Stony Brook University

OFFICIAL COPY

The official electronic file of this thesis or dissertation is maintained by the University Libraries on behalf of The Graduate School at Stony Brook University.

© All Rights Reserved by Author.

Population Genetics and Experiments in Metabolic Enzymes of *Drosophila melanogaster*

A Dissertation Presented

by

Erik Lavington

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Genetics

Stony Brook University

August 2015

Stony Brook University

The Graduate School

Erik Lavington

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Walter Eanes – Dissertation Advisor

Professor and Chair, Department of Ecology and Evolution

John True – Chairperson of Defense

Associate Professor, Department of Ecology and Evolution

J. Peter Gergen – Committee Member

Professor, Department of Biochemistry and Cell Biology

Lawrence Harshman – Committee Member

Professor, School of Biological Sciences

University of Nebraska-Lincoln

This dissertation is accepted by the Graduate School

Charles Taber

Dean of the Graduate School

Abstract of the Dissertation

Population Genetics and Experiments in Metabolic Enzymes of *Drosophila melanogaster*

by

Erik Lavington

Doctor of Philosophy

in

Graduate Program in Genetics

Stony Brook University

2015

Control of metabolic flux, the flow of metabolites through a complex metabolic network, is of importance in understanding how an organism is sensing and responding to nutrient changes in its environment. Metabolic flux control can be measured for, and a control coefficient assigned to, each enzyme in a pathway. Measuring metabolic flux control in multicellular organisms is complicated by the fact that nutrient sensing and metabolic flux control may vary by tissue type. Major effects should be detectable in genomic information, as enzymes with high control coefficients will exhibit genetic patterns of adaptation when the pathway is under selection pressure. I used genetic variation within and among populations of *Drosophila melanogaster*, as well as divergence between *D. melanogaster* and the closely related *D. simulans*, to identify candidate genes for experimental study. I then conducted experiments with candidate genes using tissue specific RNA interference knockdown, focusing on two enzymes comprising the glycerophosphate shuttle in the context of starvation resistance, adipokinetic hormone (AKH) signaling, the *Drosophila* analog of glucagon signaling, and Insulin/Insulin-like signaling. None of the genes that I studied had a significant effect on starvation resistance when knocked down in Insulin-like protein secreting cells. I found that glycerophosphate oxidase, but not glycerophosphate dehydrogenase, significantly increased the average time to death in starvation conditions when knocked down in AKH secreting cells. Because the glycerophosphate shuttle is important in transferring nicotinamide adenine dinucleotide equivalents between the cytosol and inner matrix of the mitochondrion, this result implicates the coupling of reduction-oxidation state with AKH signaling.

Table of Contents

List of Figures and Tables — Chapter 1

List of Figures and Tables — Chapter 2

List of Figures and Tables — Chapter 3

Introduction

The purpose of this dissertation was to use evidence of adaptation in the metabolic network of *Drosophila melanogaster* to find candidate genes for further testing in the context of a phenotype in which metabolism is expected to play a central role. I used a novel approach correlating expression data from one population with SNP allele frequencies for 20 populations along a latitudinal gradient. I compared these results with three widely used tests of selection to determine what, if any, patterns are shared between tests. I then tested candidate genes as to their effect, if any, on starvation resistance using tissue specific knockdown of gene activity.

Experiments designed to find candidate genes for more detailed study are typically strictly forward genetics screen on a particular phenotype. Candidate genes are those that have a major effect on the phenotype among all genes tested, often most of the genome. How the candidates arise are a function of the experimental design, from the type and amount of genetic variation to the phenotypic measures. Knock-out lines of deletions covering the genome in an otherwise isogenic background work well to determine the structure and components of a network but ignore population variation. Quantitative trait mapping using inbred lines uses population variation but for two, and only two, lines.

Tests of genetic patterns of adaptation have been developed for use prior to the relatively new era of affordable high-throughput sequencing has been applied to new genomic data sets. Fay et al (FAY *et al.* 2002) used the amino acid replacement to synonymous substitution ratio (A/S) and assumptions of the McDonald-Kreitman tests on 45 genes in *Drosophila melanogaster* to determine whether positive selection is acting on particular loci, in contrast to demographic events which effect all loci similarly. Shapiro et al (2007) took this further with 419 genes and Mackay et al (MACKAY *et al.* 2012) performed McDonald-Kreitman (MK) tests (MCDONALD and KREITMAN 1991) across the entire genome of *Drosophila melanogaster* for both coding and modified MK tests for non-coding regions (JENKINS *et al.* 1995; ANDOLFATTO 2005; EGEA *et al.* 2008). Studies using Fst outlier tests are quite common, as are software resources (BEAUMONT and NICHOLS 1996; ANTAO *et al.* 2008; FOLL and GAGGIOTTI 2008). Hohenlohe et al (2010) used both Fst scans between three populations of fresh water and marine three-spine sticklebacks. They

found that several regions of significantly increased or decreased Fst overlapped when comparing the three freshwater populations and that some of these significant outlier regions overlapped with previously described QTL linkage groups.

