glenn Foundation Scholarship to ABP, by National Science Foundation-DEB grant 0542859 to PSS, by a Federation Fellowship to AAH, and by the Australian Research Council via their Special Research Centre Scheme.

CHAPTER FIVE

CHARACTERIZATION OF AN ADAPTIVE POLYMORPHISM IN THE *DROSOPHILA* INSULIN RECEPTOR

ABSTRACT

Previous work has identified a candidate functional polymorphism in the *Insulin-like Receptor* (*InR*) by nonrandom distribution across geography. Two alternate alleles of an amino acid indel polymorphism show parallel, independent clines in natural populations of *Drosophila melanogaster* on the North American and Australian continents. The clinal genotypic variation at *InR* co-occurs with phenotypic variation in stress tolerance, fecundity, and other life history traits, which likely represent local adaptations across the climatic gradient. We hypothesize that these alternate *InR* alleles contribute to observed genetic variance for life history phenotypes, as targets of natural selection imposed by environments that range from temperate to tropical. Here, we tested the functional significance of the two *InR* alleles on lifespan, fecundity, development time, body size and stress tolerance. We find that the alleles are associated with predictable differences in multiple traits, consistent with the hypothesis that levels of insulin signaling may mediate life history adaptation in the wild.

INTRODUCTION

The highly conserved insulin/insulin-like signaling (IIS) pathway, a primary research target in model systems, mediates growth, development, metabolism, reproduction and stress tolerance in metazoans. In particular, reduction in IIS has been shown to produce extension in lifespan, reduction in reproductive success and increased stress tolerance in flies, worms and mice (Partridge & Gems 2002; Tatar et al. 2003; Giannakou & Partridge 2007). Despite these characterizations, very little is known about how IIS pathway members vary in natural populations. A reduction in insulin signaling by manipulation at almost any point in the pathway can induce these pleiotropic effects in either *D. melanogaster* or *C. elegans*, as demonstrated by genetic manipulations of multiple elements in both systems (Giannakou & Partridge 2007). However, elements located downstream in the pathway show evidence of stronger purifying selection than upstream elements, so functional variation that persists within or among populations may be more likely to occur in the receptor than in factors mediating later steps of the signaling cascade (Alvarez-Ponce et al. 2009).

Across latitudinal clines, natural populations of *D. melanogaster* exhibit substantial variation in life history traits. At high latitudes, populations exhibit longer lifespan, lower fecundity, higher incidence of reproductive diapause, larger body size, and higher cold stress tolerance relative to low latitude populations (Capy et al. 1993; Mitrovski & Hoffmann 2001; De Jong & Bochdanovits 2003; Schmidt et al. 2005a; Trotta et al. 2006). Genetic variance for and genetic correlations among these traits indicate that selection in the local environment may act on certain phenotypes but drive expression of others through tradeoffs (David 1975;

Anderson et al. 2003; De Jong & Bochdanovits 2003; Schmidt et al. 2005b; Rako et al. 2007; Schmidt & Paaby 2008). Differences in the climatic environment co-occur with variations in multiple forms of stress resistance, and may impose selection pressures that significantly shape the evolution of life histories in these populations (Hoffmann & Harshman 1999; Hoffmann et al. 2001; Hoffmann et al. 2005; Hoffmann et al. 2007). This framework suggests a hypothetical selection regime: high latitude, seasonally cold climates impose seasonal stress and favor genotypes that confer stress tolerance; correlated traits, which may evolve as co-adapted responses to the same selection regime or by indirect selection via pleiotropy, are characterized by better overwintering ability, larger body size, longer lifespan, slower development and lower fecundity relative to low latitude populations (Paaby & Schmidt 2009).

Despite the well-characterized, hypothetically adaptive phenotypic variation in life history traits *D. melanogaster* populations exhibit across heterogeneous environments, few genes or polymorphisms have been identified that contribute to genetic variance for these phenotypes (De Luca et al. 2003; Carbone et al. 2006; Paaby & Schmidt 2008; Schmidt et al. 2008). However, genes in the IIS pathway are good candidates because the pleiotropic determination of multiple phenotypes by IIS mirrors the life history tradeoffs observed in natural populations. Previous examination of natural genetic variation at the *Insulin-like Receptor* (*InR*) in wild populations of *D. melanogaster* revealed reciprocal clines for the two most common alleles of an amino acid indel polymorphism (Paaby et al. in press). This polymorphism disrupts a region of glutamine-histidine repeats in the first exon, which encodes the N-terminal region of the protein (Guirao-Rico & Aguadé 2009). In both North

America and Australia, the *InR* "248" allele (named for its PCR fragment length) is at low frequency in tropical and subtropical populations and increases in frequency with latitude, showing highest frequency in temperate populations; the "254" allele exhibits the inverse cline. These patterns appear to be responses to similar but independent selection pressures between the continents: fly populations in North America and Australia were founded at different times and from different source populations (Bock & Parsons 1981; David & Capy 1988), and nucleotide variants on either side of the indel polymorphism show neutral patterns across geography. Cursory examination of the functional effects of these alleles showed that the 248 allele is associated with greater stress tolerance, and the 254 allele with higher fecundity (Paaby et al. in press), results consistent with our predictions about selection pressures in temperate versus tropical environments. Here, we more fully explored the phenotypic effects of these two alleles, using more robustly replicated test lines, on a spectrum of life history traits that vary in natural populations and may be mediated by IIS. We also tested whether the indel polymorphism at *InR* produces detectable differences in levels of IIS, by measuring transcript abundance for seven genes that have been shown to be direct targets of dFOXO, a central transcription factor in the pathway that is inhibited by ligand binding (Puig et al. 2003; Wang et al. 2005; Casas-Tinto et al. 2007; Flatt et al. 2008; Vihervaara & Puig 2008; Matilla et al. 2009).

MATERIALS AND METHODS

Genotyping. The *InR* indel polymorphism was genotyped by determining the length

of amplified fragments (either 248 or 254 base pairs) following polymerase chain reaction (PCR) with a fluorescent-tagged primer, using an Applied Biosystems 3100 capillary sequencer. PCR conditions, including primer sequences, are described in Paaby et al. (in press).

Fly stocks. To measure the effects of the *InR* alleles, we generated a series of test lines spanning three genotypic classes: homozygous 248, homozygous 254, and the heterozygote. Each line carried X and second chromosomes that were isogenic within and across lines, and third chromosomes for which genetic variation outside the *InR* locus was randomly distributed across lines. To generate these test lines, we selected a single 248 line and a single 254 line as parentals from stocks that were already genotyped at *InR* and were carrying isogenic third chromosomes following balancer extraction (originally described in Paaby et al. in press). The X and second chromosomes were replaced using stocks available at the Bloomington Stock Center (for the X chromosome, stock 2475; for the second chromosome, stock 6326). These two parental lines were crossed and the offspring were permitted to recombine for four generations, so that genetically variable regions from the two parental third chromosomes were distributed as randomized blocks across lines. Individual third chromosomes in the F4 generation were again extracted with balancers, each line was genotyped for the *InR* indel polymorphism, and test stocks were established. This entire scheme was replicated by generating a second set of test lines derived from a second natural population. All of the assays were performed on test lines derived from a population in Mount Sinai, NY (40.95°N, 72.84°W); a subset of the assays were also performed on test lines derived from a population in Bowdoinham, ME (44.01°N latitude, 69.90°W longitude).

Once the test stocks were established, either of two methods was used to generate the flies for the assays. In the first method, we exerted nominal control over the background variation in the test stocks by crossing 14 lines of each genotype (248 or 254) in a round-robin design to generate a total of 42 eventual lines, comprised of 14 lines of each genotypic class (two homozygotes and the heterozygote). This design produced non-independent replication within each genotypic class, and because test stocks were never mated to themselves, this design ensured that no individuals were strictly isogenic at the third chromosome. In the second method, we permitted test stocks to continue recombining freely (with individuals of the same *InR* genotype, either 248 or 254) for multiple generations. To rear flies for the assays, virgin females and males were collected and mated in fresh bottles, either within or across genotype class, to generate the two homozygote and the heterozygote genotypes for the assays.

Phenotype assays. For all phenotype assays, flies were reared and assays were performed (unless otherwise specified) on standard cornmeal-molasses media at room temperature and subject to ambient light cycles. Larval density was kept low to reduce effects of overcrowding. All statistics were performed using using JMPv7 (SAS Institute, Cary, NC). *Lifespan and lifetime fecundity*. Bottle cages were populated with 30 virgin females and 30 males, collected over 24 hours, from lines generated by method one from the Mount Sinai population. A total of 42 bottles, comprising 14 of each genotypic class, were maintained by changing standard media egg-laying plates every day (days 1-26) or every other day (days 28 onward) until all flies were dead. Bottles were inverted, so that dead flies collected on the media plates. Dead flies and eggs laid were scored at every plate change.

Lifespan was analyzed by a proportional hazards model, and cumulative fecundity was analyzed by ANOVA. *Quick fecundity*. To measure how quickly young flies lay eggs upon access to fresh food, virgin females and males of each genotypic class were collected over 6 hours, separated into single-sex vials of 5 females or 3 males each, and aged for 3 days on standard media with topical yeast. Flies were transferred in mixed-sex batches (five females and three males) to media-free vials and starved for 4 hours, then transferred again to fresh media vials and permitted to lay eggs. The number of eggs laid was recorded after 12 hours and analyzed by ANOVA. Vials were replicated 10 times within each genotype, using flies generated by method two described above. This assay was performed on test lines derived from both the Mount Sinai and Bowdoinham populations. *Development time.* Young (3-6 days old) males and females (typically 5 flies of each sex) were transferred from standard media vials to media-free vials and starved for four hours, then transferred again to fresh media vials and permitted to lay eggs, at densities of $\leq$30 eggs per vial, for up to four hours before being removed. Each of the three genotypic classes was replicated 14 times using method one described above for flies derived from the Mount Sinai population, and using method two for flies derived from the Bowdoinham population. The number of adults per vial was recorded daily at 8 AM, 12 PM and 8 PM for all emerging adults, and eclosion time was analyzed by ANOVA. *Body weight, lipid content and body size*. Using methods and replication schemes identical to those described for development time, eggs were reared and permitted to eclose as adults. After eclosion, adults were transferred to fresh media vials containing topical yeast and aged for 24 hours, then were frozen at -80°C. To measure dry weight, frozen flies were dried for 24 hours at 55°C and weighed to 0.0001 g in single-sex

batches of 7 individuals on a Sartorius scale (1-2 batches per original vial cross). The batches were transferred to 1.5 mL tubes and lipids were extracted by adding 400 μL diethyl ether and gently rocking the tubes for 24 hours. Flies were dried again overnight at 55°C and batches re-weighed to determine dry weight without lipids; lipid content was computed by subtraction. Weights were analyzed by ANOVA. To measure wing vein length as an indicator of body size, left wings of frozen flies were removed and affixed to an index card using clear tape. Images were captured using a DC 300 camera (Leica) affixed to a dissecting microscope. Veins L2 and L3 were measured using landmarks 6, 12 and 13 (Gidaszewski et al. 2009) on 3-8 male and 3-8 female wings from each original vial cross, using tpsDigv2.12 (F.J. Rohlf, 2008). Vein lengths were analyzed by ANOVA. *Chill coma recovery*. Freshly eclosed flies were collected over 24 hours from flies generated by method one from the Mount Sinai population, and sorted into standard media vials in batches of 5 males and 5 females each. After 3 days, males were sorted into fresh vials and aged another day. A total of 42 vials of 5 males each were used in the assay, comprised of 14 vials for each genotypic class. To induce chill coma, vials were completely covered in ice and placed at 4°C for three hours, then restored to room temperature. Time to recovery (transition to the upright position) was recorded using a video camera and analyzed by ANOVA. *Cold shock and starvation*. Freshly eclosed flies were collected over 24 hours from flies generated by method one from the Mount Sinai population, and sorted into standard media vials in batches of 5 males and 5 females and aged for 3 days. A total of 42 vials, comprised of 14 vials for each genotypic class, was replicated twice for each assay. To induce cold shock, vials were inverted and frozen at -20°C for 25 minutes. After 24 hours, individuals were scored as dead

or alive. To measure starvation resistance, flies were transferred into media-free vials containing a cotton ball saturated with 2 mL of water. At 51 hours, males were scored as dead or alive; at 64 hours, females were scored as dead or alive. Both assays were analyzed by nominal logistic regression modeling the log odds of mortality/survivorship. *Heat shock*. Freshly eclosed flies were collected over 24 hours and sorted into mixed-sex standard media vials of 5 males and 5 females each. Flies were generated by method two from the Mount Sinai population, and each genotypic class was comprised of 15-20 replicate vials. Flies were aged for 4 days and then transferred to fresh media vials. Vials were inverted in racks and placed in a 25°C incubator; the incubator was programmed to increase in temperature to 37°C; after reaching 37°C (about 20 minutes), flies were subject to heat shock for another 25 minutes and then removed to room temperature to recover. After 24 hours, flies were scored as dead or alive and analyzed by nominal logistic regression modeling the log odds of mortality/survivorship.

Quantitative PCR. To measure the relative levels of IIS between the three genotypic classes of the *InR* indel polymorphism, we used quantitative (real-time) polymerase chain reaction (qPCR) to determine the relative abundance of 7 transcriptional targets of dFOXO, a central transcription factor within the IIS pathway that is repressed by InR activation (Puig et al. 2003; Jünger et al. 2003). Total RNA was prepared from flies generated by method two described above, from flies derived from the Mount Sinai population. Thirty virgin females from each of the three genotypic classes were collected over 24 hours, aged on standard media, snap frozen, and processed using RNeasy (Qiagen). RNA was reverse transcribed using iScript cDNA synthesis kit (Bio-Rad), following the

manufacturer's protocol. Relative abundance of transcript levels was determined using an ABI 7500 Fast Real-Time PCR machine and SYBR Green PCR Master Mix (Applied Biosystems) by the $\Delta\Delta C_T$ relative quantitation method. Four technical replicates were used for each sample and relative abundance was normalized by using *GAPDH2* as an endogenous control. Other work has demonstrated that *GAPDH2* is an appropriate control for measuring relative levels of insulin signaling (Hwangbo et al. 2004, Flatt et al. 2008). Primer sequences for the qPCR reaction are as follows: *GAPDH2* forward, GCGGTAGAATGGGGTGAGAC; *GAPDH2* reverse, TGAAGAGC GAAAACAGTAGC; *4E-BP* forward, GAAGGTTGTCAT CTCGGATC; *4E-BP* reverse, ATGAA AGCCCGCTCGTAGA; l*(2)efl* forward, AGGGACGATGTGACCGTGTC; *l(2)efl* reverse, CGAAGCAGACGCGTTTATCC; *ac76e* forward, CAGGATGAATGACGCCCTTTCGG; *ac76e* reverse, ATGGACACAACACATG CCAGCAGC; *dLip4* forward, GATAGCAATGTGCGGTTG GA; *dLip4* reverse, TCATCCGTCTCCAAGGTGTG; *InR T1* forward, CACAAGCTGGAAAGA AAGTGC; *InR T1* reverse, CAAACACGTTTCGATAATATTTTTCT; *InR T2* forward, GCCTCGCACTTTGCTTATGT; *InR T2* reverse, AAAAACAACGACAGCGACAA; *InR T3* forward, TTACGCCACTGCATTCGTTC; *InR T3* reverse, ATGGCCTCTCTCTCCGTCTC.

RESULTS

Flies carrying the 248 allele, which is common in temperate, high-latitude environments, showed phenotypic differences that are consistent with canonical life history

trade-offs and an environmentally-imposed selection regime when compared to flies carrying the tropical 254 allele. The 248 flies better tolerated multiple forms of stress and showed some evidence of longer lifespan, while the 254 flies developed faster and were larger and more fecund. Heterozygotes typically exhibited highest fitness, or fitness equivalent to the fitter of the other two genotypes; these high fitness outcomes may be a result of heterosis at other loci. The phenotypic differences between the genotypes are likely determined by different levels of insulin signaling: the 248 allele appears to mediate lower levels of signaling than the 254 allele, as transcript abundance of multiple targets of the transcription factor dFOXO, which is repressed by InR activity, was highest in 248 genotypes. For five of seven of these transcriptional targets, the 248 genotype showed approximately twofold higher mRNA abundance than the 254 genotype, while the heterozygote showed intermediate abundance.

Lifespan. The 248 and 254 genotypes show little difference in rates of aging over the majority of lifespan, and there was no significant difference in survivorship between 248 and 254 females ($p=0.89$) (Figure A, Table 1). However, 248 males exhibited a slowing in aging late in life relative to the 254 males, producing a significantly higher rate of survivorship ($p<0.01$) (Figure 1B, Table 1). For both sexes, the heterozygotes were significantly longer lived ($p<0.01$) (Figure 1, Table 1).

Fecundity. Surprisingly, the *InR* genotypes did not show significant differences in total fecundity: a fly carrying the 254 allele laid on average only 5.3 more eggs over her lifetime than a fly carrying the 248 allele ($p=0.69$, Table 1). Heterozygotes laid the most eggs, but there were no significant differences among the genotypes under the whole model

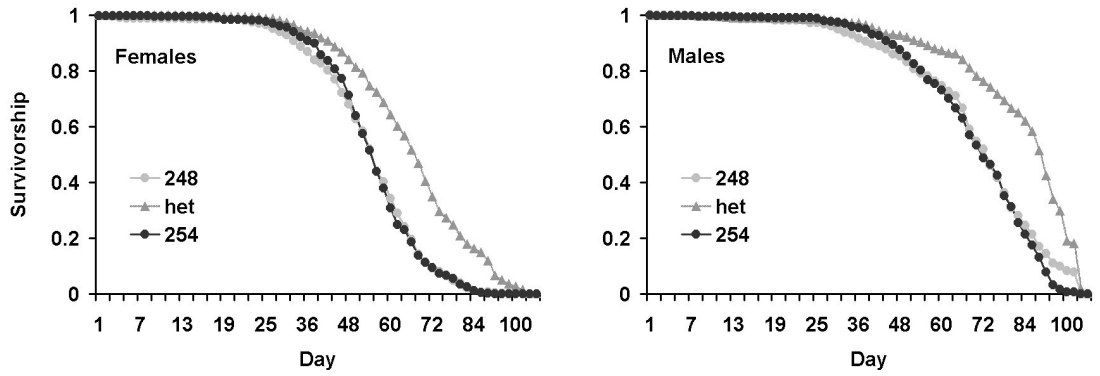


FIGURE 1. Average survivorship curves for the 248, heterozygote and 254 genotypes, from flies derived from the Bowdoinham population.