In contrast to commonly used methods using evidence of adaptation to find candidate genes for further testing, I used a novel approach focusing on a limited set of genes, rather than the whole genome, correlating expression data from one population with SNP allele frequencies for 20 populations along a latitudinal gradient. While it has been common to use one or two tests of selection, I also compared these results with three widely used tests of selection to determine what, if any, patterns are shared between tests. I then tested candidate genes as to their effect, if any, on starvation resistance using tissue specific knockdown of gene activity.

Metabolism is of central importance to life. Understanding the control of flux through metabolic pathways is of interest for a broad scope of study. It is believed that most enzymes in the pathway have little effect on flux when their activity is modified and therefore have little *control* of flux (Fell 1998; Olson-Manning *et al.* 2013). Observations of large whole metabolic shifts or individual metabolic enzyme activity associations with pathology have moved, with our increased understanding, from assayable effects to defining causes and targets of treatment (Martins *et al.* 2006; Kroemer and Pouyssegur 2008; Fontana *et al.* 2010, to name a few). For example, cancer metabolism includes the founding, and the subsequent modification, of the Warburg effect which was initially thought to be the down regulation of mitochondrial oxidative phosphorylation, but now is accepted as the up regulation of the glycolytic pathway (Bensinger and Christofk 2012). In addition, metabolic enzymes have become potential targets for drug therapies alongside classical signaling pathways (Wu and Zhao 2013; Losman and Kaelin 2013). Research in longevity has implicated metabolic enzymes indirectly as downstream targets of the Insulin/Insulin-like signaling pathway (Partridge *et al.* 2011) and lifespan studies in our lab with *Drosophila melanogaster* have also shown direct effects of central metabolic enzyme activity on lifespan (Talbert et al., in review). Consequently, a better understanding of metabolic flux control and the role of metabolic enzymes in life history traits promises to yield insights into how these genes may be used as targets to treat human disease.

Experimental studies to evaluate flux control through a metabolic network require a great deal of prior knowledge and complex experiments for each enzyme. The goal of such metabolic control analysis is to measure the "control coefficients" for each enzyme by perturbing enzyme activity and assessing flux or some proxy. The control exercised by a step enzyme is calculated as the change in defined output of the network (flux) given the either increased or reduced activity of that enzyme. To understand the relative control across an entire network, this must be repeated for a range of perturbations for each enzyme (Fell 1992).

In an example of estimating the flux control from sucrose to starch in a potato tuber (Geigenberger *et al.* 2004) used data collected from several studies to determine control based on an approximate control coefficient. They calculated the control coefficients from each of the studied enzymes. The data from experiments using tubers or isolated discs of tuber tissue were all normalized to respective full activity ("wild-type") levels for comparison across all of the enzymes. The control coefficients were calculated as the slope of curves near full activity of the enzyme, with the expectation that most would be small. With only limited information about the structure of the network, one might expect flux from sucrose to starch to be controlled at a step closer to the branch points of sucrose or glucose-6-phosphate (G6P) and not the amyloplastidial ATP/ADP translocator (AATPT), yet this is what they found. This counterintuitive conclusion would also be missed if one were to focus on only the direct enzymatic pathway from sucrose to starch, thus highlighting the importance of including as many steps involved in the measured flux as possible.

Given the difficulty of experimentally quantifying flux control even for a single enzyme through a network as large and complex as central metabolism, it would be important to utilize other observations to produce a candidate list of enzymes that are most likely to control flux. We might assume that natural selection acts upon the flux of a pathway and the response to selection and associated evidence appears in the genes with the largest control coefficients (Eanes 2011). Again, we would expect that most enzymes have low control coefficients (Kacser and Burns 1981) and would not stand out when looking at molecular data. Dykhuizen et al. (Dykhuizen *et al.* 1987) first combined metabolic flux control theory with fitness measures (assayed as growth rate) to predict the control coefficients of two enzymes involved in lactose metabolism in *Escherichia coli,* β-galactosidase and β-galactoside permease. Using chemostats, with lactose as the limiting nutrient, they measured relative fitness of strains with known enzyme activities. Except for genetic variants

with very low β-galactosidase activity, they were able to predict fitness of *E.coli* strains via flux control (most showed low control coefficients). Changes in β-galactoside permease, having a much larger control coefficient than β-galactosidase, resulted in large changes in fitness whereas even very large changes in β-galactosidase activity resulted in only small changes in fitness until activity levels were extremely low. Consequently, if the environment favored shifting flux in the pathway one might expect genetic variation in expression to appear at the permease step.

With an association between fitness and metabolic flux control, we can connect flux control with molecular signatures of selection within a network of interest. Given the central importance of metabolism, we could reasonably expect variation in glycolytic flux to be a target of selection for many traits, including starvation resistance, lifespan, fertility and fecundity. However, the presence or absence of flux control can be inferred in cases where the footprint of selection can be observed in molecular data (Eanes 2011). For example, when Flowers et al (2007) investigated patterns of adaptive evolution in glycolytic enzymes of *D.melanogaster* they found these signals clustered around the major branch point glucose-6-phosphate. While current high-throughput sequencing methods have made it relatively easy to scan the entire genomes of both model and non-model organisms for genetic variation (Stinchcombe and Hoekstra 2008; Fabian *et al.* 2012) such a wide view over networks as large and complex as metabolism would be expected to yield large numbers of non-interacting targets. We chose to focus on a particular portion of the metabolic pathway to find candidate genes that may have local control so as to generate fine-grain analyses.