(Table 1). This was unexpected because we routinely observed substantially higher population numbers in our 254 bottle cultures, and previous results showed that the 254 genotype was more fecund (Paaby et al. in press). Examination of the number of eggs laid per day, however, reveals that the 254 genotype exhibits higher fecundity at earlier ages (Figure 2). We performed another assay to explicitly measure "quick" fecundity, or whether there was a difference in how many eggs young flies laid upon first access to fresh food. Here, we found that the 254 genotype is nearly six times more fecund than the 248 genotype in the first 12 hours of egg-laying ($p<0.01$, Table 1, Figure 2). As in the lifetime fecundity assay, the heterozygotes also laid the most eggs ($p<0.01$) (Table 1, Figure 2B). The high fecundity of the 254 genotype was also observed in a replicate assay using flies derived from the second population ($p<0.01$), although in this test the heterozygotes showed intermediate fecundity (Table 1, Figure 2).

Development time. From egg to emerging adult, flies derived from two separate

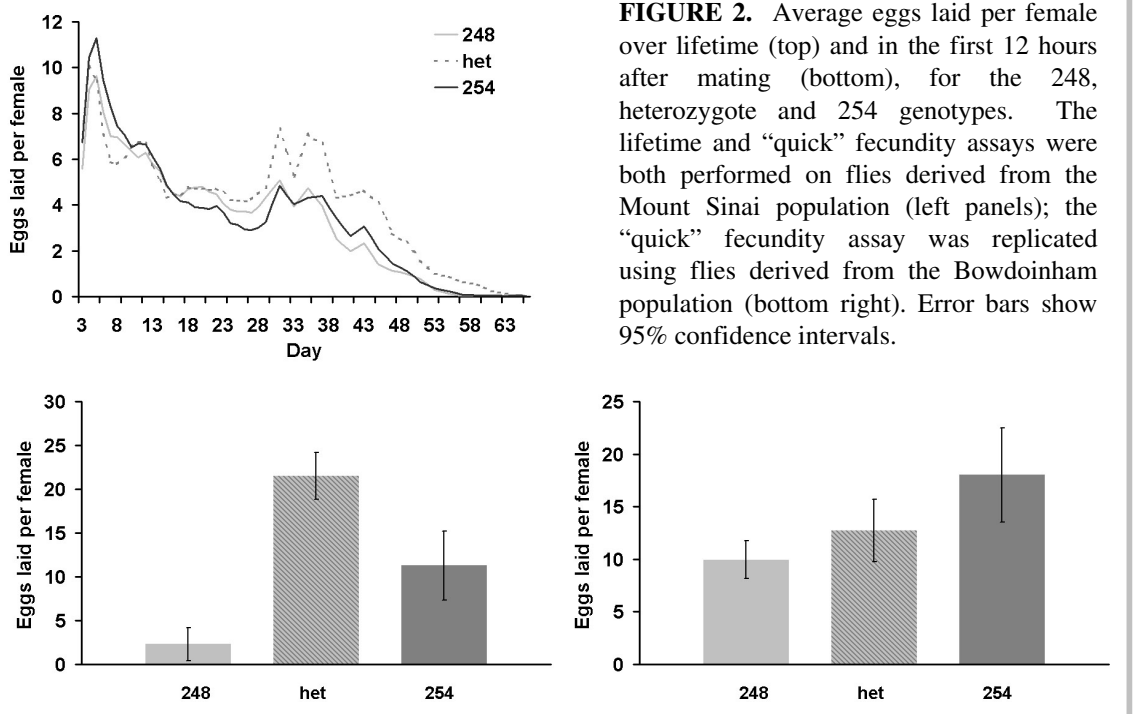


FIGURE 2. Average eggs laid per female over lifetime (top) and in the first 12 hours after mating (bottom), for the 248, heterozygote and 254 genotypes. The lifetime and "quick" fecundity assays were both performed on flies derived from the Mount Sinai population (left panels); the "quick" fecundity assay was replicated using flies derived from the Bowdoinham population (bottom right). Error bars show 95% confidence intervals.

populations demonstrated fastest development in the 254 genotype. In the first population, males and females carrying the 254 allele developed an average of 6.0 hours faster than the 248 allele ($p<0.01$); heterozygotes developed even faster, at an average of 7.1 hours ahead

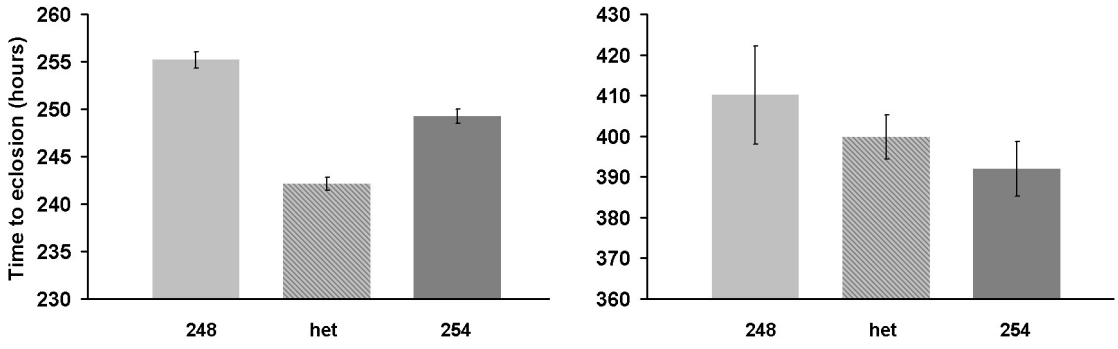


FIGURE 3. Average development time for the 248, heterozygote and 254 genotypes, in flies derived from the Mount Sinai (left) and Bowdoinham (right) populations. Error bars show 95% confidence intervals.

of the 254 homozygotes ($p<0.01$). In the second population, development time was not significantly different among male genotypes, but 254 females emerged an average of 18.2 hours ahead of 248 females ($p<0.01$); heterozygotes showed intermediate development time (Table 1, Figure 3).

Body weight and size. Because decreases in insulin signaling reduce body size (Clancy et al. 2001; Tatar et al. 2001), we predicted that individuals carrying the 254 *InR* allele would be larger than individuals carrying the 248 allele. Indeed, the average dry body weight of 254 females and males was 15.5% and 10.9% heavier ($p<0.01$ for both) than the

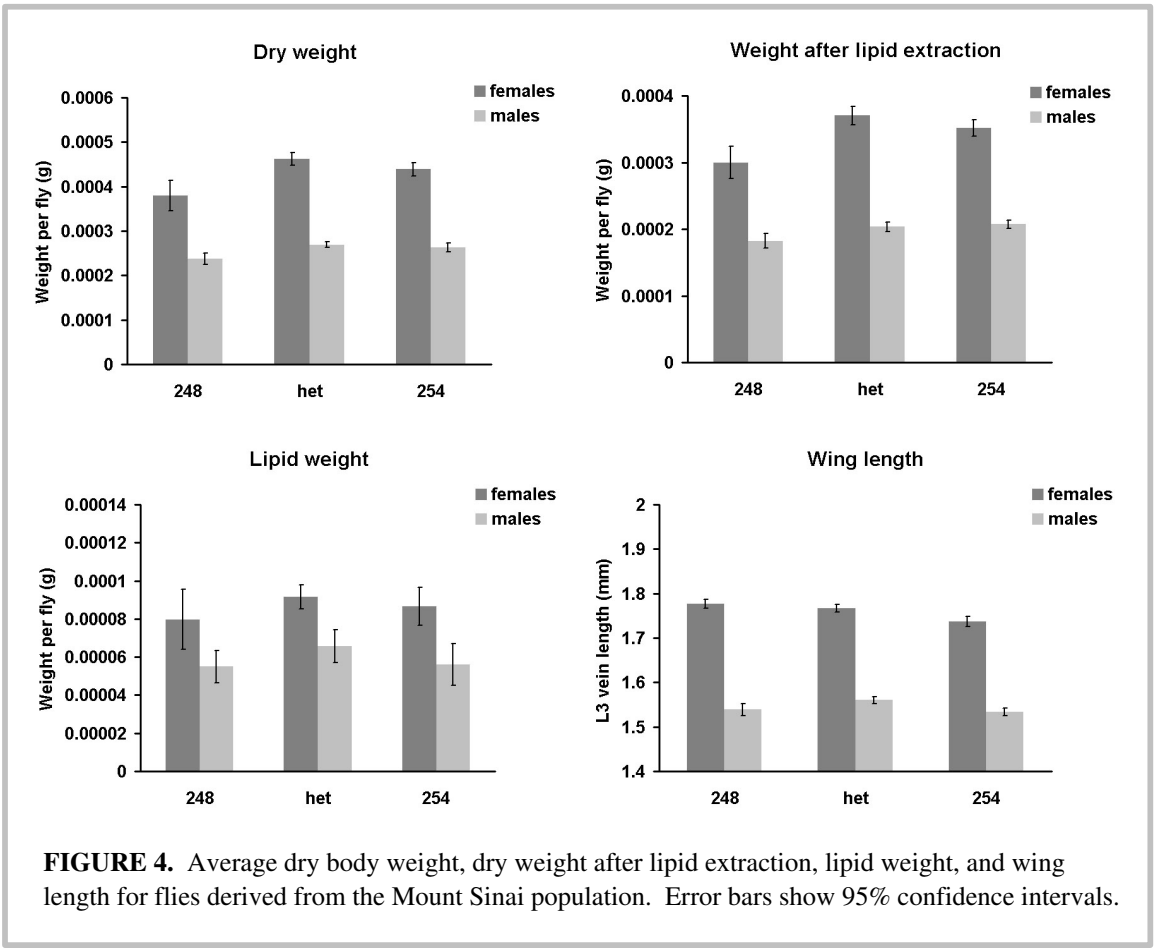


FIGURE 4. Average dry body weight, dry weight after lipid extraction, lipid weight, and wing length for flies derived from the Mount Sinai population. Error bars show 95% confidence intervals.

dry weight of 248 females and males derived from the Mount Sinai population (Table 2, Figure 4). Likewise, average dry weight of 254 females and males was 26.1% and 18.6% heavier ($p<0.01$ for both) than the weight of 248 females and males derived from the Bowdoinham population (Table 3, Figure 5). For both sexes in both populations, the 254 genotype was also significantly heavier than the 248 genotype after lipid extraction (Tables 2 and 3, Figures 4 and 5). In all comparisons, the heterozygote genotype was heaviest, or equally heavy to the 254 genotype. Total lipid content was not significantly different among genotypes from the Mount Sinai population, and was less significantly

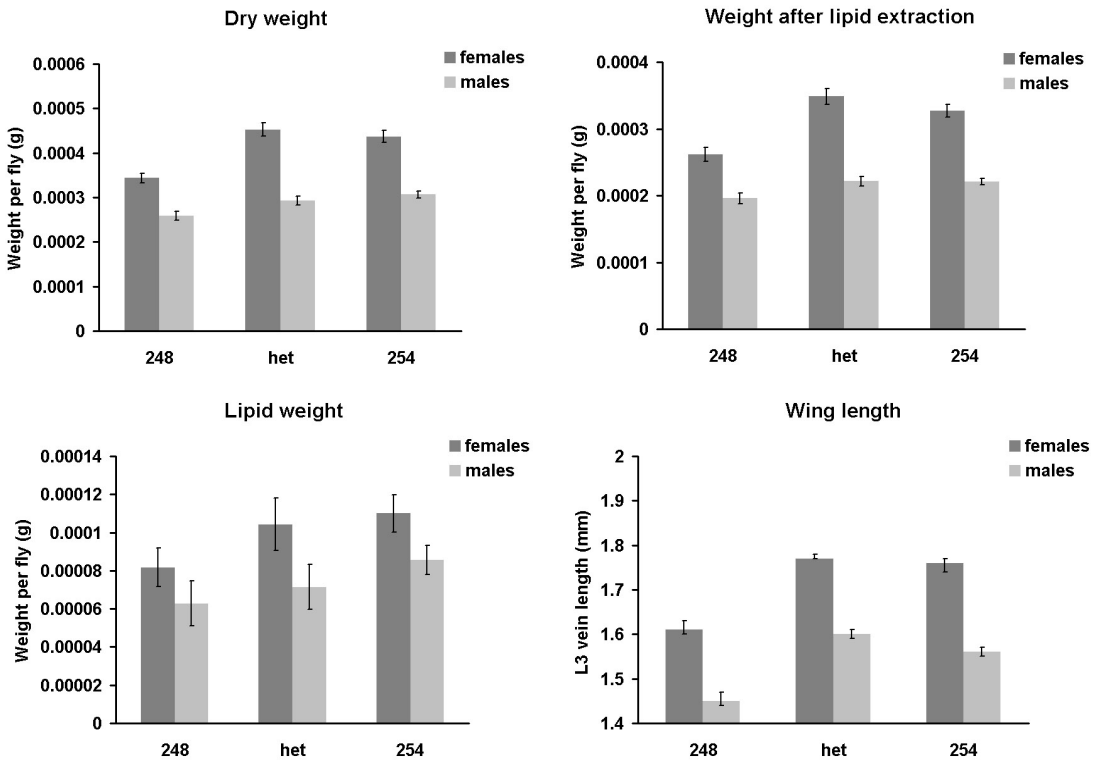


FIGURE 5. Average dry body weight, dry weight after lipid extraction, lipid weight, and wing length for flies derived from the Bowdoinham population. Error bars show 95% confidence intervals.

different among genotypes from the Bowdoinham population, indicating that relative lipid content differs among the lines (Tables 2 and 3, Figures 4 and 5). Although the 254 genotype was heavier than the 248 genotype in all comparisons, implicating larger body size, the L3 wing vein lengths of 254 flies were only longer than the wing vein lengths of 248 flies in the Bowdoinham population ($p<0.01$ for females and males) (Table 3, Figure 5). In flies derived from the Mount Sinai population, the reverse was true ($p<0.01$ for both) (Table 2, Figure 4). In all comparisons, the heterozygotes exhibited longest vein length, or length equal to the 254 genotype. These patterns were qualitatively identical using measurements from a shorter longitudinal vein, L2 (data not shown).

Stress tolerance. Reduction in insulin signaling confers increased resistance to stress (Giannakou & Partridge 2007), and our results for three stress tolerance assays are consistent with these findings (Figure 6). Flies carrying the 248 allele recovered faster from chill coma than did flies carrying the 254 allele ($p<0.01$), and the heterozygotes were not statistically distinct from the 254 homozygotes (Table 4). Likewise, the 248 genotypes survived significantly better than the 254 or heterozygote genotypes after cold shock: 254 females showed 3.8 times the odds of dying and heterozygous females showed 3.7 times the odds of dying relative to 248 females; 254 and heterozygous males showed 2.7 and 2.0 times the odds of dying, respectively, relative to 248 males (Table 4). Male resistance to starvation was consistent with this pattern as well: 254 and heterozygous males showed 1.8 and 1.9 times the odds of dying, respectively, relative to the 248 genotype. In females, the 254 genotype was weaker at resisting starvation than the 248 allele, showing 1.3 times the odds of dying. However, in females, heterozygotes exhibited strongest starvation resistance, as the

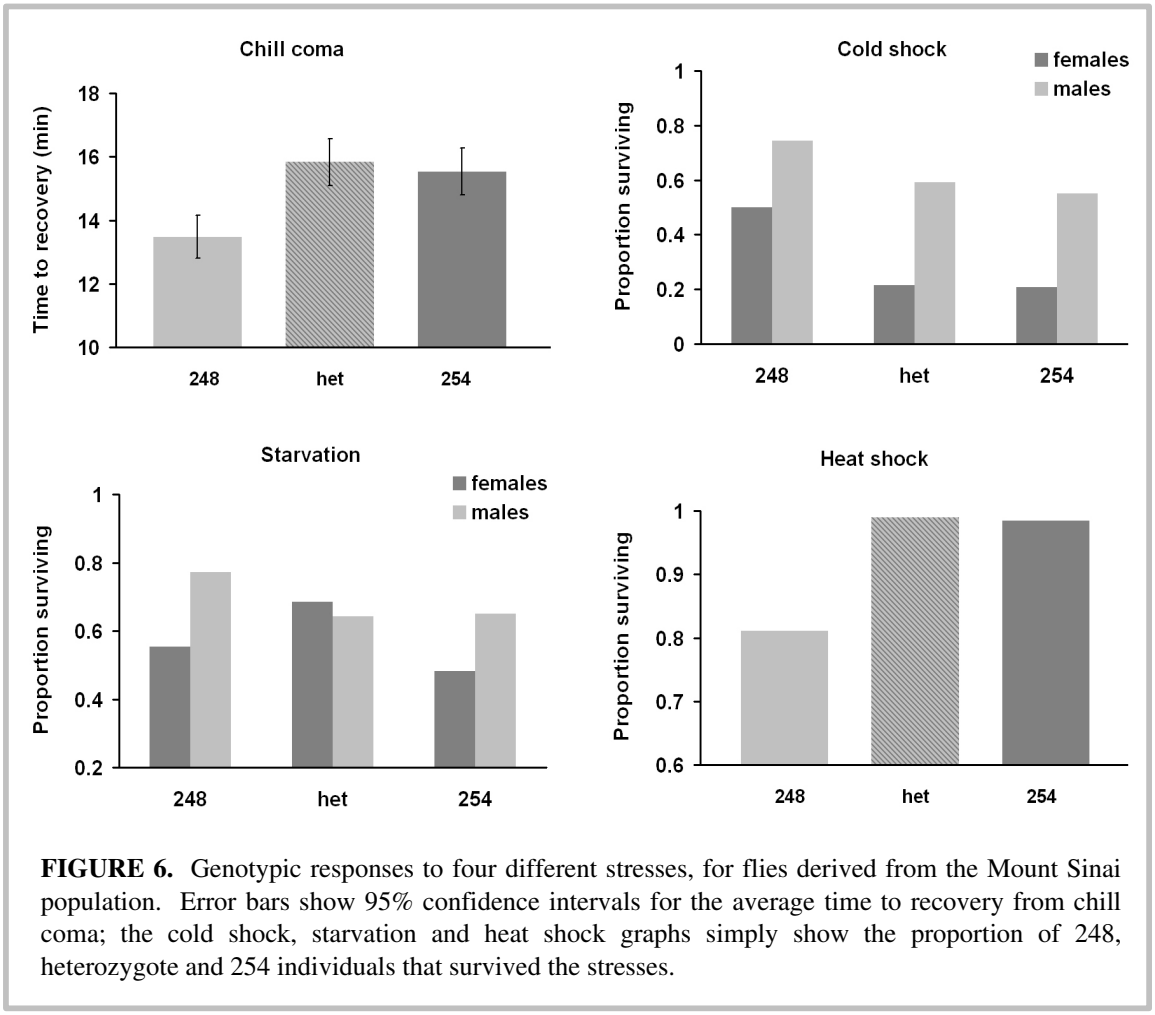


FIGURE 6. Genotypic responses to four different stresses, for flies derived from the Mount Sinai population. Error bars show 95% confidence intervals for the average time to recovery from chill coma; the cold shock, starvation and heat shock graphs simply show the proportion of 248, heterozygote and 254 individuals that survived the stresses.

248 genotype showed 1.8 times the odds of dying relative to the heterozygote (Table 4). The ability of these genotypes to tolerate heat stress was the reverse of the pattern shown for the first three stresses: the 248 genotype showed far higher susceptibility to dying of heat shock, exhibiting 15.3 and 21.7 times the odds of dying relative to the 254 and heterozygote genotypes, respectively, with no differences between the sexes (Table 4). This pattern is consistent with other examples of reciprocal tolerance for environmentally imposed thermal stresses (Hoffmann et al. 2002; Hoffmann et al. 2005), and is also consistent with the hypothetical selection regime that may be maintaining the *InR* polymorphism across

geography.

Levels of insulin signaling. The phenotypic differences between the *InR* genotypes may be mediated by different levels of insulin signaling, as five of seven dFOXO transcriptional targets, which should should exhibit increased expression under reduced IIS activity, showed predictably higher abundance in flies carrying the 248 allele. Transcript levels of 4E-BP, l(2)efl, ac76e, dLip4 and InR T1 were lowest in the 254 allele, intermediate in the heterozygote, and highest in the 248 allele (Figure 7); all but 4E-BP showed statistical significance between the two homozygous genotypes. This consistent pattern of transcript

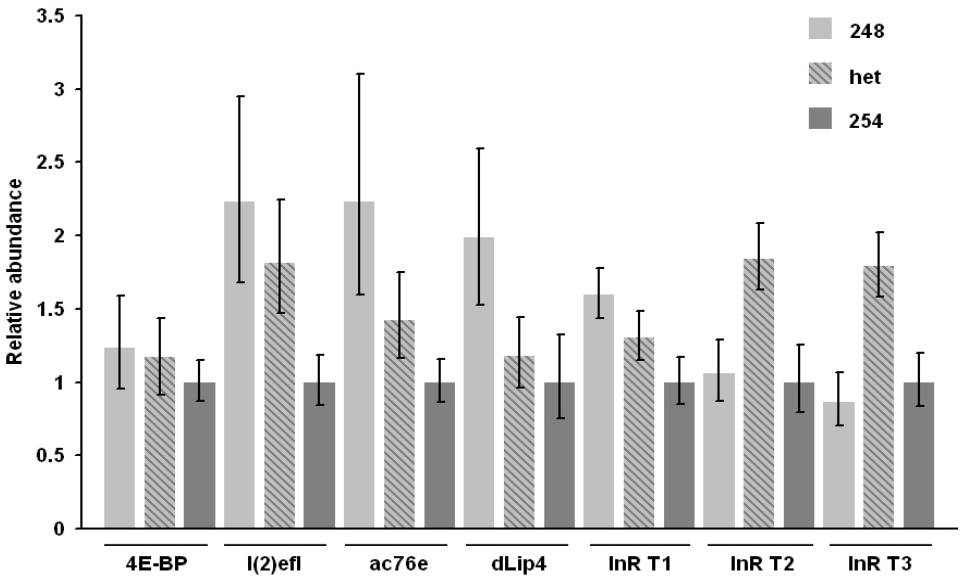


FIGURE 7. Relative abundance of seven transcriptional targets of dFOXO, a central transcription factor in the IIS pathway that is repressed by InR activity. Transcripts 4E-BP, l(2)efl, ac76e, dLip4 and InR T1 all show higher abundance in the 248 genotype, indicating reduced insulin signaling, and intermediate abundance in the heterozygote relative to the 254 genotype (though 4E-BP is not statistically significant). Transcripts InR T2 and InR T3, however, depart from this pattern and exhibit significantly higher abundance in the heterozygote genotype than in either of the homozygotes. Transcripts were measured from flies derived from the Mount Sinai population. Error bars show 95% confidence intervals.

abundance across genotypes is especially encouraging considering that the 248 genotype exhibits only about twofold higher levels relative to the 254 genotype—a difference that may be biologically meaningful, but close to the limit of detection by qPCR methods. Two other targets, transcripts of *InR* itself, show no difference in abundance level between the homozygote genotypes, but significantly higher abundance in the heterozygote (Figure 7).

TABLE 1. Statistical results for the fecundity, lifespan and development time assays.

Lifespan: Proportional hazards model effect likelihood ratio tests (Mount Sinai)				
Source	DF		χ^2	p
genotype (f)	2		155.20	<0.0001***
248 × 254 (f)	1		0.02	0.89
genotype (m)	2		165.85	<0.0001***
248 × 254 (m)	1		6.73	0.0095**
Lifetime fecundity: ANOVA (Mount Sinai)				
Source	DF	SS	F	p
genotype	2	6160.87	2.40	0.10
p.c. 248 × 254	1, 39	200.18	0.16	0.69
p.c. 248 × het	1, 39	5466.55	4.26	0.05*
p.c. het × 254	1, 39	3574.57	2.79	0.10
Quick fecundity: ANOVA (Mount Sinai)				
Source	DF	SS	F	p
genotype	2	1469.05	55.85	<0.0001***
p.c. 248 × 254	1, 21	299.77	22.79	0.0001***
p.c. 248 × het	1, 21	1451.04	110.31	<0.0001***
p.c. het × 254	1, 21	443.77	33.74	<0.0001***
Quick fecundity: ANOVA (Bowdoinham)				
Source	DF	SS	F	p
genotype	2	335.23	8.02	0.0018**
p.c. 248 × 254	1, 27	324.82	15.55	0.0005***
p.c. 248 × het	1, 27	38.64	1.85	0.19
p.c. het × 254	1, 27	139.39	6.67	0.02*

Development time: ANOVA (Mount Sinai)

Source	DF	SS	F	p
genotype	2	81440.13	270.48	<0.0001***
sex	1	271.65	1.81	0.18
genotype × sex	2	245.64	0.82	0.44
p.c. 248 × 254	1, 3006	17049.77	113.23	<0.0001***
p.c. 248 × het	1, 3006	80455.71	534.30	<0.0001***
p.c. het × 254	1, 3006	27147.15	180.28	<0.0001***

Development time: ANOVA (Bowdoinham)

Source	DF	SS	F	p
genotype	2	2742.59	1.24	0.29
sex	1	2449.16	2.22	0.14
genotype × sex	2	9928.10	4.50	0.01**
p.c. 248 × 254 (f)	1, 244	9383.14	9.07	0.0029**
p.c. 248 × het (f)	1, 244	3402.43	3.29	0.07
p.c. het × 254 (f)	1, 244	3029.58	2.93	0.09
p.c. 248 × 254 (m)	1, 278	1652.35	1.42	0.23
p.c. 248 × het (m)	1, 278	2163.43	1.86	0.17
p.c. het × 254 (m)	1, 278	30.21	0.03	0.87

TABLE 2. Statistical results for the body weight, lipid content and body size assays (Mount Sinai).

Dry body weight: ANOVA					
Source		DF	SS	F	p
genotype	(f)	2	2.57e-6	19.51	<0.0001***
	(m)	2	4.14e-7	14.00	<0.0001***
p.c. 248 × 254	(f)	1, 45	1.10e-6	16.73	0.0002***
	(m)	1, 44	2.22e-7	14.99	0.0004***
p.c. 248 × het	(f)	1, 45	2.55e-6	38.75	<0.0001***
	(m)	1, 44	3.93e-7	26.61	<0.0001***
p.c. het × 254	(f)	1, 45	2.20e-7	3.34	0.07
	(m)	1, 44	1.46e-8	0.99	0.33
Dry body weight after lipid extraction: ANOVA					
Source		DF	SS	F	p
genotype	(f)	2	1.89e-6	21.61	<0.0001***
	(m)	2	2.41e-7	11.15	0.0001***
p.c. 248 × 254	(f)	1, 45	8.57e-7	19.58	<0.0001***
	(m)	1, 44	2.02e-7	16.68	<0.0001***
p.c. 248 × het	(f)	1, 45	1.87e-6	42.74	<0.0001***
	(m)	1, 44	1.73e-7	15.98	0.0002***
p.c. het × 254	(f)	1, 45	1.37e-7	3.13	0.08
	(m)	1, 44	5.15e-9	0.48	0.49
Lipid content: ANOVA					
Source		DF	SS	F	p
genotype	(f)	2	5.32e-8	1.63	0.21
	(m)	2	5.86e-8	1.97	0.15
p.c. 248 × 254	(f)	1, 45	1.54e-8	0.95	0.34
	(m)	1, 44	4.58e-10	0.03	0.86
p.c. 248 × het	(f)	1, 45	5.29e-8	3.26	0.08
	(m)	1, 44	4.48e-8	3.01	0.09
p.c. het × 254	(f)	1, 45	9.75e-9	0.60	0.44
	(m)	1, 44	3.71e-8	2.50	0.12
Body size by L3 wing vein length: ANOVA					
Source		DF	SS	F	p
genotype	(f)	2	0.0573	16.29	<0.0001***
	(m)	2	0.0320	9.15	0.0002***
p.c. 248 × 254	(f)	1, 230	0.0504	28.63	<0.0001***
	(m)	1, 218	9.41e-4	0.54	0.46
p.c. 248 × het	(f)	1, 230	0.0040	2.27	0.13
	(m)	1, 218	0.0163	9.32	0.0025**
p.c. het × 254	(f)	1, 230	0.0370	21.02	<0.0001***
	(m)	1, 218	0.0269	15.42	0.0001***

TABLE 3. Statistical results for the body weight, lipid content and body size assays (Bowdoinham).

Dry body weight: ANOVA

Source		DF	SS	F	p
genotype	(f)	2	2.21e-6	80.95	<0.0001***
	(m)	2	4.28e-7	33.52	<0.0001***
p.c. 248 × 254	(f)	1, 37	1.42e-6	103.96	<0.0001***
	(m)	1, 39	4.03e-7	63.17	<0.0001***
p.c. 248 × het	(f)	1, 37	1.94e-6	141.70	<0.0001***
	(m)	1, 39	2.06e-7	32.23	<0.0001***
p.c. het × 254	(f)	1, 37	4.32e-8	3.16	0.08
	(m)	1, 39	3.29e-8	5.16	0.03*

Dry body weight after lipid extraction: ANOVA

Source		DF	SS	F	p
genotype	(f)	2	1.29e-6	81.32	<0.0001***
	(m)	2	1.52e-7	22.32	0.0001***
p.c. 248 × 254	(f)	1, 37	6.93e-7	86.69	<0.0001***
	(m)	1, 39	1.12e-7	32.91	<0.0001***
p.c. 248 × het	(f)	1, 37	1.22e-6	152.71	<0.0001***
	(m)	1, 39	1.16e-7	34.03	<0.0001***
p.c. het × 254	(f)	1, 37	8.04e-8	10.06	0.0030**
	(m)	1, 39	3.2e-11	<0.01	0.92

Lipid content: ANOVA

Source		DF	SS	F	p
genotype	(f)	2	1.42e-7	7.56	0.0018**
	(m)	2	9.21e-8	5.46	0.0081**
p.c. 248 × 254	(f)	1, 37	1.38e-7	13.80	0.0007***
	(m)	1, 39	9.03e-8	10.70	0.0022**
p.c. 248 × het	(f)	1, 37	8.26e-8	8.79	0.0053**
	(m)	1, 39	1.29e-8	1.52	0.22
p.c. het × 254	(f)	1, 37	5.71e-9	0.61	0.44
	(m)	1, 39	3.50e-8	4.15	0.05*

Body size by L3 wing vein length: ANOVA

Source		DF	SS	F	p
genotype	(f)	2	0.5989	210.98	<0.0001***
	(m)	2	0.4775	176.68	<0.0001***
p.c. 248 × 254	(f)	1, 117	0.4080	287.45	<0.0001***
	(m)	1, 123	0.2519	186.41	<0.0001***
p.c. 248 × het	(f)	1, 117	0.5071	357.29	<0.0001***
	(m)	1, 123	0.4386	324.59	<0.0001***
p.c. het × 254	(f)	1, 117	0.0058	4.11	0.04*
	(m)	1, 123	0.0257	19.04	<0.0001***

TABLE 4. Statistical results for the stress assays (Mount Sinai).

Chill coma: ANOVA				
Source	DF	SS	F	p
genotype	2	829073.78	12.66	<0.0001***
p.c. 248 × 254	1, 207	534075.78	16.31	<0.0001***
p.c. 248 × het	1, 207	698516.78	21.33	<0.0001***
p.c. het × 254	1, 207	11018.31	0.34	0.56
Cold shock: Effect likelihood ratio tests				
Source	DF		χ^2	p
genotype	2		45.26	<0.0001***
sex	1		93.80	<0.0001***
genotype × sex	2		3.05	0.22
genotype (f)	2		35.70	<0.0001***
(m)	2		12.64	0.0018**
Starvation: Effect likelihood ratio tests				
Source	DF		χ^2	p
genotype (f)	2		12.36	0.0021**
(m)	2		7.03	0.0298*
Heat shock: Effect likelihood ratio tests				
Source	DF		χ^2	p
genotype	2		55.01	<0.0001***
sex	1		0.05	0.83
genotype × sex	2		0.65	0.72

DISCUSSION

Our results are consistent with the hypothetical selection regime that may be maintaining life history variation across geography (Paaby & Schmidt 2009). The *InR* allele common at high latitudes, where seasonal stress may be imposed by colder temperatures and

ephemeral food resources, is associated with better tolerance of cold shock, chill coma and starvation; the allele common at warmer low latitudes, where interspecific competition may select for fast reproduction, is associated with stronger resistance to heat stress, higher fecundity and faster development time. We conclude that the *InR* indel polymorphism is adaptive, and contributes to the observed variations in life history phenotypes by differentially mediating levels of insulin signaling.

Interestingly, the alleles appear to contribute only subtle differences to lifespan. We detected no significant differences in survivorship among females, and the differences in male longevity were the result of changes in the rate of aging very late in life. Likewise, the differences in fecundity among the alleles were nuanced, with no significant effect on the number of eggs laid over lifetime. Given the role of insulin signaling in determining lifespan and fecundity (Clancy et al. 2001; Tatar et al. 2001), we expected the allelic differences to be more robust in these gross traits. However, phenotypic differences among naturally occurring alleles may be hardest to detect in the most quantitative traits. Our results emphasize the value of defining phenotypes carefully when characterizing phenotypic tradeoffs. For example, the highly significant difference in egg-laying immediately after mating may have substantial fitness consequences in the wild, and yet it is undetectable in assays that measure fecundity over several days. Consequently, estimates of total lifetime fecundity, which may be less adaptively meaningful, may also obscure functional relationships in evaluations that seek to characterize genotypic and phenotypic determinants of life history.

We expected that the allele mediating higher levels of IIS, 254, would be associated

with larger body size, based on observations of substantial decreases in body size after reduction of IIS by mutational genetics (Giannakou & Partridge 2007). This was borne out by the heavier body weight observed in the 254 genotypes, but not by the shorter wing length observed in the Mount Sinai population. However, allometric relationships between wing length and thorax width may govern the determination of wing size in a way too complex for our assay to parse. The wing/thorax ratio is negatively correlated with wing loading, and low wing load may be an adaptation to flying in cold temperatures (Stalker 1980; Pétavy et al. 1997); wing/thorax ratios are highest in temperate environments, and decrease with latitude (Capy et al. 1993; Capy et al. 1994). The larger wing size observed in the Mount Sinai 248 genotypes is probably associated with a narrower thorax and an overall smaller body size, since these 248 flies weighed less, but we did not measure thorax width. Furthermore, the expectation that the temperate, high latitude 248 allele from the Mount Sinai population confers smaller body size conflicts with body size clines, which routinely show increasing size with latitude (Capy et al. 1993; Capy et al. 1994; Karan et al. 1998; Hoffmann et al. 2001). Such phenotypic variation is likely adaptive, as colder temperatures have induced genetic responses for smaller body size under laboratory conditions (Anderson et al. 1973; Powell et al. 1974; Cavicchi et al. 1985; Partridge et al. 1994). Given the strength of the allelic responses to stress, we hypothesize that *InR* genotypes that confer better stress resistance but smaller body size still experience positive selection at high latitudes. Nevertheless, the highly pleiotropic nature of IIS presents an interesting case in which functional nucleotides may be inversely distributed according to fitness, deserving of future exploration.

We also predicted that 248 flies would have higher lipid content, since IIS mutants, while smaller, show much higher lipid levels (Böhni et al. 1999; Brogiolo et al. 2001; Clancy et al. 2001; Tatar et al. 2001; Hwangbo et al. 2004). Indeed, the equivalency in absolute lipid content between the alleles demonstrates that the smaller 248 genotypes carry higher lipid content relative to the 254 genotypes. This outcome is consistent with the demonstrated tradeoff between starvation resistance and early reproductive success (Wayne et al. 2006), since our starvation-resistant 248 females showed markedly lower early fecundity relative to the 254 females. However, other studies have demonstrated direct links between starvation resistance and fat content (David et al. 1975; Service at al. 1987; Zwaan et al. 1991), and higher starvation resistance in tropical populations relative to temperate populations (Da Lage et al. 1990; Karan et al. 1998). Consequently, the association of the high lipid, high latitude 248 allele with increased starvation resistance in our study is at odds with phenotypic variation for starvation resistance observed elsewhere. However, reports on these phenotypic associations are inconsistent, and imply complexities that are not yet understood (Hoffmann & Harshman 1999). One step towards finer resolution of this problem would be more complete characterization of geographical variation in lipid content.

Likewise, our observation that the low latitude 254 allele is associated with faster development time is inconsistent with reports of clinal variation in this trait. Positive phenotypic correlations have been demonstrated for starvation resistance, slow development and reduced early age reproduction (Hoffmann & Harshman 1999); in our study, this suite of life history traits is consistent with phenotypes associated with the high latitude 248 allele. However, like starvation resistance, faster development time has been observed in tropical

populations relative to temperate populations (James and Partridge 1995; van't Land 1999). We have hypothesized previously that faster development may be a high fitness strategy for tropical populations with more consistent access to food resources, assuming incorruptible tradeoffs with cold and starvation tolerance (Paaby & Schmidt 2009). If true, the *InR* indel polymorphism would be expected to experience positive selection on each of these traits within each local environment. Regardless, further investigations are required to more fully describe the physiological correlations and selection pressures that shape the life histories of natural populations.

Although we sought to evaluate dominance dynamics of the *InR* indel polymorphism, many of our phenotypic assays revealed significant heterosis. In particular, flies derived from the Mount Sinai population showed consistent heterosis. The higher fitness of the heterozygotes is probably not due to intermediate levels of IIS, as heterosis was most evident in the most quantitative traits, like lifespan and fecundity, and least evident in the least quantitative traits, like body weight and transcript abundance. This suggests that contributions from other loci may be more strongly affecting phenotypes in which the heterozygotes are fittest. In principle, randomization of the naturally occurring genetic variation on the remainder of the third chromosome should have made the lines equivalent everywhere but the target locus. However, the genome regions close to *InR* almost certainly contained blocks in linkage disequilibrium with the indel polymorphism. Since we selected test lines based on the *InR* allele identity, the 248/254 genotypes would have carried higher heterozygosity in the regions still in disequilibrium with the indel, which likely accounts for the heterosis observed in our assays. The consistently intermediate phenotypes demonstrated

by the heterozygotes derived from the Bowdoinham population probably reflect variation between the source populations at loci neighboring *InR*. Consequently, the recombinant scheme we employed is not sufficient to test for dominance between the alleles. Likewise, we cannot definitively claim that the phenotypic differences between the homozygote genotypes are directly mediated by the *InR* indel polymorphism, since other, unknown alleles must also be associated with these phenotypes.