With this focus on the glycolytic pathway, our lab has used data from high-throughput sequencing techniques and the population genetics of the system to predict the enzymes that are under adaptive selection in a relatively recent evolutionary time scale in *Drosophila melanogaster*. This may identify loci that we suspect have significant flux control (Eanes 2011). The populations in our study were sampled along a latitudinal gradient on the eastern coast of North America, from the southern end of Florida to Maine, as well as from further inland in Sudbury, Ontario. This historically tropical species has adapted to temperate climates since its colonization in North America less than 200 years ago, and maintains stable local populations in the northern as well as the southern parts of its range (Reaume and Sokolowski 2006). Many of the genes in this pathway possess SNPs that vary in a latitude-dependent fashion (Sezgin *et al.* 2004) that posits a gradient in local tropical-temperate selection. Specifically, we used sequence data from the DGRP lines

(Mackay *et al.* 2012) and calculated an allele effect on cis-acting whole-fly adult gene expression (Ayroles *et al.* 2009) for each SNP. We then used bulk pyrosequencing to estimate SNP allele frequencies across the 20 populations (Lavebratt and Sengul 2006; Doostzadeh *et al.* 2008). Results from a principal component analysis of the variance-covariance matrix of population SNP-associated expression variation across the north-south cline provided a multivariate description of patterns of expression change with latitude. It predicted a general down-regulation of enzymes at the entrance to the glycolytic pathway and an up-regulation of cofactor shuttles, *Gpdh* and *Gpo*-1, as well as on the mitochondrial side of the glutamate-aspartate shuttle. The genes studied here, *Gpdh*, *Gpo-1*, and *HexA*, are significant players in that analysis.

Additionally, experiments in our lab (Talbert et al., in revision.) have shown that metabolic enzyme activity perturbation can have a significant effect on lifespan, oxidative stress resistance and starvation resistance. The experiments were performed with lines of *D. melanogaster* that compared 'full' activity against 'reduced' activity genotypes that were generated by P-element mutagenesis. For each gene, many mutant alleles were screened and 'full' activity and 'knockdown' or 'knockout' alleles were saved, assayed, and placed in identical genetic backgrounds using a controlled cross design(Merritt *et al.* 2006). Because dietary restriction is an important feature in life span extension (CHIPPINDALE *et al.* 1993; PARTRIDGE *et al.* 2005; TATAR 2011; PIPER *et al.* 2014), a diet effect was also tested in the lifespan experiments. Results show a significant effect of *Gpdh*, *HexA*, *Gdh*, and *Men* activity on lifespan for one or both diets. Reduction of *Idh* activity had no effect on lifespan, even though the expected dietary effect was still observed. *Mdh2* activity perturbation had a significant effect on low nutrient food, but not with high nutrient food. Also of note is that there was no diet effect (wherein flies are expected to be longer lived on low nutrient food) (PARTRIDGE *et al.* 1987; CHIPPINDALE *et al.* 1993), of *Gpdh* genotypes. The results of starvation resistance are all significant. The result of *HexA* hints at a possible trade-off between longevity and starvation resistance. It is of special interest that *Gpdh* showed a significant extension of life span in the genotype with lowered activity. Reduced NAD^+/NADH ratio was first associated with starvation in rat livers(KREBS 1967) and later confirmed in *D.melanogaster* (ZHU and RAND 2012). *Gpdh* and *Gpo-1* represent the cytosolic and mitochondrial, respectively, sides of the glycerol-phosphate shuttle that is used to transfer

NAD/NADH equivalents and thereby should play a direct role in the reduction-oxidation (redox) state of the cell (O'Brien and MacIntyre 1972).

Starvation resistance is an important phenotype in wild *Drosophila.* Heritability for starvation resistance has been reported as near 1.0 (SERVICE and ROSE 1985), indicating that the vast majority of variation is genetic. Selection experiments have also shown that there is great genetic variation for starvation resistance in wild populations (HOFFMANN and HARSHMAN 1999). While latitudinal variation for starvation resistance has been seen in Indian populations (KARAN *et al.* 1998), it has not found in South America or Australia (ROBINSON *et al.* 2000; HOFFMANN *et al.* 2005). Although high heritability and lack of latitudinal variation outside of India may diminish the importance of starvation resistance in natural populations, starvation resistance remains an attractive study system to test candidate genes given and association of metabolism to starvation resistance in QTL studies (HARBISON 2004; AYROLES *et al.* 2009).