ACKNOWLEDGEMENTS

Several people generously assisted in the completion of this work. We thank Emily Behrman and Katherine O'Brien for measuring the fly wings, Thomas Flatt for advice, and Li Yang, Nancy Bonini and the members of her lab for sharing their expertise and resources for qPCR.

CHAPTER SIX

DISSECTING THE GENETICS OF LONGEVITY IN *DROSOPHILA MELANOGASTER*†

ABSTRACT

Drosophila melanogaster has been an historically important system for investigating the genetic basis of longevity, and will continue to be valuable as new technologies permit genomic explorations into the biology of aging. The utility of *D. melanogaster* resides in two resources: its powerful genetic tools as a model system, and a natural ecology that provides substantial genetic variation across significant environmental heterogeneity. Here we provide a review of the genetics of longevity in *D. melanogaster*, in which we describe the characterization of individual aging genes, the complexity of the genetic architecture of this quantitative trait, and the evaluation of natural genetic variation in the evolution of life histories.

INTRODUCTION

Organism longevity is a quantitative trait determined by both environmental and genetic components. *Drosophila melanogaster* has proved one of the most useful model

† This chapter has been published as a journal article: Paaby, A.B. and P.S. Schmidt. 2009. Dissecting the genetics of longevity in *Drosophila melanogaster*. *Fly* 3(1): 29-38.

systems for exploring the genetic determination of lifespan, both by identification of candidate aging genes by classical genetics approaches, and by characterization of the contribution of natural genetic variation to longevity phenotypes, artificial selection responses, and natural selection responses. The laboratory lifespan of *D. melanogaster* is on the order of eight weeks, and is highly responsive to manipulations like induced mutations or artificial selection regimes—long-lived strains can show twice the lifespan of short-lived strains. But despite the tractability of this system to experimentation, a few critical questions regarding the genetics of aging remain largely unanswered. Is the genetic determination of longevity principally governed by many genes of small effect, or by a few genes of large effect? And are the aging genes discovered by mutational analyses the same genes that contribute to differences in longevity phenotypes in natural populations?

Although these questions remain, work in *D. melanogaster* has led to substantial advances in our understanding of the biology of aging. Repeated identification of candidate aging genes by independent approaches demonstrates the efficacy of these methods and suggests that comprehensive characterization of the most important genetic determinants is possible. Furthermore, the rapid development of genomic techniques will facilitate exploration of the complex genetic architecture of lifespan, which, as a highly quantitative trait, can only be fully understood on a genome-wide scale. Identification of specific aging genes has also permitted evaluation of how these loci contribute to the observed genetic variance for lifespan in natural populations. Comparison of the subset of genes shown to affect natural variation in lifespan to the subset of genes shown to extend lifespan by genetic manipulation will yield critical insights into the utility of these different approaches to

characterize the genetics of lifespan. For example, a gene that extends lifespan by laboratory induced mutation might regulate aging, but not necessarily contribute to observed differences in lifespan in the wild: such a gene may be under strong natural selection constraint and vary little at the nucleotide level; it may harbor substantial allelic variation that does not affect the phenotype and is segregating neutrally; or, alternately, it may harbor functionally significant allelic variation that responds to natural selection. Characterizing the natural genetic variation that affects lifespan can reveal subtleties in the mechanisms of lifespan mediation, and will help us better understand how mutational analyses identify and describe genetic determination of quantitative traits.

In addition to dissecting the genetics of longevity, *D. melanogaster* has been invaluable in exploring the physiology of aging (Kirkwood and Austad 2000; Partridge and Gems 2002; Helfand and Rogina 2003; Tu et al. 2006). This research, along with characterization of the types and rates of senescence and the importance of genes and gene classes in hypothesized mechanisms of aging regulation, is beyond the scope of this review. Instead, we describe the identification of specific aging genes in *D. melanogaster*, which show general effects on organismal longevity, or lifespan, and discuss the complexity of the genetic architecture of lifespan as it affects genetics research and the evolution of longevity phenotypes. We also describe the specific cases where allelic variation has been shown to affect lifespan, and provide a hypothetical framework for the maintenance of longevity phenotypes by natural selection in the wild.

IDENTIFICATION OF AGING GENES

Identification of specific genes that regulate lifespan in *D. melanogaster* has been achieved by two processes: mutational analysis (Ford and Tower 2006), in which manipulation of gene or pathway function has demonstrated lifespan extension, and quantitative trait locus (QTL) analysis (Mackay et al. 2006), in which genic elements affecting natural variation in longevity have been mapped to specific positions along the chromosomes. Tables 1 and 2 summarize the characterization of specific genes shown to extend lifespan in *D. melanogaster* by decreased or increased gene activity, respectively, and Table 3 summarizes the characterization of genes for which allelic variation is associated with variation in longevity.

Mutational analysis. *Genes involved in stress response.* The association between stress and lifespan has motivated the identification of many aging genes. The hypothesis that reactive oxygen species (ROS) cause aging (e.g. Harmon 1981) led to tests for lifespan extension by increased activity of genes that promote antioxidant defenses. Overexpression of both *Catalase* (*Cat*) and *Superoxide dismutase* (*SOD*) has demonstrated increased organismal longevity (Orr and Sohal 1994; Parkes et al. 1998), although these effects are highly dependent upon the sex and genetic background of the strains being tested (Sun and Tower 1999; Spencer et al. 2003; Orr and Sohal 2003). The enzyme peptide methionine sulfoxide reductase A counteracts oxidative damage by catalyzing the repair of oxidized methionine, and overexpression of *msrA* in the nervous system can extend lifespan (Ruan et al. 2002). Enhancement of the redox process by overexpression of *glucose-6-phosphate*

dehydrogenase (*G6PD*) also increases lifespan (Legan et al. 2008). Accumulation of isoaspartyl residues in cellular proteins is a degenerative process that affects protein function. Carboxyl methyltransferase (*Pcmt*) counteracts this aging process by modifying isoaspartyl residues, and ubiquitous overexpression of this protein can extend lifespan at elevated temperatures (Chavous et al. 2001). A screen for genes that show differences in gene expression between normal and stress conditions identified several loci that had already been shown to affect lifespan, as well as two additional candidates, heat-shock protein (hsp) genes *hsp26* and *hsp27*; independent overexpression of both genes extends lifespan (Wang et al. 2004). Tests on other molecular chaperones that are induced in response to stress have shown that increasing copy numbers of *hsp70* reduces the mortality rate after a non-lethal induction of stress (Tatar et al. 1997), and that overexpression of *hsp22*, either ubiquitously or in motor neurons, also extends lifespan (Morrow et al. 2004). Similarly, an extra copy of *mei-41*, which may repair DNA damaged by oxidative stress, increases longevity (Symphorien and Woodruff 2003).

Genes involved in insulin signaling. Examination of members in the insulin/insulin-like signaling (IIS) pathway has identified a suite of genes that can extend lifespan by reduction of insulin signaling (Giannakou and Partridge 2007). The role of IIS in nutrient sensing, metabolism and determination of body size, processes which may regulate the well-characterized lifespan extensions by dietary restriction (e.g. Weindruch et al. 1986), and the identification of the nematode insulin receptor homolog *daf-2* as an aging gene (Kenyon et al. 1993), led to tests for lifespan extensions via this pathway in *D. melanogaster*. Independent disruption of the *Insulin-like Receptor* (*InR*) or the *InR* substrate, *chico*, or

TABLE 1. Characterization of genes in *D. melanogaster* that extend lifespan when gene activity is decreased.

Gene	Characterization	Reference
chico	Females homozygous for $chico^{1}$ show a median lifespan increase of 43%; heterozygous females and males show median lifespan increases of 13% and 36%, respectively; homozygous males do not live longer, but age at a slower demographic rate	Clancy et al. 2001; Tu et al. 2002
dilp genes	Ablation of cells producing *dilp2*, *dilp3* and *dilp5* increases median lifespan by 10.5% in males, 18.5% in virgin females and 33.5% in mated females	Broughton et al. 2005
DTS-3	Heterozygous *DTS-3* females, but not males, show a mean lifespan increase of 42% at 29°C; females show increases at lower temperatures if also exposed to higher temperature early in life	Simon et al. 2003
EcR	Males and females heterozygous for multiple alleles disrupting the *EcR* locus show lifespan increases of up to 50%	Simon et al. 2003
Indy	Enhancer trap lines heterozygous for P-element insertions at *Indy* show a doubling of mean lifespan in males and females, but these lifespan extensions are likely artifacts of genetic background and *Wolbachia* infection	Rogina et al. 2000; Toivonen et al. 2007
InR	Heteroallelic, hypomorphic InR^{p5545}/InR^{E19} females show a lifespan extension of 85%; after survival to day 10, males show a lifespan extension of 43%	Tatar et al. 2001
mth	Males and females homozygous for a P-element insertion at *mth* show an average lifespan increase of 35%; UAS/Gal4-mediated expression of an antagonist peptide extended mean lifespan by 38% at 29°C	Lin et al. 1998; Ja et al. 2007
ovo	Females heterozygous for ovo^{D1} show significant lifespan extension; extension is greater in lines selected for short life than lines selected for long life	Sgro and Partridge 1999
puc	Flies heterozygous for loss-of-function alleles $puc^{A251.1}$ or puc^{E69} show extensions of median and maximum lifespan	Wang et al. 2003
rpd3	Males heterozygous for hypomorphic $rpd3^{P\text{-}UTR}$ or null $rpd3^{def24}$ alleles show lifespan extensions of 33% or 41-47%, respectively; heterozygous $rpd3^{P\text{-}UTR}$ females show extensions of 52%	Rogina et al. 2002
sun	Females heterozygous for sun^{EM67} or sun^{Y6} mutations show increases in average lifespan of 25-51%; *sun* resides on the X chromosome and mutations in males are lethal	Cvejic et al. 2004

downstream transcription factor, or *PTEN*, which promotes nuclear localization of endogenous dFOXO, all reduce insulin signaling and extend lifespan (Tatar et al. 2001,

Clancy et al. 2001, Hwangbo et al. 2004, Giannakou et al. 2004). Ablation of cells that produce insulin-like peptides (encoded by *dilp* genes) also increases longevity (Broughton et al. 2005).

Genes that interact with the IIS pathway. Other potential aging genes have been evaluated based on their role in hypothesized mechanisms of aging, including members of a number of pathways that appear to interact with IIS. The elongation factor *EF-1α* is required for protein synthesis, and reduction of both EF-1α and other proteins is associated with senescence; ubiquitous overexpression of *EF-1α* can extend lifespan (Shepherd 1989). The tradeoff between longevity and reproduction prompted evaluation of *ovo*, the disruption of which confers female sterility and lifespan extension (Sgro and Partridge 1999). The target of rapamycin (TOR) pathway interacts with IIS and also regulates body size (e.g. Marygold and Leevers 2002), and inhibition of TOR signaling by single gene manipulation also extends lifespan, including by overexpression of *dTsc1* or *dTsc2* or by expression of dominant-negative forms of *dTOR* or *dS6K* (Kapahi et al. 2004A; Kapahi et al. 2004B). The Jun-N-terminal Kinase (JNK) pathway, which is activated in response to stress, antagonizes IIS and causes nuclear localization of dFOXO (Wang et al. 2005). Extension of lifespan by increased JNK signaling, which is dependent upon *dFOXO* (Wang et al. 2005), has been demonstrated by independent overexpression of *hemipterous* (*hep*) and *hsp68*, and by disruption of *puckered* (*puc*) (Wang et al. 2003). Overexpression of *Drosophila Plenty of SH3s* (*DPOSH*) activates *puc* and also extends lifespan (Seong et al. 2001).

Lifespan extension has been demonstrated by two other genes that may operate through mechanisms related to IIS and dietary restriction, *rpd3* and *Sir2*. Reduction in levels

of the histone deacytlase Rpd3 extends lifespan, but not under dietary restriction, suggesting a common mechanism between these two processes (Rogina et al. 2002). Reduction of *rpd3*

TABLE 2. Characterization of genes in *D. melanogaster* that extend lifespan when gene activity is increased.

Gene	Characterization	Reference
Cat	Transgenic flies with P-element insertions contriuting one extra copy of *Cat* and one extra copy of *SOD* showed median lifespan increases of 6-33%	Orr and Sohal 1994
Cctl	Conditional overexpression via the DOX-dependent P{PdL} system produces an average lifespan increase of 7%	Landis et al. 2003
dFOXO	Conditional overexpression of *dFOXO* in the adult fat body via the mifepristone inducible-Gal4 system and via induction of UAS-dFOXO by P{Switch}$S_1$106 increases median lifespan in both sexes by 35-56% and in females by 22-52%, respectively, but demographic error may account for lifespan extension via the UAS-dFOXO by P{Switch}$S_1$106 system	Hwangbo et al. 2004; Giannakou et al. 2004; Tatar 2005
DPOSH	Neural-specific overexpression of *DPOSH* extends mean lifespan by 14% at 25°C	Seong et al. 2001
dS6K	Inhibition of *TOR* pathway signaling by ubiquitous overexpression with the *da*-Gal4 driver of the dominant-negative UAS-$dS6K^{STDETE}$ produces a mean lifespan increase of 12-34%	Kapahi et al. 2004
dTOR	Inhibition of *TOR* pathway signaling by ubiquitous overexpression with the *da*-Gal4 driver of UAS-$dTOR^{FRB}$ produces a mean lifespan increase of 24-26%	Kapahi et al. 2004
dTsc1	Inhibition of *TOR* pathway signaling by ubiquitous overexpression with the *da*-Gal4 driver of a UAS construct containing *dTsc1* extends mean lifespan by 14%	Kapahi et al. 2004
dTsc2	Inhibition of *TOR* pathway signaling by ubiquitous overexpression with the *da*-Gal4 driver of a UAS construct containing *dTsc2* extends mean lifespan by 12%	Kapahi et al. 2004
EF-1α	Ubiquitous overexpression of *EF-1α* extends male lifspan by 18% at 25°C and 41% at 29.5°C; females were not tested	Shepherd et al. 1989
filamin	Conditional overexpression via the DOX-dependent P{PdL} system produces an average lifespan increase of 8.5%	Landis et al. 2003
fwd	Conditional overexpression via the DOX-dependent P{PdL} system produces an average lifespan increase of 8%	Landis et al. 2003
G6PD	Overexpression of G6PD via the UAS-Gal4 system increases mean lifespan by 16-38% among multiple driver and responder genotypes in males and females	Legan et al. 2008

hebe	Conditional overexpression via the DOX-dependent P{PdL} system increases lifespan in males and females; multiple strains show median lifespan increases of 2.2-31%	Li and Tower 2008
hep	Constitutive overexpression of *hep* in neuronal tissue via the UAS-Gal4 system extends lifespan in males; females were not assayed	Wang et al. 2003
hsp22	Overexpression using the UAS-Gal4 system, either ubiquitously or in motor neurons, increases mean lifespan by approximately 32%	Morrow et al. 2004
hsp26	Overexpression using the UAS-Gal4 system increased mean lifespan by 30-31%	Wang Benzer et al. 2004
hsp27	Overexpression using the UAS-Gal4 system increased mean lifespan by 27-31%	Wang Benzer et al. 2004
hsp68	Constitutive overexpression of *hsp68* via the UAS-Gal4 system extends lifespan in males; females were not assayed	Wang et al. 2003
hsp70	Expression of *hsp70* 10-15% above normal reduces the mortality rate two weeks after non-lethal heat stress, but does not necessarily extend total lifespan	Tatar et al. 1997; Minois et al. 2001
magu	Conditional overexpression via the DOX-dependent P{PdL} system increases lifespan in males and females; multiple strains show median lifespan increases of 2.2-18%	Li and Tower 2008
mei-41	Transgenic flies with one extra copy (but not two extra copies) of wild-type *mei-41* conferred by a P-element transposon show an increase in lifespan	Symphorien and Woodruff 2003
msrA	Overexpression of *msrA* in the nervous system using via the UAS-Gal4 system produces a median lifespan increase of 70%	Ruan et al. 2002
Pcmt	Ubiquitous overexpression of *Pcmt* via the UAS-Gal4 system extends average lifespan by 32-39% at 29°C, but not 25°C	Chavous et al. 2002
PTEN	Overexpression of *dPTEN* in the adult fat body via UAS-*dPTEN* induction by the P{Switch} strain $S_1$32 increases lifespan by 19.6% in females and 17.4% in males	Hwangbo et al. 2004
Sir2	Ubiquitous overexpression of multiple *dSir2* constructs via the UAS-Gal4 system increases lifespan by 18-57%; neuronal overexpression increases lifespan by 20-52%	Rogina and Helfand 2004
SOD	Overexpression of *SOD* by multiple gene copies or by expression of human *SOD1* in adult motor neurons increases mean adult lifespan up to 40%, but lifespan extension by *SOD* is highly dependent upon sex and genetic background	Orr and Sohal 1994; Parkes et al. 1998; Sun and Tower 1999; Spencer et al. 2003; Orr and Sohal 2003
Sug	Conditional overexpression via the DOX-dependent P{PdL} system produces an average lifespan increase of 6%	Landis et al. 2003
VhaSFD	Conditional overexpression via the DOX-dependent P{PdL} system produces an average lifespan increase of 8%	Landis et al. 2003

expression also increases RNA levels of the histone deacetylase Sir2 (Rogina et al. 2002), the direct overexpression of which also extends lifespan (Helfand and Rogina 2004). The activity of these histone deacytlases are further linked together and to IIS: a decrease in Sir2 levels prevents lifespan extension by either *rpd3* or caloric restriction (Helfand and Rogina 2004), lifespan extension in *C. elegans* by *sir-2.1* requires the ortholog to *dFOXO*, *daf-16*, and the mammalian ortholog of *Sir2* regulates the activity of *FOXO* family members (Partridge et al. 2005). IIS is a likely regulator of the synthesis of secondary hormones like juvenile hormone and ecdysone, which are well-evidenced determinants of insect life history (Tu et al. 2006). Mutations that disrupt the ecdysone receptor, *EcR*, and *DTS-3*, which is involved in ecdysone biosynthesis, also extend lifespan (Simon et al. 2003).