This study investigates what effect, if any, perturbation of the metabolic pathway has in the context of adipokinetic hormone (AKH) signaling during starvation. AKH producing cells in the corpora cardiaca (CC) are required for starvation induced hyperactivity (Lee and Park 2004). Furthermore, metabolic state sensing has been linked to AKH release from the CC cells in response to starvation through AMPK (Braco *et al.* 2012). *HexA* uses ATP as a cofactor and is near the major branch point of glucose-6-phosphate (Flowers *et al.* 2007) and *Gpdh*, and *Gpo-1* are partners in the glycerophosphate shuttle and use NAD(P) and FAD as cofactors, respectively. Position in the overall pathway, the use of highly connected cofactors and predicted metabolic flux control makes these genes attractive candidates to affect starvation response by control of metabolic flux and therefore metabolic state.

To test the effect of candidate genes in AKH signaling starvation response, I have used tissue specific RNAi knockdown using RNAi constructs from the Transgenic RNAi Project (TRiP) under the control of *GAL4* UAS and *GAL4* drivers expressed under the control of *Akh* (Lee and Park 2004) and *dILP2* (Rulifson *et al.* 2002) *Drosophila melanogaster* promoters. Insulin signaling is known to play a role in starvation resistance (Clancy *et al.* 2001; Ikeya *et al.* 2002; Broughton *et al.* 2005; Wang *et al.* 2008) although coupling of metabolic state with dILP release lies outside of the dILP producing cells (Geminard *et al.* 2009). The assay used is the same as used with the P-

element lines, but this tests the effect of enzyme activity knockdown, by RNAi knockdown, on cells that are required for a starvation-induced behavior. The P-element excision lines reduce activity in all cells at all developmental times and could reflect the effect of perturbation of the *response* to a change in AKH signaling, rather than a direct effect on AKH signaling (e.g. timing or magnitude) (BHARUCHA *et al.* 2008).

Acknowledgements:

Rodrigo Cogni and Spencer Koury collected most of the population of flies in collaboration with Paul Schmidt and his students, Emily Behrman and Katherine O'Brien. Thomas Merritt donated the Sudbury, ON population. John True and Joe LaChance donated populations from New York. The population from Raleigh, NC is the DGRP and live flies were acquired from the Bloomington Drosophila Stock Center. Fly husbandry to generate isofemale lines was performed by Maria Amella and Spencer Koury. Fly husbandry to generate the *Akh-GAL4* and *dILP2-GAL4* lines was performed by Maria Amella and Matt Talbert. Pyrosequencing primer design and sample preparation was performed by Kate Kuczynski and Rodrigo Cogni. I collected DGRP sequence and expression data and prepared sequence and expression data for our analyses. I designed and set up, including all fly husbandry, all experiments in Chapter 3. Eugene Brud and Spencer Koury helped to collect data during starvation experiments. I ran all data analyses.

Chapter 1 of this dissertation was originally published by Oxford University Press in Molecular Biology and Evolution, volume 31, pages 2032-2041 and is referred to as "Lavington et al (2014)" throughout the remainder of this dissertation.

Chapter 1 - A Small System—High-Resolution Study of Metabolic Adaptation in the Central Metabolic Pathway to Temperate Climates in Drosophila melanogaster

Erik Lavington,1 Rodrigo Cogni,z,1 Caitlin Kuczynski,1 Spencer Koury,1 Emily L. Behrman,2 Katherine R. O'Brien,2 Paul S. Schmidt,2 and Walter F. Eanes*,1

1Department of Ecology and Evolution, Stony Brook University

2Department of Biology, University of Pennsylvania

zPresent address: Department of Genetics, University of Cambridge, Cambridge, United Kingdom

*Corresponding author: E-mail: walter.eanes@stonybrook.edu.

Associate editor: Michael Purugganan

Abstract

In this article, we couple the geographic variation in 127 single-nucleotide polymorphism (SNP) frequencies in genes of 46 enzymes of central metabolism with their associated cis-expression variation to predict latitudinal or climatic-driven gene expression changes in the metabolic architecture of Drosophila melanogaster. Forty-two percent of the SNPs in 65% of the genes show statistically significant clines in frequency with latitude across the 20 local population samples collected from southern Florida to Ontario. A number of SNPs in the screened genes are also associated with significant expression variation within the Raleigh population from North Carolina. A principal component analysis of the full variance–covariance matrix of latitudinal changes in SNP-associated standardized gene expression allows us to identify those major genes in the pathway and its associated branches that are likely targets of natural selection. When embedded in a central metabolic context, we show that these apparent targets are concentrated in the genes of the upper glycolytic pathway and pentose shunt, those controlling glycerol shuttle activity, and finally those enzymes associated with the utilization of glutamate and pyruvate. These metabolites possess high connectivity and thus may be the points where flux balance can be best shifted. We also propose that these points are conserved points associated with coupling energy homeostasis and energy sensing in mammals. We speculate that the modulation of gene expression at specific points in central metabolism that are associated with shifting flux balance or possibly energy-state sensing plays a role in adaptation to climatic variation.

Key words: energy sensing, metabolism, life history, clines, gene expression variation.