Genes in uncharacterized pathways. Mutation screens for longevity genes have identified a handful of candidates not otherwise characterized as members of known pathways that mediate lifespan. P-element insertion screens have shown that disruption of the G-protein coupled receptor *methuselah* (*mth*) and the Krebs cycle cotransporter *I'm not dead yet* (*Indy*) extend lifespan (Lin et al. 1998; Rogina et al. 2000), although the effect at *Indy* appears to be an artifact of genetic background and Wolbacia infection, not activity at the gene itself (Toivonen et al. 2007). Mutations at *stunted* (*sun*), which encodes endogenous peptide ligands of Mth, have also been shown to increase lifespan (Cvejic et al. 2004). Generation of a P-type transposable element with a doxycycline-inducible promoter has permitted, in genome-wide screens, the identification of genes that extend lifespan by overexpression; aging genes identified by this method include *CTP:phosphocholine cytidylyltransferase-I* (*Cctl*), *filamin*, *four wheel drive* (*fwd*), *Sugar baby* (*Sug*), *VhaSFD*

(Landis et al. 2003), *hebe* and *magu* (Li and Tower 2008). Forced misexpression via a genome-wide P-element gene search has also identified 23 genic elements, out of a total of 646 inserts, that are correlated with relatively longer lifespan (Seong et al. 2001); this method may also prove useful in the characterization of specific genes that mediate longevity.

QTL analysis. QTL mapping has identified a substantial number of specific genes that affect longevity, and even more genic regions that contain unexamined candidate loci (for a detailed review, see Mackay et al. 2006). This approach identifies aging loci by localizing differences in longevity between natural strains to chromosomal regions, either flanked by known molecular markers (recombination mapping) (e.g. Falconer and Mackay 1996), or uncovered by deficiency chromosomes containing genomic deletions (deficiency mapping) (e.g. Pasyukova et al. 2000). For example, QTLs have been identified in recombinant inbred lines (RILs) derived from laboratory strains selected for different lifespans, including redundant identification of the same regions by either recombination or selection mapping (Curtsinger and Khazaeli 2002; Forbes et al. 2004; Valenzuela et al. 2004); in RILs derived from recently collected wild strains (Wang et al. 2004; Nuzhdin et al. 2005); and by deficiency mapping to localized chromosomal regions (Pasyukova et al. 2000; Vieira et al. 2000; De Luca et al. 2003; Wilson et al. 2006). Complementation tests using chromosomal deficiencies spanning aging genes already identified by mutational analysis have demonstrated the efficacy of deficiency mapping in identifying longevity QTL (Geiger-Thornsberry and Mackay 2004). RILs derived from *Oregon* (*Ore*) and *2b* strains have also been used to identify dozens of lifespan QTL via recombination mapping (Nuzhdin et al. 1997; Vieria et al. 2000; Leips and Mackay 2000; Leips and Mackay 2002; Leips

TABLE 3. Characterization of genes in *D. melanogaster* for which natural allelic variation is associated with variation in longevity.

Gene	Characterization	Reference
Catsup	Deficiency complementation tests between strains *Ore* and *2b* identify a lifespan QTL containing *Catsup*; complementation tests using $Catsup^{1}$, $Catsup^{cs1}$, and $Catsup^{cs2}$ mutant alleles demonstrate that allelic differences at *Catsup* affect variation in male lifespan; naturally-occurring polymorphisms at *Catsup* demonstrate association with variation in longevity	Mackay et al. 2006; Carbone et al. 2006
cpo	Diapause expression extends lifespan and delays senescence; diapause genotype, independent of diapause expression, is associated with natural variation in longevity; QTL mapping identifies *cpo* as a major determinant of diapause	Tatar Priest et al. 2001 Schmidt and Paaby 2008; Schmidt et al. 2008
Ddc	Deficiency complementation tests between strains *Ore* and *2b* identify a lifespan QTL containing *Ddc*; three natural polymorphisms segregating at *Ddc* account for 15.5% of the genetic contribution to lifespan from chromosome 2	Pasyukova et al. 2000; Leips and Mackay 2002; De Luca et al. 2003
Dox-A2	Deficiency complementation tests between strains *Ore* and *2b* identify a lifespan QTL containing *DOX-A2;* complementation tests using $Dox\text{-}A2^{1}$, $Dox\text{-}A2^{2}$ and $Dox\text{-}A2^{mfs1}$ mutant alleles demonstrate that allelic differences at *Dox-A2* affect variation in male lifespan	Mackay et al. 2006
Lim3	Deficiency complementation tests between strains *Ore* and *2b* identify a lifespan QTL containing *Lim3*; complementation tests using $Lim3^{1}$ and $Lim3^{2}$ mutant alleles demonstrate that allelic differences at *Lim3* affect variation in male lifespan	Mackay et al. 2006
ms(2)35Ci	Deficiency complementation tests between strains *Ore* and *2b* identify a lifespan QTL containing *ms(2)35Ci*; a complementation test using the $ms(2)35Ci^{02316}$ mutant allele demonstrates that allelic differences at *ms(2)35Ci* affect variation in male lifespan	Pasyukova et al. 2000; Pasyukova et al. 2004
stc	Deficiency complementation tests between strains *Ore* and *2b* identify a lifespan QTL containing *stc*; complementation tests using stc^{6} and stc^{11112} mutant alleles demonstrate that allelic differences at *stc* affect sex- and allele-specific variation in lifespan	Pasyukova et al. 2000; Pasyukova et al. 2004
tup (isl, l(2)37Aa)	Deficiency complementation tests between strains *Ore* and *2b* identify a lifespan QTL containing *tup*; a complementation test using the $tup^{isl\text{-}1}$ mutant allele demonstrates that allelic differences at *tup* affect variation in male lifespan	Mackay et al. 2006

and Mackay 2006). Between these same strains, deficiency complementation mapping has uncovered more than a dozen QTL affecting lifespan (Pasyukova et al. 2000; Vieira et al.

2000; De Luca et al. 2003). While this approach yields much higher resolution, the average deficiency QTL still contains about 50 genes (Mackay et al. 2006). However, for genes within these QTL for which mutant alleles are available, gene-specific complementation tests have identified seven genes that contribute to longevity variation between *Ore* and *2b*: *Catsup* (Carbone et al. 2006), *Dopa decarboxylase* (*Ddc)* (De Luca et al. 2003)*, Diphenol oxidase A2* (*Dox-A2*) (Mackay et al. 2006), *Lim3* (Mackay et al. 2006), *ms(2)35Ci* (Pasyukova et al. 2004), *shuttle craft* (*stc*) (Pasyukova et al. 2004) and *tailup* (*tup*) (Mackay et al. 2006).

Like many insects, *D. melanogaster* is capable of expressing a form of diapause, a neuroendocrine mediated physiological syndrome that results in reproductive quiescence and organismal persistence over long periods of suboptimal conditions (Denlinger 2002). Diapause expression extends lifespan and delays senescence (Tatar Priest et al. 2001), but the propensity to express diapause is variable within and among *D. melanogaster* populations: diapause genotype, independent of diapause expression, is associated with natural variation in longevity (Schmidt and Paaby 2008). QTL and complementation mapping identified *couch potato* (*cpo*) as a major determinant of diapause, and consequently as a likely candidate for determining lifespan phenotypes in natural populations (Schmidt et al. 2008). Variation in diapause expression has also been linked to allelic variation at both *timeless* (*tim*), a light-dependent component of the circadian clock (Tauber et al. 2007 Science) and the insulin-regulated phosphatidylinositol 3-kinase (PI3-kinase) gene *Dp110* (Williams et al. 2006). Diapause in the house mosquito *Culex pipiens* is also regulated by *Dp110* (Sim and Denlinger 2008 PNAS), and the homolog of *Dp110* in *C. elegans*, *age-1*, is well established as a gene that affects aging and dauer formation in worms. This suggests that insect diapause and the

dauer phenotype in *C. elegans* may be in part regulated by homologous pathways (Tatar and Yin 2001 Exp Gerontol); while direct mutational and functional tests of diapause genes on longevity phenotype are lacking, such analyses have the potential to confirm *tim*, *cpo* and *Dp110* as aging genes in *Drosophila*.

COMPLEXITY OF GENETIC ARCHITECTURE

Epistasis and background effects. The characterization of single genes that affect longevity has been invaluable in identifying underlying pathways and mechanisms. However, because lifespan is a highly quantitative trait, its genetics can only be characterized comprehensively on a genome-wide level. Aging genes span many functional classes and implicate diverse biochemical and physiological processes in the mediation of longevity; future identification of aging genes will broaden these lists, and it is likely that a substantial fraction of the genome participates in the regulation of lifespan (Mackay et al. 2006). Screens that have evaluated genome-wide patterns of gene expression have also identified a wide range of gene classes that are differentially expressed between young and old ages, including several specific loci that have already been shown to mediate lifespan (Pletcher et al. 2002; Seroude et al. 2002; Landis et al. 2004); these approaches should prove useful in the continued identification of aging genes and the characterization of genetic interactions. The broad range of loci, pathways and processes that affect longevity means that investigation of any subset of these mechanisms will always be influenced by other factors, and the importance of background effects has been demonstrated by every approach used to explore

lifespan. In addition to genetic background, the interaction of any one gene with the rest of the genome, or epistasis, these effects also include sex, age, and environment. Consequently, the complexity of genetic determination of lifespan is not only compounded by epsistatic interactions among genes, but also by interactions between those genes and the sex and age of the organism, and between genes and environmental conditions like diet quality and temperature. This complexity can frustrate the identification and characterization of genic elements that affect longevity, and make comparisons between experimental findings difficult.

The importance of genetic background in the expression of lifespan phenotypes is not surprising given the quantitative nature of the trait, and its relevance has been well-demonstrated in the specific cases for which it has been tested. The original characterization of *Indy* as an aging gene showed that the robust lifespan extension in short-lived strains is weaker in strains selected for long life (Rogina et al. 2000); lifespan extension by *ovo* is also shorter in lines selected for long life compared to lines selected for short life (Sgro and Partridge 1999); and lifespan extension by overexpression of *SOD* was also shown to be dependent upon the laboratory strain used (Sun and Tower 1999). Additional work has further characterized longevity expression by *SOD* as highly dependent on genetic background and sex: overexpression of human *SOD* in the motorneurons in ten lines extends lifespan in some but not others, and these effects are different between males and females (Spencer et al. 2003). Evaluation of the functional significance of allelic variation at *mth* has also shown that the contribution of different alleles to longevity is significantly affected by epistasis and sex (Paaby and Schmidt 2008). Epistatic interactions in the expression of

longevity have also been demonstrated among lifespan QTL and among marker pairs within inbred lines, some of which show sensitivity to mating status (Mackay et al. 2006).

Individual aging genes identified by mutational (Shepherd et al. 1989; Clancy et al. 2001; Tatar et al. 2001; Chavous et al. 2002; Rogina et al. 2002; Simon et al. 2003; Hwangbo et al. 2004; Broughton et al. 2005) and QTL (Nuzhdin et al. 1997; Leips and Mackay 2000; Pasyukova et al. 2000; Vieira et al. 2000; Leips and Mackay 2002; Pasyukova et al. 2004) analysis also routinely show effects by temperature and sex. For example, lifespan is extended by overexpression of *Pcmt* at 29°C, but not 25°C (Chavous et al. 2002); allelic variation at *ms(2)35Ci* affects male, but not female, longevity (Pasyukova et al. 2004). The cross-sex genetic correlation (r_{GS}) describes the degree to which one locus affects both sexes in a quantitative analysis; r_{GS} values range between -1 (if the same genes affect variation in longevity in males and females, but in opposite directions), 0 (if different genes affect variation in longevity between the sexes), and 1 (if the same genes affect the sexes in the same direction). Estimates of r_{GS} from virgin RILs derived from *Ore* and *2b* strains were approximately 0.2, indicating some, but not complete, consistency between the sexes; when these same lines were mated, and/or reared in stressful conditions, estimates of r_{GS} increased (Mackay et al. 2006). Age also appears to affect the expression of genes that mediate longevity: QTL that affect lifespan have also been shown to affect other traits in an age-dependent manner, including fecundity (Leips and Mackay 2002; Leips et al. 2006) and metabolic rate (Khazaeli et al. 2005); QTL that directly affect mortality also show dependence on organism age (Nuzhdin et al. 2005). These results underscore the as yet uncovered complexity of interactions at loci that, in the majority of cases, have only been

narrowly explored. Furthermore, the sensitivity of individual genes to background effects means that detection of candidate aging genes can be obscured or spurious, depending upon conditions: screening for lifespan effects in laboratory strains that have evolved high early fecundity and shortened lifespan may be biased towards the identification of genes that restore normal lifespan, or show disproportionate effects in those backgrounds (Linnen et al. 2001).

Correlations and pleiotropy. Just as the contribution of any one gene to the expression of longevity may be dependent upon other factors, the expression of longevity itself is complicated by pleiotropic effects of its determinants. Two major correlations underly the biology of aging: the negative correlation between lifespan and reproduction (Reznick 1985; Stearns 1991), and the positive correlation between lifespan and stress tolerance (Vermeulen and Loeschcke 2006). While comprehensive characterization of these correlations is beyond the scope of this review, consistent demonstration of associations between these traits clearly implicates shared genetic and physiological mechanisms and provides a likely foundation for the evolution of longevity phenotypes. In fact, the ability to identify aging genes by extended longevity mutant phenotypes underscores the existence of antagonistic pleiotropy at these loci: gene disruption that extends life must also compromise fitness, or else functional copies would not persist in populations. Despite some evidence that lifespan and reproduction can be decoupled, (Partridge et al. 2005; Carbone et al. 2006; Li and Tower 2008), multiple analyses have revealed previously undetected tradeoffs under specific conditions (Marden et al. 2003; Mockett and Sohal 2006). Furthermore, the close association between stress tolerance and longevity has permitted the use of this trait as a

proxy for long-lived phenotypes in studies that examine the genetic basis of lifespan (e.g. Wang et al. 2004).

Artificial selection experiments have been instrumental in demonstrating genetic correlations between lifespan, reproduction and stress tolerance (e.g. Luckinbill et al. 1984; Service et al. 1985; Rose et al. 1992; Chippendale 1998; Mocket et al. 2001; Bubliy and Loeschcke 2005), and single-gene manipulations have routinely demonstrated decreases in reproductive success and/or increases in stress tolerance with longevity extension (e.g. Lin et al. 1998; Clancy et al. 2001; Tatar et al. 2001; Cvejic et al. 2004; Hwangbo et al. 2004). Gene expression assays have also demonstrated correlations between stress and longevity: genes associated with relatively longer lifespan in a misexpression screen were positively correlated with oxidative stress tolerance, and of those with known function, half are involved in stress resistance or redox balance (Seong et al. 2001); of genes showing differential expression between young and old ages, about one third of those expressed at old ages also respond to oxidative stress (Zou et al. 2000; Landis et al. 2004). Longevity QTL have also been mapped to the same regions as QTL for fertility and stress tolerance (Nuzhdin et al. 1997; Leips and Mackay 2002; Leips et al. 2006). Pleiotropy at genic elements that mediate lifespan should exist so long as mechanisms are shared among traits. Parsing those mechanisms into discreet components that behave singularly may be possible in some cases, but identification of elements that induce shorter lifespan should only be possible under two conditions: 1) pleiotropy, where the element antagonistically affects another trait, and is maintained by positive selection for that trait; 2) neutrality, where the element does not affect fitness, and is never subject to natural selection. Given the persistence of many functional aging genes over

historical time, the importance of pleiotropy is clear in the maintenance of genetically mediated longevity.

VARIATION IN NATURAL POPULATIONS

Phenotypic variation. *D. melanogaster* shows significant variation in longevity within and among natural populations. Between populations, patterns of longevity correlate with latitude, and are possibly driven by differential selection imposed by variation between tropical and temperate climates. Isofemale lines derived from high latitude populations from the United States east coast show longer lifespan than lines from low latitude populations, and lines along this gradient show covariance between longevity and other life history traits, including incidence of reproductive diapause, triglyceride content, oocyte development, ovariole number and fecundity (Schmidt et al. 2005). Recently derived lines from temperate European populations and tropical Central American and African populations also show differences in mean lifespan, and mean lifespan under different thermal environments (Trotta et al. 2006). Differences in longevity have also been observed between inbred lines recently derived from natural populations near Ankara, Turkey (Cakir and Bozcuk 2000). Longevity also varies significantly within populations: although lifespan varies predictably with geography in North America, diapause genotype explains the majority of the variation in longevity and other associated traits, including variation between individuals from the same population (Schmidt and Paaby 2008). While longevity estimates from recently collected wild lines provide useful measures for experimentation in the laboratory, they do not

necessarily represent actual age distributions in the wild. Mueller *et al*. (2004) demonstrate a method for determining age-specific survival and mortality in natural populations by marking individuals sampled from the wild at unknown age and subsequently constructing life tables from recorded times-of-death. This technique has been used to describe the survival schedule of the medfly *Ceratitis capitata*, and could be of great utility in describing the age structures of natural populations of *D. melanogaster*.

Genetic variation. *D. melanogaster* exhibit robust genetic variance for lifespan, or evidence that nucleotide variation affects longevity phenotypes. Phenotype means for lifespan and other life history traits from inbred lines derived from North American populations show genetic variance for and genetic correlations among these traits (Schmidt et al. 2005 Hes). QTL analyses and artificial selection regimes have also demonstrated that flies derived from the wild harbor allelic variation that affects lifespan: the QTL approach to the identification of aging genes, which has produced more candidates than can be easily tested (Mackay et al. 2006), is predicated upon alleles for longevity segregating in natural populations, and laboratory populations of *D. melanogaster* have shown robust responses to myriad artificial selection regimes affecting longevity, indicating the existence of sufficient additive genetic variance for these traits (Harshman and Hoffman 2000). Artificial selection on multiple traits has demonstrated phenotypic responses in longevity, including selection on time of reproduction (Luckinbill et al. 1984; Rose 1984; Luckinbill and Clare 1985; Service et al. 1985; Arking 1987; Fukui et al. 1995; Partridge et al. 1999), stress tolerance (Rose et al. 1992; Chippendale et al. 1998), and response to diet quality (Chippendale et al. 1993). While inconsistencies have been observed (Baret and Lints 1993; Baret et al. 1996; Harshman et al.

1999; Norry and Loeschcke 2002), these studies clearly demonstrate that wild-derived lines harbor genetic variation at loci that affect aging, and that longevity tends to correlate negatively with reproduction and positively with stress and starvation tolerance. One study has performed direct artificial selection on lifespan in *D. melanogaster* (Zwaan et al. 1995), and results support these same correlations between longevity, reproduction and stress tolerance.