Introduction

A goal of modern evolutionary genetics should be to integrate our understanding of the causes of genetic and molecular variation among genes into larger functional contexts. The promise of exploring adaptive natural selection on all genes in the genome (Clark et al. 2007; Greenberg et al. 2008; Andres et al. 2009; Liti et al. 2009) has led to the expectation that we can predict how the inherent roles, properties, network, and pathway context of enzymes determine their relative participation in adaptive evolution (Lu and Rausher 2003; Cork and Purugganan 2004; Flowers et al. 2007, 2009; Alvarez-Ponce

et al. 2009, 2012; Obbard et al. 2009; Ramsay et al. 2009; Jumbo-Lucioni et al. 2010; Eanes 2011; Clark et al. 2012). Using interspecific and intraspecific contrasts, it is claimed that a significant portion of the amino acid replacements among Drosophila species are mutations involved in adaptive change (Smith and Eyre-Walker 2002; Shapiro et al. 2007), but little progress has been made in defining those context-dependent properties that determine the likelihood that a given gene participates in adaptive response. Although much of the focus has been on amino acid mutations, where functional effect is ambiguous in most cases, only a handful of studies have focused on geographic variation in expression polymorphism (Whitehead and Crawford 2006; Fraser et al. 2010; Fraser 2013).

Drosophila melanogaster is one of the best models with which to explore natural selection in a geographic and ecological context. In temperate regions, the population is envisioned as a seasonal metapopulation, where populations die back during the winter followed by local reestablishment of populations in the spring that are seeded through survivors in basements, barns, and compost piles, and thus contributing to a local genetic continuity through time (Ives 1945, 1954; Reaume and Sokolowski 2006; Shpak et al. 2010; Garrigan et al. 2010). This local survival of populations has led to the expectation that there are adaptations that are associated with the colonization of cosmopolitan populations that span a wide range of environments from subtropical to temperate and that often are associated with seasonal variation in seasonal nutrient availability. Many adaptations must involve changes in metabolic architecture, as energy tradeoffs are known to shift along the climatic gradient. The presence of latitudinal clines in the frequencies of alleles in many metabolic genes (Sezgin et al. 2004) suggests this spatial-seasonal model is plausible for studying genetic response in a defined context, namely the pathways of central metabolism. There has been considerable progress in describing variation at the full genome level in D. melanogaster (Clark et al. 2007; Kolaczkowski et al. 2011; Vishnoi et al. 2011; Fabian et al. 2012). These studies have explored genome-wide patterns of differentiation at very low resolution across the entire genome at geographic range extremes and matched them to coarse functional classes but have not examined patterns of variation in finer geographical detail in well-defined functional contexts.

The goal of this study is to evaluate clinal variation in single-nucleotide polymorphisms (SNPs) associated with cis-expression variation in central metabolic genes across D. melanogaster populations spanning a seasonal climate gradient in the eastern US and use this knowledge to predict geographic changes in central metabolic pathway architecture. It is expected that this can be used to identify gene targets and mechanisms of adaptation. The pathway represents the central flow and partitioning of energy as nutrient levels change (Gershman et al. 2007), and it is likely that this partitioning varies geographically with those life history challenges associated with somatic maintenance and reproduction. Moreover, aside from simple energy partitioning, it is well known that particular metabolites of central metabolism play key signaling roles in energy-state sensing. It follows that variation in expression of associated enzymes in close pathway proximity to these metabolites may respond to changing selection pressures to reset signal levels that couple nutritional state to overall downstream metabolic and stress responses (Moore et al. 2003; Kim and Dang 2005; Kim et al. 2008; Rathmell and Newgard 2009; Wellen et al. 2009).

Results

General Summary of Clines and Expression Quantitative Trait Nucleotides

We used bulk pyrosequencing to examine allele frequencies for 127 SNPs in 46 genes of central metabolism across 20 local populations spanning a latitudinal gradient from southern Florida to Ontario. SNP selection was not intended to be exhaustive but to use the bulk pyrosequencing approach to screen several SNPs per gene. The number of SNPs screened per gene depended partly on the resident variation, which varies considerably among loci. For example, six genes possess a single screened SNP, and several have as many as seven screened SNPs; the average number per gene is 2.6 SNPs. Each SNP was tested for clinal variation by a linear regression of allele frequency against latitude. We find that 53 of 127

SNPs are significantly clinal in 30 of 46 genes at $a < 0.05$. We might have expected six to seven significant SNPs at this type I error rate, so the prevalence of clines in these genes is clear. Our tests of SNPs within genes are not statistically independent, and the possibility of reporting a gene with a significant cline simply because of type 1 error obviously increases with the number of SNPs sampled per gene. We are unable to test for a gene-wise error by controlling for the linkage correlation because the bulk pyrosequencing produces only a mean frequency estimate for each population. Overall, using a value cutoff of $q < 0.04$ (Storey and Tibshirani 2003) that takes into account our multiple tests of clinal significance (but ignores within gene linkage disequilibrium), we still predict a significant cline in one or more SNPs in 30 of 46 genes (supplementary table S1, Supplementary Material online).

Among genes, there is evidence for cis-expression effects for many of these SNPs. Our SNPs were selected for pyrosequencing focusing on amino acid changes and those SNPs acting best as proxies for haplotype structure. They were not preselected for cis-expression association. Many potential SNPs cannot be screened because they possess closely linked polymorphisms that overlap pyrosequencing primers or are in close proximity and interfere with the pyrosequencing estimate. Finally, a small proportion fails for unknown reasons.

For each screened SNP, we tested association with expression variation using a nested analysis of variance (ANOVA) of the Affymetrix Drosophila 2.0 array expression data reported in Ayroles et al. (2009) and the original 37 sequences of the Drosophila melanogaster Genetic Reference Panel (DGRP) first released in 2009 (Mackay et al. 2012). We observe cases of significant expression quantitative trait nucleotides (eQTNs) in 21 genes if we ignore nonindependence and multiple tests. Using multiple test criteria within each gene to control for a gene-wise error rate (Stranger et al. 2007), we find that 12 of 46 genes possess at least one significant eQTN in the coding region of one or both sexes using the FDR of 10%. Ten of these 12 genes possess a significant cline in the most significant eQTN (supplementary table S1, Supplementary Material online). We should note that the relative allelic expression variation for several genes is not subtle. For example, in the DGRP, the relative differences in expression associated with the SNP alleles in Raleigh, NC, at Gpdh and Got2 are 47% and 29%, whereas those associated with the alleles at Idh-f3, Hex-A, and Pgd are 18%. For each gene, the SNP with the largest cis-acting effect was used with the SNP allele frequency to predict, assuming additivity, the mean gene expression expected for each gene and each collection locality (see Materials and Methods). This matrix of population-by-gene mean expression was used in the principal component analysis (PCA).

PCA

The PCA of the full geographic-expression data of 45 genes (Adh was removed) allows us to identify the major patterns of across population gene expression variation and covariation along the eastern US climatic gradient. We have carried out the PCA on the variance–covariance matrix of standardized eQTN expression. High gene loadings on the first principal component will reflect those genes with both strong allele frequency variation and the large average allelic effects on expression. The percent variances associated with the first three PCs were 50%, 12%, and 8%, respectively. Table 1 shows the individual gene expression loadings associated with the first three principal components, as ranked by by sign respectively, decreasing and increasing mean expression with increasing latitude. The population scores for the first PC axis shows a strong correlation with latitude (fig. 1; $r = 0.833$, $P < 0.0001$). This is because the most of the variance–covariance among genes in mean population expression is associated with the latitudinal

distribution of population allele frequencies. The six genes with the highest positive loadings and significant eQTNs are Men-b, Gpdh, Got-2, Gdh, Pdk, and Acon. The five genes with the most negative loadings and significant eQTNs are Idh-f3, Pgd, Tpi, UGPase, and Hex-A.

To show overall patterns of increased and decreased expression, these gene loadings are placed into a central– peripheral metabolic pathway context in figure 2. The support associated with each step is scaled by the format of the arrow and the direction of expression change with latitude is indicated by color (see fig. 2).

Discussion

The goal of this study is to identify steps or nodes in the central and peripheral metabolic pathway of glycolysis where SNP-associated expression variation is an apparent target of natural selection in response to environmental correlates of latitude. In the apparent absence of isolating structures, geographic clines have frequently been interpreted as evidence of selection and adaptation to climate and associated variables. Our question is whether adaptation associated with the clines involves the modulation of gene expression in this pathway, and whether this modulation is focused on particular nodes, networks, or genes. In this study, we use a two-step approach where both latitudinal patterns and observed SNP effects on expression are integrated and then evaluated in the context of the central metabolic pathway to identify those steps that are most likely to be involved in climatic adaptation and suggest a mechanism.

Several steps within this well-described pathway predict interesting shifts in metabolic architecture (highlighted in fig. 2) that we propose could reflect selection on either of two related phenotypes. The first phenotype is a shift in flux balance or the partitioning of flux. The genes most strongly associated with clinal climatic change are those linked with metabolites with the highest connectivity in the overall metabolic network. Wagner and Fell (2001) in assessing the network architecture of central metabolism in Escherichia coli (similar in metazoans) placed glutamate and pyruvate first and second in their connectivity rank, followed by CoA, 2-oxglutarate, glutamine, and aspartate (see fig. 2). Gdh, Got2, and Idh-f3 directly affect glutamate and 2-oxglutarate, and Got2 also affects aspartate well. Pdk, PC, and Men-b all impact the metabolism of pyruvate. Wagner and Fell (2001) did not address the energy cofactors, NAD/NADH, NADP/ NADPH, and ATP/ADP because those possess even greater connectivity. However, Hex-A, Pgd, Gdh, Gpdh, and Gpo-1 use these cofactors that reflect energy state. This observation raises the speculation that expression selection for genes at the nodes with the highest connectedness is central to changing metabolic flux balance, and it is expression variation in these genes that is responding to selection along the cline to shift this balance.