Identification of aging genes by induced mutation, transgenics or QTL analysis has begun to improve our understanding of the number, type, and magnitude of effect of genes that affect lifespan. However, these approaches do not test whether natural populations harbor segregating allelic variation, the substrate upon which natural selection acts, at these identified loci (Flatt 2004). Evaluating these specific loci for natural allelic variation can test the importance of candidate aging genes in the evolution of longevity, and also reveal subtleties in the mechanisms of lifespan mediation that laboratory-induced mutations cannot. To date, natural allelic variation has been characterized at a handful of loci identified as aging genes. The G-protein coupled receptor, *mth*, identified by mutational screen (Lin et al. 1998), shows a cline in frequency of the most common haplotype across the latitudinal gradient of the U.S. east coast (Schmidt et al. 2000). This haplotype is comprised of five SNPs across the coding region; no individual SNP shows this pattern across geography, but decay of linkage disequilibrium 5' and 3' of the *mth* locus (Duvernell et al. 2003) suggests that selection is acting on one or more unidentified polymorphisms within the gene. Allelic variation at *mth* is functionally significant: in a modified quantitative complementation scheme, wild-derived *mth* alleles showed significant differences in lifespan, fecundity and

resistance to oxidative stress (Paaby and Schmidt 2008). These results support the conclusion that some natural genetic variation at *mth* is adaptive, and current work is exploring sequence variation in UTRs in an effort to identify the functional polymorphism(s) within the locus.

Sequence polymorphism data are also available for genes in the insulin signaling pathway, one of which shows evidence of recent adaptive evolution. *InR* and *chico* were sequenced in wild lines from North America, and while both exhibit substantial polymorphism, only *InR* shows a significant pattern of allele frequency across latitude (Paaby et al. In Review). Specifically, an indel polymorphism in the first exon, which disrupts a region of glutamine-histidine repeats, shows significantly increasing and decreasing frequencies across latitude for the two most common alleles. Additional sequencing and polymorphism screening in Australia shows a nearly identical pattern on that continent (Paaby et al. In Review). These results suggest that this polymorphism is under direct selection across heterogeneous environments, and current work is testing the effect of this polymorphism on levels of insulin signaling and life history phenotypes. Although hypomorphic mutations at both *InR* and *chico* reduce insulin signaling and produce similar mutant phenotypes (Tatar et al. 2001; Clancy et al. 2001), only *InR* demonstrates an adaptive response at the nucleotide level in natural populations: analysis of wild-derived *chico* sequences show evidence of neutral evolution (Paaby et al. 2009). This suggests that while reduction of insulin signaling at almost any point in the pathway can extend lifespan (Giannakou and Partridge 2007), different pathway members may be under different constraints. The downstream transcription factor *dFOXO* is an important component of the insulin signaling pathway: overexpression of *dFOXO* extends lifespan (Hwangbo *et al*. 2004),

upregulation of *dFOXO* may be required for lifespan extension by other genes in the pathway, and the *C. elegans* ortholog *daf-16* is essential for lifespan extension by insulin signaling (reviewed in Giannakou and Partridge 2007). Natural variation at the human ortholog of *dFOXO*, *FOXO3A*, contributes significantly to differences in longevity in human populations (Willcox *et al*. 2008), and characterization of natural genetic variation at *daf-16* and other IIS pathway members in *C. elegans* is underway (R. Jovelin and P. Phillips, personal communication). Variation at this key transcription factor in other species is unknown, but future characterization and comparisons across taxa could yield insight into how a highly pleiotropic pathway can respond to natural selection . Specifically, evaluation of allelic variation at *dFOXO* in natural populations of *D. melanogaster* will deepen our understanding of how this pathway mediates longevity, especially since other pathway members have already been characterized in this system.

Transgenic experiments have shown that *Dp110*, another member of the insulin signaling pathway, affects the expression of reproductive diapause (Williams *et al*. 2006). In *D. melanogaster*, diapause produces a significant extension in lifespan and is associated with genetic variance for longevity, fecundity, development time, lipid content and stress tolerance (Schmidt et al. 2005; Schmidt et al. 2005; Schmidt and Conde 2006; Schmidt and Paaby 2008). Two natural *Dp110* alleles which differentially affect diapause have been sequenced, though there is an absence of pronounced polymorphism between them: of 20 polymorphisms detected, none affect the amino acid sequence nor levels of RNA (Williams *et al*. 2006). However, a single nucleotide polymorphism at *cpo* shows strong control over the ability of *D. melanogaster* to enter diapause (Schmidt et al. 2008). The SNP shows a

significant cline in frequency across latitude in the eastern U.S., suggesting a functional response to heterogeneous selection pressure. While neither of these genes have been demonstrated as aging genes in the laboratory, the association of *Dp110* with the insulin signaling pathway, a major mediator of lifespan, and the association of both genes with diapause, a process that directly mediates lifespan and is strongly genetically correlated with natural variation in longevity (Tatar Priest et al. 2001; Schmidt and Paaby 2008), suggests that they may play an important role in determining lifespan phenotypes in wild populations.

Ddc was identified as a candidate gene for aging by QTL mapping and complementation tests, and linkage disequilibrium mapping shows that it is associated with natural variation for lifespan (De Luca et al. 2003). Sequence data from natural lines revealed high levels of polymorphism at the locus, including at the promoter, and high linkage disequilibrium between sites that is suggestive of balancing selection or a recent selective sweep. Frequency spectra of polymorphic sites also support balancing selection, with apparent selection on both long- and short-lived phenotypes. These data demonstrate that natural allelic variation is segregating at *Ddc*, and that specific polymorphisms within the gene are likely targets of selection in the wild. The gene *Catsup* is an aging gene that has pleiotropic effects on multiple traits, but individual polymorphisms within the locus show an absence of pleiotropy: identified from wild-derived *Catsup* alleles, the polymorphisms show independent effects on longevity, locomotor behavior and sensory bristle number (Carbone et al. 2006). In addition to identifying *Catsup* as a potentially important contributor to genetic variance for lifespan in the wild, these results also expose details in mechanisms of trait determination that may only be revealed by the subtle variations exhibited by natural

mutations.

Evolution in natural populations. Two major theories have been proposed to explain the evolution of lifespan, and they have been extensively reviewed in the literature (Rose 1991, Charlesworth 1994, Promislow and Tatar 1998, Partridge and Mangel 1999, Zwaan 1999, Kirkwood and Austad 2000). These theories are not mutually exclusive, and they rely on the assumption that the strength of natural selection decreases with organism age (Medawar 1952, Williams 1957). This decrease is evident even in populations which show no age-related decline in reproductive fitness or other traits: Medawar's imaginary population of non-senescing test tubes, for example, will experience weakest selection on the oldest age class. The weak selection is a function of age structure in the population, as the oldest age class will always be the smallest (Medawar 1952).

The mutation accumulation theory posits that while mutations that deleteriously affect reproduction and survival will be quickly eliminated from a population if they are expressed early in life, similar mutations with late-life expression may escape natural selection (Medawar 1952). Consequently, an accumulation of late-acting deleterious mutations may be responsible for the senescent phenotypes observed in most metazoans. Multiple predictions accompany this model (for a complete review, see Promislow and Tatar 1998), including an increase in additive genetic variance for fitness traits with age and an effect by induced or accumulated mutations on patterns of longevity. *D. melanogaster* has been a major empirical system in testing these predictions. Increases in additive genetic variance at later ages has been demonstrated for several traits, including lifespan (Hughes and Charlesworth 1994; Charlesworth and Hughes 1996, Nuzhdin et al. 1997), fecundity

(Engstrom et al. 1989, Tatar et al. 1996, Leips et al. 2006), and male mating ability (Kosuda 1985, Hughes 1995). However, these results are not definitive. Other studies have failed to find increases in additive genetic variance with age (Rose and Charlesworth 1980, Rose and Charlesworth 1981b, Promislow et al. 1996), or suggest that high estimates may be artifacts of statistical methodology (Tatar et al. 1996, Promislow and Tatar 1998). Experiments that have induced mutations or permitted the accumulation of natural mutations in laboratory populations show some evidence of mutational variance in age-related traits (Houle et al. 1994, Clark and Guadalupe 1995), but these results do not appear sufficient to fully explain the persistence of senescence. However, the characterization of many lifespan QTL that only show late age-of-onset effects provides strong evidence that the accumulation of late-acting mutations contributes to the evolution of senescence (Nuzhdin et al. 2007).

The theory of antagonistic pleiotropy (Williams 1957) differs from the theory of mutation accumulation in the expectation that late-acting, deleterious mutations may have beneficial, rather than neutral, effects early in life. For example, an allele that promotes senescence might experience positive selection if it produces a fitness benefit at early age, such as increased reproductive success. The identification of aging genes by mutational analysis supports the theory of antagonistic pleiotropy: functional, wild-type copies of aging genes would not persist in nature if null copies produce only putatively beneficial lifespan extensions. In fact, as previously discussed, most lifespan extension mutations produce additional costs to fitness. The observation of additive genetic variance for lifespan also supports the theory of antagonistic pleiotropy. Pleiotropic alleles that produce only beneficial or only deleterious effects should be fixed or purged in populations, reducing variation; but

different alleles that affect fitness both positively and negatively may be maintained, preserving variation (Williams 1957). Even then, it is likely that variation in selection pressure, across temporally or spatially heterogeneous environments, plays a role in the maintenance of this allelic variation at longevity loci. Many studies have contributed positive evidence for the theory of antagonistic pleiotropy, including genetic correlations among life history traits by artificial selection (e.g. Luckinbill et al. 1984; Service et al. 1985; Rose et al. 1992; Chippendale 1998; Mocket et al. 2001; Bubliy and Loeschcke 2005), QTL analyses (e.g. Nuzhdin et al. 1997; Leips and Mackay 2002; Leips et al. 2006) and characterization of wild lines (e.g. Schmidt et al. 2005; Schmidt and Paaby 2008). However, characterization of the functional effects of allelic variation at individual loci and specific nucleotides within those loci can offer precise examples of how longevity genes evolve in natural populations.

CONCLUSION

The utility of *D. melanogaster* for investigating the genetic basis of longevity lies in its complementary resources: it offers both powerful genetic tools and a natural ecology that effectively provides a grand naturalistic experiment. *D. melanogaster* originated in tropical Africa and has colonized temperate regions, including the European, American and Australian continents, within the last several thousand years (David and Capy 1988). Surveys at multiple loci have revealed patterns of variation along latitudinal clines in which the frequencies of the derived

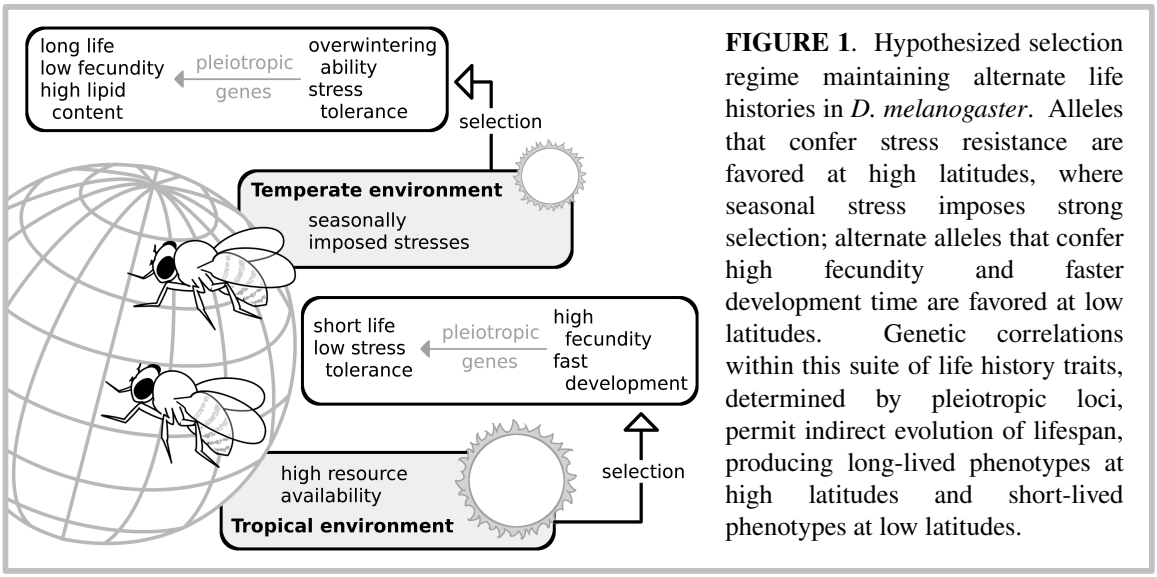


FIGURE 1. Hypothesized selection regime maintaining alternate life histories in *D. melanogaster*. Alleles that confer stress resistance are favored at high latitudes, where seasonal stress imposes strong selection; alternate alleles that confer high fecundity and faster development time are favored at low latitudes. Genetic correlations within this suite of life history traits, determined by pleiotropic loci, permit indirect evolution of lifespan, producing long-lived phenotypes at high latitudes and short-lived phenotypes at low latitudes.

alleles increase with latitude, suggesting adaptation to temperate habitats (Sezgin et al. 2004). Moreover, the changing patterns of longevity, fecundity, stress resistance, development time and other life history variables exhibited by natural populations of *D. melanogaster* along such latitudinal gradients co-occur with hypothesized changes in environmental selection pressures (Schmidt *et al*. 2005). A hypothetical selection regime, which imposes seasonal stresses at high latitudes, may favor stress resistant alleles in some environments and highly fecund alleles in others, indirectly driving the evolution of longevity and maintaining the distribution of lifespan phenotypes that we observe (Figure 1). The identification of aging genes by mutational or QTL analysis offers an opportunity for fine-scale characterization of genetic variation for longevity using a gene-targeted approach. The demonstration of allelic patterns of variation at *mth* (Schmidt et al. 2000), *InR* (Paaby et al. In Review) and *cpo* (Schmidt et al. 2008) across heterogeneous environments shows how individual genes may contribute to the determination of these divergent life histories. Moreover, evaluation of how

polymorphic alleles (Paaby and Schmidt 2008) or individual polymorphisms (Carbone et al. 2006) affect phenotype can elucidate both how life histories evolve in natural populations, and how genotypes translate into phenotypes.

ACKNOWLEDGEMENTS

We thank one anonymous reviewer for helpful comments on the manuscript.

LITERATURE CITED

Alvarez-Ponce D, Aguadé M, Rozas J. 2009. Network-level molecular evolutionary analysis of the insulin/TOR signal transduction pathway across 12 *Drosophila* genomes. *Genome Research* 19: 234-242.

Anderson AR, Collinge JE, Hoffmann AA, Kellett M, McKechnie SW. 2003. Thermal tolerance tradeoffs associated with the right arm of chromosome 3 and marked by the *Hsr-omega* gene in *Drosophila melanogaster*. *Heredity* 90: 195-202.

Anderson AR, Hoffmann AA, McKechnie SW, Umina PA, Weeks AR. 2005. The latitudinal cline in the *In(3R)Payne* inversion polymorphism has shifted in the last 20 years in Australian *Drosophila melanogaster* populations. *Molecular Ecology* 14: 851-858.

Anderson WW. 1973. Genetic divergence in body size among experimental populations of Drosophila subobscura kept at different temperatures. *Evolution* 27: 278-284.

Arking R. 1987. Successful selection for increased longevity in *Drosophila*: analysis of the survival data and presentation of a hypothesis on the genetic regulation of longevity. *Experimental Gerontology* 22: 199-220.

Baldal EA, Baktawar W, Brakefield PM, Zwaan BJ. 2006. *Methuselah* life history in a variety

of conditions, implications for the use of mutants in longevity research. *Experimental Gerontology* 41: 1126-1135.

Baret P, Lints FA. 1993. Selection for increased longevity in *Drosophila melanogaster*: a new interpretation. *Gerontology* 39: 252-259.

Baret PV, Le Bourg E, Lints FA. 1996. Selection for increased longevity in *Drosophila melanogaster*: reflections on new data. *Gerontology* 42: 14-17.

Berry A, Kreitman M. 1993. Molecular analysis of an allozyme cline: alcohol dehydrogenase in *Drosophila* melanogaster on the east coast of North America. *Genetics* 134: 869-893.

Bock IR, Parsons PA. 198. Species of Australia and New Zealand. In: *Genetics and Biology of Drosophila,* Vol. 3a (eds. Asherburner M, Carson HL, Thompson JN) Academic Press, London, U.K.

Böhni R, Riesgo-Escovar I, Oldham S, Brogiolo W, Stocker H, Andruss BF, Beckingham K, Hafen E. 1999. Autonomous control of cell and organ size by CHICO, a *Drosophila* homolog of vertebrate IRS1-4. *Cell* 97: 865–875.

Brogiolo W, Stocker H, Ikeya T, Rintelen F, Fernandez R, Hafen E. 2001. An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth

control. *Current Biology* 11: 213–221.

Broughton SJ, Piper MDW, Ikeya T, Bass TM, Jacobson J, Driege Y, Martinez P, Hafen E, Withers DJ, Leevers SJ, Partridge L. 2005. Longer lifespan, altered metabolism, and stress resistance in *Drosophila* from ablation of cells making insulin-like ligands. *Proceedings of the National Academy of Sciences, USA* 102: 3105-3110.

Bubliy OA, Loeschcke V. 2005. Correlated responses to selection for stress resistance and longevity in a laboratory population of *Drosophila melanogaster*. *Journal of Evolutionary Biology* 18: 789-803.

Cakir S, Bozcuk AN. 2000. Longevity in some wild type and hybrid strains of *Drosophila melanogaster*. *Turkish Journal of Biology* 24: 321-329.

Capy P, Pla E, David JR. 1993. Phenotypic and genetic variability of morphometrical traits in natural populations of *Drosophila melanogaster* and *D. simulans*. I. Geographic variations. *Genetics Selection Evolution* 25: 517-536.

Capy P, Pla E, David JR. 1993. Phenotypic and genetic variability of morphometrical traits in natural populations of *Drosophila melanogaster* and *Drosophila simulans*. *Evolution* 25: 517-536.

Capy P, Pla E, David JR. 1994. Phenotypic and genetic variability of morphometrical traits in natural populations of *Drosophila melanogaster* and *D. simulans*. II. Within population variability. *Genetics Selection Evolution* 26: 15-28.

Carbone MA, Jordan KW, Lyman RF, Harbison ST, Leips J, Morgan TJ, De Luca M, Awadalia P, Mackay TFC. Phenotypic variation and natural selection at *Catsup*, a pleiotropic quantitative trait gene in *Drosophila*. *Current Biology* 16: 912-919.

Casas-Tinto S, Marr MT II, Andreu P, Puig O. 2007. Characterization of the *Drosophila* insulin receptor promoter. *Biochimica et Biophysica Acta* 1769: 236-243.

Cavicchi SD, Guerra G, Giorgi G, Pezzoli C. 1985. Temperature-related divergence in experimental populations of *Drosophila melanogaster*. 1. Genetic and developmental basis of wing size and shape variation. *Genetics* 109: 665-689.

Charlesworth B, Hughes KA. 1996. Age-specific inbreeding depression and components of genetic variance in relation to the evolution of senescence. *Proceedings of the National Academy of Sciences, USA* 93: 6149-6145.

Charlesworth B. 1994. Evolution in Age-Structured Populations. Cambridge: Cambridge University Press.