A second phenotype could be the cellular levels of the highly connected metabolites themselves and their effect on energy-state sensing in response to changing nutrient levels. The nutrient response networks are early features of evolution and appear widely conserved. In these networks, metabolite levels trigger downstream transcriptional changes to shifting nutrient input. The detection of nutrient state and its response generally acts through the levels of metabolites that are most reactive to nutrient input. For example, in plants where light, temperature, and CO2 affect photosynthesis, it is the sugars that signal energy levels and initiatedownstream gene transcription that determines plastic growth responses (e.g., Heisel et al. 2013; Xiong et al. 2013). Many studies in mammals have shown that perturbation of the enzymes of the central metabolic pathway sets energy state and determine insulin secretion especially through mitochondrial function (e.g., Guay et al. 2007). In D. melanogaster, there are comparable pancreatic models of energy or nutrient sensing that involve neurosecretory cells (Toivonen and Partridge 2008; Nassel and Winther 2010).

This introduces the hypothesis that natural genetic variation in key metabolic enzymes may play a role not only in shifting flux balance but also in setting homoeostatic limits, acting as "energy-stats" that determine the nutritional set points that trigger downstream responses of the wellestablished sensing pathways in Drosophila (Baker and Thummel 2007; Savraj 2009). If the sensing mechanisms are universally conserved around the same common metabolites, then we should have in Drosophila the same interest in the enzymes associated with cofactor shuttles (see Eto et al. [1999], as well as certain branches that share metabolites with high connectivity [e.g., glutamate and pyruvate]). As discussed in more individual detail below, these shifts should involve the hexokinases, those enzymes of the NAD/ NADH cofactor or redox shuttles, and those enzymes at the glutamate node. Moreover, enzymes with function in the mitochondria also hold special interest because of well-established signaling responses to nutrient levels associated with mitochondrial function (Wiederkehr and Wollheim 2006; Baltzer et al. 2010; Cho et al. 2011). Therefore, we might propose that expression variation observed in genes encoding these enzymes may respond to the climatic variation along a latitudinal tropical-temperate gradient that varies seasonal and locally in the availability of nutrients. The unparalleled biochemical and physiological knowledge of many steps in the central metabolic pathway allows a functional interpretation of the consequences of expression variation.

Overall, in the upper glycolytic pathway and its branches, there appears to be coordinated decreases in expression with latitude. Both of the principal hexokinase genes, Hex-A and Hex-C (Duvernell and Eanes 2000), show lower SNP-associated expression with increasing latitude, as do the enzymes of the pentose shunt (G6pd, Pgd, TA), Tpi, fbp, Gapdh1, and UGPase. These reductions along with reduced Pepck suggest an overall decrease in gluco- and glyconeogenesis with latitude with an increase in fatty acid synthesis. The hexokinases have singly emerged as having jack-of-all-trade function in both animals and plants (Cho and

Yoo 2011) and especially in the roles of nutrient signaling associated with glucose and glucose- 6-phosphate (Moore et al. 2003; Cho and Yoo 2011). The role of glucokinase mutations in the misreading of glucose sensing in humans is very well established; many regulatory mutations in glucokinase reset the blood glucose levels where insulin in secreted, creating hypo- and hyperglycemia (Matschinsky 2005; Matschinsky et al. 2006).

Here, the two major genes (Gdh and Got2) for enzymes coupled to glutamate metabolism in the mitochondria both show increases in population-level expression with latitude. Glutamate is important because it stands at the intersection of carbohydrate and amino acid metabolism and will clearly reflect nutrition status (Brosnan 2000). Both enzymes control the entry of amino acids into carbohydrate and energy metabolism. It is well established that glutamate is an important signaling molecule for energy state (Karaca et al. 2011). Moreover, regulatory mutations in human GDH are also associated with hyperinsulinism; they reset energy-state signaling associated with amino acid levels (Stanley 2004). Given GDH's regulatory sensitivity to redox state, the parallel and elevated level of expression with the glycerol shuttle genes (Gpdh and Gpo1) is particularly intriguing.

The third general observation is the increasing expression of Gpdh and Gpo-1 with latitude. Both enzymes are associated with the essential glycerol shuttle that transfers NAD/NADH equivalents into the mitochondria for subsequent use in oxidative phosphorylation. This is the major cofactor shuttle in insects (O'Brien and MacIntyre 1972; MacIntyre and Davis 1987; Carmon and MacIntyre 2010) and should control the redox balance in the mitochondria. The essential role of the NAD/NADH shuttles in nutrient sensing in mammals has been noted repeatedly (Eto et al. 1999), and it is well established that starvation in Drosophila significantly changes the redox ratio upward (Zhu and Rand 2012). Furthermore, by setting NAD/NADH ratio, this shuttle should couple metabolic status with transcriptional control through the sirtuins, the NAD-dependent histone deacetylases (Imai et al. 2000) shown to affect chromatin silencing and impact life span in number of organisms (Imai 2011). Both members, Gpdh and Gpo-1, show enhanced activity, but this does not predict direction, just an increase in shuttle function in the north.

Although a number of robust and novel connections emerge from our analysis, we note several caveats. The whole body measures of adult expression make it difficult to predict functional responses for genes whose enzyme products possesses roles in many tissues and often are functionally and physiologically reversible, possessing both energy producing and energy consuming roles (e.g., PGI). Other genes are more specialized (e.g., PEPCK in the fat body) and are effectively irreversible in those tissues, so the predicted response is less equivocal. Nevertheless, as first entries into these questions, we believe that we have recovered informative signal from the expression that is averaged across many tissues.