Charlesworth J, Eyre-Walker A. 2008. The McDonald-Kreitman test and slightly deleterious mutations. *Molecular Biology and Evolution* 25: 1007-1015.

Chavous DA, Jackson FR, O'Connor CM. 2001. Extension of the *Drosophila* lifespan by overexpression of a protein repair methyltransferase. *Proceedings of the National Academy of Sciences, USA* 98: 14814-14818.

Chippendale AK, Gibbs AG, Sheik M, Yee KJ, Djawdan M, Bradley TJ, Rose MR. 1998. Resource acquisition and the evolution of stress resistance in *Drosophila melanogaster*. *Evolution* 52: 1342-1352.

Chippindale AK, Leroi AM, Kim SB, Rose MR. 1993. Phenotypic plasticity and selection in *Drosophila* life-history evolution. I. Nutrition and the cost of reproduction. *Journal of Evolutionary Biology* 6: 171-193.

Clancy DJ, Gems D, Harshman LG, Oldham S, Stocker H, Hafen E, Leevers SJ, Partridge L. 2001. Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein. *Science* 292: 41-43.

Clark AK, Guadalupe RN. 1995. Probing the evolution of senescence in *Drosophila melanogaster* with p-element tagging. *Genetica* 96: 225-234.

Curtsinger JW, Khazaeli AA. 2002. Lifespan, QTLs, age-specificity, and pleiotropy in *Drosophila*. *Mechanisms of Ageing and Development* 123: 81-93.

Curtsinger JW. 2003. Peeking under QTL peaks. *Nature Genetics* 34: 358-359.

Cvejic S, Zhu Z, Felice SJ, Berman Y, Huang X-Y. 2004. The endogenous ligand Stunted of the GPCR *Methuselah* extends lifespan in *Drosophila*. *Nature Cell Biology* 6: 540-546.

Da Lage JL, Capy P, David JR. 1990. Starvation and desiccation resistance in *Drosophila melanogaster*: differences between European, North Africa and Afrotropical populations. *Genes Selection Evolution* 22: 381-391

David J, Cohet Y, Fouillet P. 1975. Physiologie de l'anition et utilisation des reserves chez les adultes de *Drosophila melanogaster*. *Archives de Zoologie Experimentale et Generale* 116: 579-590.

David JR, Capy P. 1988. Genetic variation of *Drosophila melanogaster* natural populations. *Trends in Genetics* 4: 106-111.

David JR. 1975. Evolution of a cosmopolitan species: genetic latitudinal clines in *Drosophila melanogaster* wild populations. *Experientia* 31: 164-166.

De Jong G, Bochdanovits Z. 2003. Latitudinal clines in *Drosophila melanogaster*: body size, allozyme frequencies, inversion frequencies, and the insulin signalling pathway. *Journal of Genetics* 82: 207-223.

De Luca M, Roshina NV, Geiger-Thornsberry GL, Lyman RF, Pasyukova EG, Mackay TFC. 2003. Dopa decarboxylase (*Ddc*) affects variation in *Drosophila* longevity. *Nature Genetics* 34: 429-433.

Denlinger DL. 2002. Regulation of diapause. *Annual Reviews Entomology* 47: 93-122.

Duvernell DD, Schmidt PS, Eanes WF. 2003. Clines and adaptive evolution in the *methuselah* gene region in *Drosophila melanogaster*. *Molecular Ecology* 12: 1277-1285.

Endler JA. 1986. Natural Selection in the Wild. Princeton, NJ, Princeton University Press.

Engstrom G, Liljedahl L-E, Rasmuson M, Bjorklund T. 1989. Expression of genetic and environmental variation during ageing. *Theoretical Applied Genetics* 77: 119-122.

Falconer DSF, Mackay TFC. 1996. Introduction to Quantitative Genetics. Fourth Edition. Harlow, Essex, Addison Wesley Longman.

Flachsbart F, Caliebe A, Kleindorp R, Blanché H, von Eller-Eberstein H, Nikolaus S,

Schreiber S, Nebel A. 2008. Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proceedings of the National Academy of Sciences, USA* 106: 2700-2705.

Flatt T. 2004. Assessing natural genetic variation in genes affecting *Drosophila* lifespan. *Mechanisms of Ageing and Development* 125: 155-159.

Forbes SN, Valenzuela RK, Keim P, Service PM. 2004. Quantitative trait loci affecting life span in replicated populations of *Drosophila melanogaster*. I. Composite interval mapping. *Genetics* 168: 301-311.

Ford D, Tower J. 2006. Genetic manipulation of lifespan in *Drosophila melanogaster*. In: Masaro EJ, Austad SN, eds. Handbook of the Biology of Aging, Sixth Edition. Burlington, MA, Elsevier Press, 400-414.

Fu Y-Y, Li W-H. 1993. Statistical tests of neutrality of mutations. *Genetics* 133, 693-709.

Fukui HH, Pletcher SD, Curtsinger JW. 1995. Selection for increased longevity in *Drosophila melanogaster*: a response to Baret and Lints. *Gerontology* 41: 65-68.

Garofalo RS. 2002. Genetic analysis of insulin signaling in *Drosophila*. *Trends in Endocrinology and Metabolism* 13: 156−162.

Geiger-Thornsberry G, Mackay TFC. 2004. Quantitative trait loci affecting natural variation in *Drosophila* longevity. *Mechanisms of Ageing and Development* 125: 179-189.

Giannakou ME, Goss M, Juenger MA, Hafen E, Leevers SJ, Partridge L. 2004. Long-lived *Drosophila* with overexpressed dFOXO in adult fat body. *Science* 305: 361.

Giannakou ME, Partridge L. 2007. Role of insulin-like signalling in *Drosophila* lifespan. *Trends in Biochemical Science* 32: 180-188.

Gidaszewski NA, M Baylac, CP Klingenberg. 2009. Evolution of sexual dimorphism of wing shape in the *Drosophila melanogaster* subgroup. *BMC Evolutionary Biology* 9: 110.

Guirao-Rico S, Aguadé M. 2009. Positive selection has driven the evolution of the *Drosophila Insulin-like Receptor* (*InR*) at different timescales. *Molecular Biology and Evolution* 26: 1723-1732.

Harmon D. The aging process. *Proceedings of the National Academy of Sciences, USA* 78: 7134-7128.

Harshman LG, Hoffmann AA. 2000. Laboratory selection experiments using *Drosophila*: what do they really tell us? *Trends in Ecology and Evolution* 15: 32-36.

Harshman LG, Moore KM, Sty MA, Magwire MM. 1999. Stress resistance and longevity in selected lines of *Drosophila melanogaster*. *Neurobiology of Aging* 20: 521-529.

Harshman LG, Zera AJ. 2007. The cost of reproduction: the devil in the details. *Trends in Ecology and Evolution* 22: 80-86.

Hartl DL, Moriyama EN, Sawyer SA. 1994. Selection intensity for codon bias. *Genetics* 138: 227-234.

Helfand SL, Rogina B. 2003. Genetics of aging in the fruit fly, *Drosophila melanogaster*. *Annual Reviews Genetics* 37: 329-348.

Hoffmann AA, Anderson A, Hallas R. 2002. Opposing clines for high and low temperature resistance in *Drosophila melanogaster*. *Ecology Letters* 5: 614-618.

Hoffmann AA, Hallas R, Sinclair C, Mitrovski P. 2001. Levels of variation in stress resistance in *Drosophila* among strains, local popualtions, and geographic regions: patterns for desiccation, starvation, cold resistance and associated traits. *Evolution* 55: 1621-1630.

Hoffmann AA, Harshman LG. 1999. Desiccation and starvation resistance in *Drosophila*: patterns of variation at the species, population and intrapopulation levels. *Heredity* 83: 637-643.

Hoffmann AA, Parsons PA. 1989. An integrated approach to environmental stress tolerance and life history variation: desiccation tolerance in *Drosophila*. *Biological Journal of the Linnean Society* 37: 117-136.

Hoffmann AA, Sgro CM, Weeks AR. 2004. Chromosomal inversion polymorphisms and adaptation. *Trends in Ecology and Evolution* 19: 482-488.

Hoffmann AA, Shirriffs J, Scott M. 2005. Relative importance of plastic *vs* genetic factors in adaptive differentiation: geographical variation for stress resistance in *Drosophila melanogaster* from eastern Australia. *Functional Ecology* 19: 222-227.

Hoffmann AA, Weeks AR. 2007. Climatic selection on genes and traits after a 100 year-old invasion: a critical look at the temperate-tropical clines in *Drosophila melanogaster* from eastern Australia. *Genetica* 129: 133-147.

Houle D, Hughes KA, Hoffmaster DK, Ihara J, Assimacopoulos A, Canada D, Charlesworth B. 1994. The effects of spontaneous mutation on quantitative traits. I. Variances and covariances of life history traits. *Genetics* 138: 773-785.

Hughes KA, Charlesworth B. 1994. A genetic analysis of senescence in *Drosophila*. *Nature* 367: 64-66.

Hughes KA. 1995. The evolutionary genetics of male life-history characters in *Drosophila melanogaster*. *Evolution* 49: 521-537.

Hwangbo DS, Gersham B, Tu M-P, Palmer M, Tatar M. 2004. *Drosophila* dFOXO controls lifespan and regulates insulin signalling in brain and fat body. *Nature* 429: 562-566.

Ja W, West AP, Delker SL, Bjorkman PJ, Benzer S, Roberts RW. 2007. Extension of *Drosophila melanogaster* life span with a GPCR peptide inhibitor. *Nature Chemical Biology* 3: 415-419.

James AC, Partridge L. 1995. Thermal evolution of rate of larval development in *Drosophila melanogaster* in laboratory and field populations. *Journal of Evolutionary Biology* 8: 315-330.

Jenkins NL, McColl G, Lithgow GJ. 2004. Fitness cost of extended lifespan in *Caenorhabditis elegans*. *Proceedings of the Royal Society of London, Series B* 271: 2523-2526.

Jünger MA, Rintelen F, Stocker H, Wasserman JD, Végh M, Radimerski T, Greenberg ME, Hafen E. 2003. The *Drosophila* Forkhead transcription factor FOXO mediates the reduction in cell number associated with reduced insulin signaling. *Journal of Biology* 2: 20.

Kacser H, Burns JA. 1973. The control of flux. *Symposium for the Society of Experimental Biology* 27: 65-104.

Kapahi P, Zid BM, Haper T, Koslover D, Sapin V, Benzer S. 2004. Regulation of lifespan in *Drosophila* by modulation of genes in the TOR signaling pathway. *Current Biology* 14: 885-890.

Karan D, Dahiya N, Munjal AK, Gibert P, Moreteau B, Parkash R, David JR. 1998. Desiccation and starvation tolerance of adult *Drosophila*: opposite latitudinal clines in natural populations of three different species. *Evolution* 52: 825-831.

Kennington WJ, Hoffmann AA, Partridge L. 2007. Mapping regions within cosmopolitan inversion *In(3R)Payne* associated with natural variation in body size in *Drosophila melanogaster*. *Genetics* 177: 549-556.

Kennington WJ, Partridge L, Hoffmann AA. 2006. Patterns of diversity and linkage disequilibrium within the cosmopolitan inversion *In(3R)Payne* in *Drosophila melanogaster* are indicative of coadaptation. *Genetics* 172: 1655-1663.

Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R. 1993. A *C. elegans* mutant that lives twice as long as wild type. *Nature* 366: 461-464.

Khazaeli AA, Van Voorhies W, Curtsinger JW. 2005. Longevity and metabolism in *Drosophila melanogaster*: genetic correlations between life span and age-specific metabolic rate in populations artificially selected for long life. *Genetics* 169: 231-242.

Kirkwood TBL, Austad SN. 2000. Why do we age? *Nature* 408: 233-238.

Knibb WR. 1982. Chromosome inversion polymorphism in *Drosophila melanogaster*. II. Geographic clines and climatic associations in Australasia, North America and Asia. *Genetica* 58: 213-222.

Kosuda K. 1985. The aging effect on male mating activity in *Drosophila melanogaster*. *Behavioral Genetics* 15: 297-303.

Landis GN, Abdueva D, Skvortsov D, Yang J, Rabin BE, Carrick J, Tavare S, Tower J. 2004. Similar gene expression patterns characterize aging and oxidative stress in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA* 101: 7663-7668.

Landis GN, Bhole D, Tower J. 2003. A search for doxycycline-dependent mutations that increase *Drosophila melanogaster* life span identifies the *VhaSFD*, *Sugar baby*, *filamin*, *fwd* and *Cctl* genes. *Genome Biology* 4: R8.

Landry CR, Wittkopp PJ, Taubes CH, Ranz JM, Clark AG, Hartl DL. 2005. Compensatory

cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*. *Genetics* 171: 1813-1822.

Langley CH, Lazzaro BP, Phillips W, Heikkinen E, Braverman JM. 2000. Linkage disequilibria and the site frequency spectra in the *su(s)* and *su(w(a))* regions of the *Drosophila melanogaster X* chromosome. *Genetics* 156: 1837-1852.

Legan SK, Rebrin I, Mockett RJ, Radyuk SN, Klichko VI, Sohal RS, Orr WC. 2008. Overexpression of glucose-6-phosphate dehydrogenase extends the life span of *Drosophila melanogaster*. *Journal of Biological Chemistry* 283: 32492-32499.

Leips J, Gilligan P, Mackay TFC. 2006. Quantitative trait loci with age-specific effects on fecundity in *Drosophila melanogaster*. *Genetics* 172: 1595-1605.

Leips J, Mackay TFC. 2000. Quantitative trait loci for life span in *Drosophila melanogaster*: interactions with genetic background and larval density. *Genetics* 155: 1773-1788.

Leips J, Mackay TFC. 2002. The complex genetic architecture of *Drosophila* life span. *Experimental Aging Research* 28: 361-390.

Li Y, Tower J. 2009. Adult-specifc over-expression of the *Drosophila* genes *magu* and *hebe* increases life span and modulates late-age female fecundity. *Molecular Genetics and*

Genomics 281: 147-162.

Lin Y-J, Seroude L, Benzer S. 1998. Extended life-span and stress resistance in the *Drosophila* mutant *methuselah*. *Science* 282: 943-946.

Linnen C, Tatar M, Promislow D. 2001. Cultural artifacts: a comparison of senescence in natural, laboratory-adapted and artificially selected lines of *Drosophila melanogaster*. *Evolutionary Ecology Research* 3: 877-888.

Long AD, Lyman RF, Langley CH, Mackay TFC. 1998. Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* 149: 999-1017.

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

Luckinbill LS, Clare MJ. 1985. Selection for life span in *Drosophila melanogaster*. *Heredity* 55: 9-18.

Mackay TFC, Roshina NV, Leips JW, Pasyukova EG. 2006. Complex genetic architecture of *Drosophila* longevity. In: Masaro EJ, Austad SN, eds. Handbook of the Biology of Aging, Sixth Edition. Burlington, MA, Elsevier Press, 181-216.

Mackay TFC. 2002. The nature of quantitative genetic variation for *Drosophila* longevity. *Mechanisms of Ageing and Development* 123: 95-104.

Marden JH, Rogina B, Montooth KL, Helfand SL. 2003. Conditional tradeoffs between ageing and organismal performance of *Indy* long-lived mutant flies. *Proceedings of the National Academy of Sciences, USA* 100: 3369-3373.

Martin TE. 1995. Avian life history evolution in relation to nest sites, nest predation, and food. *Ecological Monographs* 65: 101-127.

Marygold SJ, Leevers SJ. 2002. Growth signaling: TSC takes its place. Current Biology 12: R785-R787.

Mattila J, Bremer A, Ahonen L, Kostiainen R, Puig O. 2009. *Drosophila* FoxO regulates organism size and stress resistance through adenylate cyclase. *Molecular and Cellular Biology* 29: 5357-5365.

Matzkin LM, Merritt TJS, Zhu C-T, Eanes WF. 2005. The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion in *In(3R)Payne* in *Drosophila melanogaster*. *Genetics* 170: 1143-1152.

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*.

Nature 351: 652-654.

Medawar PB. 1952. An unsolved problem of biology. London, HK Lewis & Company.

Minois N, Khazaeli AA, Curtsinger JW. 2001. Locomotor activity as a function of age and life span in *Drosophila melanogaster* overexpressing hsp70. Experimental Gerontology 36: 1137-1153.

Mitrovski P, Hoffmann AA. 2001. Postponed reproduction as an adaptation to winter conditions in *Drosophila melanogaster*: evidence for clinal variation under semi-natural conditions. *Proceedings of the Royal Society of London, Series B* 268: 2163-2168.

Miyashita NT, Aguade M, Langley CH. 1993. Linkage disequilibrium in the *white* locus region of *Drosophila melanogaster*. *Genetical Research* 62: 101-109.

Mockett RJ, Orr WC, Rahmandar JJ, Sohal BH, Sohal RS. 2001. Antioxidant status and stress resistance in long- and short-lived lines of *Drosophila melanogaster*. *Experimental Gerontology* 36: 441-463.

Mockett RJ, Sohal RS. 2006. Temperature-dependent trade-offs between longevity and fertility in the *Drosophila* mutant, *methuselah*. *Experimental Gerontology* 41: 566-573.

Morrow G, Samson M, Michaud S, Tanguay RM. 2004. Overexpression of the small mitochondrial Hsp22 extends *Drosophila* life span and increases resistance to oxidative stress. *FASEB Journal* 18(3): 598-599.

Mueller H-G, Wang J-L, Carey JR, Caswell-Chen EP, Chen C, Papadopoulos N, Yao F. 2004. Demographic window to aging in the wild: constructing life tables and estimating survival functions from marked individuals of unknown age. *Aging Cell* 3: 125-131.

Mullen LM, Hoekstra HE. 2008. Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution* 62: 1555-1570.

Norry FM, Loeschcke VR. 2002. Longevity and resistance to cold stress selected lines and their controls in *Drosophila melanogaster*. *Journal of Evolutionary Biology* 15: 775-783.

Nuzhdin SV, Khazaeli AA, Curtsinger JW. 2005. Survival analysis of life span quantitative trait loci in *Drosophila melanogaster*. *Genetics* 170: 719-731.

Nuzhdin SV, Pasyukova EG, Dilda CA, Zeng Z-B, Mackay TFC. 1997. Sex-specific quantitative trait loci affecting longevity in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA* 94: 9734-9739.

Nuzhdin SV, Pasyukova EG, Dilda CL, Zeng ZB, Mackay TFC. Sex-specific quantitative

trait loci affecting longevity in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA* 1997, 94: 9734-9739.

O'Neill MJ, Lawton BR, Mateos M, Carone DM, Ferreri GC, Hrbek T, Meredith RW, Reznick DN, O'Neill RJ. 2007. Ancient and continuing Darwinian selection on *insulin-like growth factor II* in placental fishes. *Proceedings of the National Academy of Sciences, USA* 104: 12404-12409.

Oakeshott JG, Gibson JB, Anderson PR, Knibb WR, Anderson DG, Chambers GK. 1982. Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. *Evolution* 36: 86-96.

Orr WC, Sohal RS. 1994. Extension of life span by overexpression of superoxide dismutase and catalase in *Drosophila melanogaster*. *Science* 263: 1128-1130.

Orr WC, Sohal RS. 2003. Does overexpression of Cu,Zn-SOD extend life span in *Drosophila melanogaster*? *Experimental Gerontology* 38: 227-230.

Paaby AB, Blacket MJ, Hoffmann AA, Schmidt PS. In Press. Identification of a candidate adaptive polymorphism for *Drosophila* life history by parallel independent clines on two continents. *Molecular Ecology*.

Paaby AB, Schmidt PS. 2008. Functional significance of allelic variation at *methuselah*, an aging gene in *Drosophila*. *PLoS ONE* 3: e1987.

Paaby AB, Schmidt PS. 2009. Dissecting the genetics of longevity in *Drosophila melanogaster*. *Fly* 3: 29-38.

Parkes TL, Elia AJ, Dickinson D, Hilliker AJ, Phillips JP, Boulianne GL. 1998. Extension of *Drosophila* lifespan by overexpression of human *SOD1* in motoneurons. *Nature Genetics* 19: 171-174.

Partridge L, Barrie B, Fowler K, French V. 1994. Evolution and development of body size and cell size in *Drosophila melanogaster* in response to temperature. *Evolution* 48: 1269-1276.

Partridge L, Gems D. 2002. Mechanisms of aging: public or private? *Nature Reviews Genetics* 3: 165-175.

Partridge L, Gems D. 2002. The evolution of longevity. *Current Biology* 12: R544-R546.

Partridge L, Gems D, Wither DJ. 2005. Sex and death: what is the connection? *Cell* 120: 461-462.

Partridge L, Mangel M. 1999. Messages from mortality: the evolution of death rates in the

old. *Trends in Ecology and Evolution* 14: 438-442.

Partridge L, Piper MDW, Mair W. 2005. Dietary restriction in *Drosophila*. *Mechanisms of Ageing and Development* 126: 938-950.

Partridge L, Prowse N, Pignatelli P. 1999. Another set of responses and correlated responses to selection on age at reproduction in *Drosophila melanogaster*. *Proceedings of the Royal Society of London, Series B* 266: 255-261.

Pasyukova EG, Roshina NV, Mackay TFC. 2004. *Shuttle craft*: a candidate quantitative trait gene for *Drosophila lifespan*. *Aging Cell* 3: 297-307.

Pasyukova EG, Vieira C, Mackay TFC. 2000. Deficiency mapping of quantitative trait loci affecting longevity in *Drosophila melanogaster*. *Genetics* 156: 1129-1146.

Pétavy G, Morin JP, Moreteau B, David JR. 1997. Growth temperature and phenotypic plasticity in two *Drosophila* sibling species: probable adaptive changes in flight capacities. *Journal of Evolutionary Biology* 10: 875–887.

Petrosyan A, Hsieh I-H, Saberi K. 2007. Age-dependent stability of sensorimotor functions in the life-extended *Drosophila* mutant *Methuselah*. *Behavioral Genetics* 37: 585-594.

Pletcher SD, Macdonald SJ, Marguerie R, Certa U, Sterns SC, Goldstein DB, Partridge L. 2002. Genome-wide transcript profiles in aging and calorically restricted *Drosophila melanogaster*. *Current Biology* 12: 712-723.

Powell JR. 1974. Temperature related divergence in *Drosophila* size. *Journal of Heredity* 65: 257-258.

Promislow DEL, Tatar M, Khazaeli AA, Curtsinger JW. 1996. Age-specific patterns of genetic variance in *Drosophila melanogaster*. I. Mortality. *Genetics* 143: 839-848.

Promislow DEL, Tatar M. 1998. Mutation and senescence: where genetics and demography meet. *Genetica* 102/103: 299-314.

Puig O, Marr MT, Ruhf ML, Tjian R. 2003. Control of cell number by *Drosophila* FOXO: downstream and feedback regulation of the insulin receptor pathway. *Genes and Development* 17: 2006-2020.

Rako L, Anderson AR, Sgro CM, Stocker AJ, Hoffmann AA. 2006. The association between inversion *In(3R)Payne* and clinally varying traits in *Drosophila melanogaster*. *Genetica* 128: 373-384.

Rako L, Blacket MJ, McKechnie SW, Hoffmann AA. 2007. Candidate genes and thermal

phenotypes: identifying ecologically important genetic variation for thermotolerance in the Australian *Drosophila melanogaster* cline. *Molecular Ecology* 16: 2948-2957.

Reznick D, Bryga H, Endler JA. 1990. Experimentally induced life-history evolution in a natural population. *Nature* 346: 357-359.

Reznick D. 1985. Costs of reproduction: an evaluation of the empirical evidence. *Oikos* 44: 257-267.

Roff DA. 1992. The Evolution of Life Histories: Theory and Analysis. New York, Chapman and Hall.

Rogina B, Helfand SL, Frankel S. 2002. Longevity regulation by *Drosophila* Rpd3 deacetylase and caloric restriction. *Science* 298: 1745.

Rogina B, Helfand SL. 2004. Sir2 mediates longevity in the fly through a pathway related to calorie restriction. *Proceedings of the National Academy of Sciences, USA* 101: 15998-16003.

Rogina B, Reenan RA, Nilsen SP, Helfand SL. 2000. Extended life-span conferred by cotransporter gene mutations in *Drosophila*. *Science* 290: 2137-2140.

Rohlf FJ. 2008. tpsDig2, version 2.12. Available from:

http://life.bio.sunysb.edu/morph/morphmet/tpsdig2w32.exe.

Rose M, Charlesworth B. 1980. A test of evolutionary theories of senescence. *Nature* 287: 141-142.

Rose MR, Charlesworth B. 1981. Genetics of life history in *Drosophila melanogaster*. II. Exploratory selection experiments. *Genetics* 97: 187-196.

Rose MR, Vu LN, Park SU, Graves JL, Jr. 1992. Selection on stress resistance increases longevity in *Drosophila melanogaster*. *Experimental Gerontology* 27: 241-250.

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

Rose MR. 1991. Evolutionary Biology of Aging. New York, Oxford University Press.

Rozas J, Rozas R. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174-175.

Ruan H, Tang XD, Chen M-L, Joiner MA, Sun G, Brot N, Weissbach H, Heinemann SH, Iverson L, Wu C-F, Hoshi T. 2002. High-quality life extension by the enzyme peptide methionine sulfoxide reductase. *Proceedings of the National Academy of Sciences, USA* 99:

2748-2753.

Schemske DW, Bierzychudek P. 2007. Spatial differentiation for flower color in the desert annual Linanthus parryae: was Wright right? *Evolution* 61: 2528-2543.

Schmidt PS, Conde DR. 2006. Environmental heterogeneity and the maintenance of genetic variation for reproductive diapause in *Drosophila melanogaster*. *Evolution* 60: 1602-1611.

Schmidt PS, Duvernell DD, Eanes WF. 2000. Adaptive evolution of a candidate gene for aging in *Drosophila*. *Proceedings of the National Academy of Sciences, USA* 97: 10861-10865.

Schmidt PS, Matzkin L, Ippolito M, Eanes WF. 2005a. Geographic variation in diapause incidence, life-history traits, and climatic adaptation in *Drosophila melanogaster*. *Evolution* 59: 1721-1732.

Schmidt PS, Paaby AB, Heschel MS. 2005b. Genetic variance for diapause expression and associated life histories in *Drosophila melanogaster*. *Evolution* 59: 2616-2625.

Schmidt PS, Paaby AB. 2008. Reproductive diapause and life history clines in North American populations of *Drosophila melanogaster*. *Evolution* 62: 1204-1215.

Schmidt PS, Zhu C-T, Das J, Batavia M, Yang L, Eanes WF. 2008. An amino acid polymorphism in the *couch potato* gene forms the basis for climatic adaptation in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA* 105: 16207-16211.

Seong K-H, Matsuo T, Fuyama Y, Aigaki T. 2001. Neural-specific overexpression of *Drosophila Plenty of SH3s* (*DPOSH*) extends the longevity of adult flies. *Biogerontology* 2: 271-281.

Seong K-H, Ogashiwa T, Fuyama Y, Aigaki T. 2001. Application of the gene search system to a screen for longevity genes in *Drosophila*. *Biogerontology* 2: 209-217.

Seroude L, Brummel T, Kapahi P, Benzer S. 2002. Spatio-temporal analysis of gene expression during aging in *Drosophila melanogaster*. *Aging Cell* 1: 47-56.

Service PM, Hutchinson EW, MacKinley MD, Rose MR. 1985. Resistance to environmental stress in *Drosophila melanogaster* selected for postponed senescence. *Physiological Zoology* 58: 380-389.

Service PM. 1987. Physiological mechanisms of increased stress resistance in *Drosophila melanogaster* selected for postponed senescence. *Physiological Zoology* 60: 321-326.

Sezgin E, Duvernell D, Matzkin LM, Duan Y, Zhu C-T, Verrelli BC, Eanes WF. 2004.

Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in *Drosophila melanogaster*. *Genetics* 168: 923-931.

Sgro CM, Patridge L. 1999. A delayed wave of death from reproduction in *Drosophila*. *Science* 286: 2521-2524.

Shepherd JCW, Walldorf U, Hug P, Gehring WJ. 1989. Fruit flies with additional expression of elongation factor EF-1α live longer. *Proceedings of the National Academy of Sciences, USA* 86: 7520-7521.

Sim C, Denlinger DL. 2008. Insulin signaling and FOXO regulate the overwintering diapause of the mosquito *Culex pipiens*. *Proceedings of the National Academy of Sciences, USA* 105: 6777-6781.

Simon AF, Shih C, Mack A, Benzer S. 2003. Steroid control of longevity in *Drosophila melanogaster*. *Science* 299: 1407-1410.

Song W, Ranjan R, Dawson-Scully K, Bronk P, Marin L, Serounde L, Lin YJ, Nie Z, Atwood HL, Benzer S, Zinsmaier KE. 2002. Presynaptic regulation of neurotransmission in *Drosophila* by the G protein-coupled receptor *methuselah*. *Neuron* 36: 105-119.

Spencer CC, Howell CE, Wright AR, Promislow DEL. Testing an 'aging gene' in long-lived

Drosophila strains: increased longevity depends on sex and genetic background. *Aging Cell* 2: 123-130.

Spencer CC, Promislow DEL. 2005. Age-specific changes in epistatic effects on mortality rate in *Drosophila melanogaster*. *Journal of Heredity* 96: 513-521.

Stalker HD. 1980. Chromosome studies in wild populations of *Drosophila melanogaster.* II. Relationships of inversion frequencies to latitude, season, wing loading and flight activity. *Genetics* 95: 211-223.

Stearns SC. 1991. Trade-offs in life-history evolution. *Functional Ecology* 3: 259-268.

Stearns SC. 1992. The Evolution of Life Histories. New York, Oxford University Press.

Suh Y, Atzmon G, Cho M-O, Hwang D, Liu Bingrong, Leahy DJ, Barzilai N, Cohen P. 2008. Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proceedings of the National Academy of Sciences, USA* 105: 3438-3442.

Sun J, Tower J. 1999. FLP recombinase-mediated induction of Cu/Zn-superoxide dismutase transgene expression can extend the life span of adult *Drosophila melanogaster* flies. *Molecular and Cellular Biology* 19: 216-228.

Symphorien S, Woodruff RC. 2003. Effect of DNA repair on aging of transgenic *Drosophila melanogaster*: I. *mei-41* locus. *Journal of Gerontology, Series A: Biological Scienes and Medical Sciences* 58: 782-787.

Tajima F. 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

Tatar M, Bartke A, Antebi A. 2003. The endocrine regulation of aging by insulin-like signals. *Science* 299: 1346-1351.

Tatar M, Chien SA, Priest NK. 2001. Negligible senescence during reproductive dormancy in *Drosophila melanogaster*. *American Naturalist* 158: 248-258.

Tatar M, Khazaeli AA, Curtsinger JW. 1997. Chaperoning extended life. *Nature* 390: 30.

Tatar M, Kopelman A, Epstein D, Tu M-P, Yin C-M, Garofalo RS. 2001. A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function. *Science* 292: 107-110.

Tatar M, Promislow DEL, Khazaeli AA, Curtsinger JW. 1996. Age-specific patterns of genetic variance in *Drosophila melanogaster* II. Fecundity and its genetic covariance with age-specific mortality. *Genetics* 143: 849-858.

Tatar M, Yin C. 2001. Slow aging during insect reproductive diapause: Why butterflies, grasshoppers and flies are like worms. *Experimental Gerontology* 36: 723-738.

Tatar, M. 2005. Comment on "Long-lived *Drosophila* with overexpressed dFOXO in adult fat body." *Science* 307: 675.

Tauber E, Zordan M, Sandrelli F, Pegoraro M, Osterwalder N, Breda C, Daga A, Selmin A, Monger K, Benna C, Rosato E, Kyriacou CP, Costa R. 2007. Natural selection favors a newly derived *timeless* allele in *Drosophila melanogaster*. *Science* 316: 1895-1898.

Toivonen JM, Partridge L. 2009. Endocrine regulation of aging and reproduction in *Drosophila*. *Molecular and Cellular Endocrinology* 299: 39-50.

Toivonen JM, Walker GA, Martinez-Diaz P, Bjedov I, Driege Y, Jacobs HT, Gems D, Partridge L. 2007. No influence of *Indy* on lifespan in *Drosophila* after correction for genetic and cytoplasmic background effects. *PLoS Genetics* 3: e95.

Trotta V, Calboli FCF, Ziosi M, Gerra D, Pezzoli MC, David JR, Cavicchi S. 2006. Thermal plasticity in *Drosophila melanogaster*: a comparison of geographic populations. *BMC Evolutionary Biology* 6: 67-79.

Tu M-P, Epstein D, Tatar M. 2002. The demography of slow aging in male and female

Drosophila mutant for the insulin-receptor substrate homologue chico. *Aging Cell* 1: 75-80.

Tu M-P, Flatt T, Tatar M. 2006. Juvenile and steroid hormones in *Drosophila melanogaster* longevity. In: Masaro EJ, Austad SN, eds. Handbook of the Biology of Aging, Sixth Edition. Burlington, MA, Elsevier Press, 415-448.

Turner TL, Levine MT, Eckert ML, Begun DJ. 2008. Genomic analysis of adaptive differentiation in *Drosophila melanogaster*. *Genetics* 179: 455-473.

Valenzuela RK, Forbes SN, Keim P, Service PM. 2004. Quantitative trait loci affecting life span in replicated populations of *Drosophila melanogaster*. II. Response to selection. *Genetics* 168: 313-324.

van't Land J, van Putten P, Zwaan B, Kamping A, van Delden W. 1999. Latitudinal variation in wild populations of *Drosophila melanogaster*: heritabilities and reaction norms. *Journal of Evolutionary Biology 12*: 222-232.

Vermeulen CJ, Loeschcke L. 2006. Longevity and the stress response in *Drosophila*. *Experimental Gerontology* 42: 153-159.

Vieria C, Pasyukova EG, Zeng Z-B, Hacket JB, Lyman RF, Mackay TFC. 2000. Genotype-environment interaction for quantitative trait loci affecting life span in *Drosophila*

melanogaster. *Genetics* 154: 213-227.

Vihervaara T, Puig O. 2008. dFOXO regulates transcription of a *Drosophila* acid lipase. *Journal of Molecular Biology* 376: 1215-1223.

Wallenfang MR, Nayak R, DiNardo S. 2006. Dynamics of the male germline stem cell population during aging of *Drosophila* melanogaster. *Aging Cell* 5: 297-304.

Wang H-D, Kazemi-Esfarjani P, Benzer S. 2004. Multiple-stress analysis for isolation of *Drosophila* longevity genes. *Proceedings of the National Academy of Sciences, USA* 101: 12610-12615.

Wang M-H, Lazebny O, Harshman LG, Nuzhdin SV. 2004. Environment-dependent survival of *Drosophila melanogaster*: a quantitative genetic analysis. *Aging Cell* 3: 133-140.

Wang MC, Bohmann D, Jasper H. 2003. JNK signaling confers tolerance to oxidative stress and extends lifespan in *Drosophila*. *Developmental Cell* 5: 811-816.

Wang MC, Bohmann D, Jasper H. 2005. JNK extends life span and limits growth by antagonizing cellular and organism-wide responses to insulin signaling. *Cell* 121: 115–125.

Wayne ML, Soundararajan U, Harshman LG. 2006. Environmental stress and reproduction in

Drosophila melanogaster: starvation resistance, ovariole numbers and early age egg production. *BMC Evolutionary Biology* 6: 57.

Weeks AR, McKechnie SW, Hoffmann AA. 2002. Dissecting adaptive clinal variation: markers, inversions and size/stress associations in *Drosophila melanogaster* from a central field population. *Ecology Letters* 5: 756-763.

Weindruch RH, Walford RL, Fligiel S, Guthrie D. 1986. The retardation of aging in mice by dietary restriction: longevity, cancer, immunity, and lifetime energy intake. *Journal of Nutrition* 116: 641-654.

West AP, Jr, Llamas LL, Snow PM, Benzer S, Bjorkman PJ. 2001. Crystal structure of the ectodomain of Methuselah, a *Drosophila* G protein-coupled receptor associated with extended lifespan. *Proceedings of the National Academy of Sciences, USA* 98: 3744-3749.

Willcox BJ, Donion TA, He Q, Chen R, Grove JS, Yano K, Masaki KH, Willcox DC, Rodriguez B, Curb JD. 2008. FOXO3A genotype is strongly associated with human longevity. *Proceedings of the National Academy of Sciences, USA* 105: 13987-13992.

Williams GC. 1957. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11: 398-411.

Williams KD, Busto M, Suster ML, So AK-C, Ben-Shahar Y, Leevers SJ, Sokolowki M. 2006. Natural variation in *Drosophila melanogaster* diapause due to the insulin-regulated PI3-kinase. *Proceedings of the National Academy of Sciences, USA* 103: 15911-15915.

Wilson RH, Morgan TJ, Mackay TFC. 2006. High-resolution mapping of quantitative trait loci affecting increased life span in *Drosophila melanogaster*. *Genetics* 173: 1455-1463.

Zou S, Meadows S, Sharp L, Jan LY, Jan YN. 2000. Genome-wide study of aging and oxidative stress response in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA* 97: 13726-13731.

Zwaan B, Bijlsma R, Hoekstra RF. 1995. Direct selection on life span in *Drosophila melanogaster*. *Evolution* 49: 649-659.

Zwaan B. 1999. The evolutionary genetics of ageing and longevity. *Heredity* 82: 589-597.

Zwaan BJ, Bijlsma R, Hoekstra RF. 1991. On the developmental theory of ageing. 1. Starvation resistance and longevity in *Drosophila melanogaster* in relation to pre-adult breeding conditions. *Heredity* 66: 29-39.