In some genes, there are undoubtedly SNPs with stronger expression effects that we have not screened. Although relatively comprehensive, the SNPs we evaluated with bulk pyrosequencing are not random. We emphasized coding regions, SNPs that are amino acid replacement polymorphisms, and those that are associated with haplotype structure. They were not initially targeted for expression effects. It is possible that although coding regions can have significant effects on gene expression (Kudla et al. 2009), we may have found better candidates for cis-expression SNPs in 50-noncoding regions (Massouras et al. 2012). Furthermore, the availability of suitable polymorphisms varies gene-by-gene, and we cannot screen every SNP with possible expression effects. In some cases, we have not screened the SNP with the greatest expression effect in a gene but a linked SNP that is in disequilibrium. In this case, our SNP is a weaker proxy for a SNP of greater effect. Therefore, we may underestimate the latitudinal change in cis-expression at some genes. Finally, we assume that the allelic effects estimated in the Raleigh DGRP do not significantly change across the cline. This is impossible to assess without independent studies as Ayroles et al. (2009) in other populations.

Many of the SNPs are amino acid polymorphisms or are in linkage disequilibrium with amino acid polymorphisms. Some amino acid changes are associated with transcript expression, but others have no expression effects. These may have functional catalytic influence and are clinal because of these catalytic differences. For example, there are no cis-expression variable sites in the Pdp gene, but there are three prominent amino acid polymorphisms (all in linkage disequilibrium) that are strongly clinal. These residues in PDP might play a role in the regulation and activation of PDH. Without a detailed functional study, this cannot be determined and in many cases the required functional characterizations would be challenging.

It is important to recognize that the absence of either cis-expression variation or clines in many genes is expected. This is because the control of flux in pathways and networks is likely to be distributed unevenly and concentrated at different steps as consequence of pathway architecture or the unique regulation of the enzymatic steps and pathway (Fell 1997; Olson-Manning et al. 2013). If extant expression variation reflects functional responses under selection, then the absence of cis-effect variation could simply reflect the lack of potential for step control at that point. Thus, the absence of effects, if real and not an artifact, is just as informative to understand the metabolic architecture as the presence. For example, Pyk and Idh, and the lower elements of the glycolytic pathway are possibly ineffective as targets of selection because they possess low control. Alternatively, some genes may possess pleiotropic constraints or tissue-specific tradeoffs or are poor mutagenic targets for cis-based expression variation. Without the independent, and albeit difficult, experimental assessment of metabolic control at each step (Eanes 2011), it is impossible to address this hypothesis. It is also unclear how increased or reduced activity in the near-equilibrium bidirectional enzymes with wide tissue specificity, such as the central glycolytic core (e.g., Tpi, Pgi, and

Eno), affects bias in glycolytic, gluconeogenic, or glycogenic flux. It is proposed theoretically (Wright and Rausher 2010) and shown empirically (Olson- Manning et al. 2013) that flux control localizes at the top of one-way pathways. However, the central pathway is not easily interpreted in that top–bottom context.

The goal of our report is to integrate the population genetics of geographic variation of SNPs with gene expression effects and then interpret this integration in the functional context of the central metabolic pathway. From this integration, interesting associations emerge that suggest hypotheses about selection on the pathway that requires further consideration. We have not proven flux balance nor energy sensing as the phenotypes that has come under selection for these genes: Rather, we have introduced them as hypotheses that emphasize different roles of expression selection on metabolic genes. For example, this is in contrast to the often advanced hypothesis that genetic variation is maintained through temperature-dependent kinetic tradeoffs that maintain constant flux and performance along a thermal cline (Place and Powers 1979; DiMichele and Powers 1982; Hall and Koehn 1983). Our hypothesis, that energy sensing and resource utilization represent the functional basis for selection on metabolic enzymes across environmental gradients, generates predictions that can be tested experimentally. Driven by the need to understand diet-associated ageing and causes of metabolic syndrome, huge progress has been made in unraveling the downstream components of energy sensing in D. melanogaster and other models (Taguchi and White 2008; Toivonen and Partridge 2008; Fontana et al. 2010), but the initial sensing mechanism is unknown. In Drosophila, experimental manipulation of gene expression is tractable using both P-element-associated knockouts of whole-body expression (Merritt et al. 2005; Eanes et al. 2006, 2008), as well as tissue-specific RNAi knockdown (Dietzl et al. 2007; Schnorrer et al. 2010). This should allow the targeted suppression and overexpression of genes of interest in the neurosecretory cells (Giannakou et al. 2004; Lee and Park 2004) and the evaluation of sensing. The effects of these manipulations on energy sensing and associated life history phenotypes can then be evaluated. The study of genetic variation in central metabolism and adaptation is a long running and multidimensional problem generally associated with energy production (Zera 2011). However, the expanding participation of metabolic enzymes in roles outside the theme of simple energy production is increasingly being recognized (Kim and Dang 2005; Marden 2013), and this model advances a new role for expression and catalytic based genetic variation in metabolic genes in adaptation to changing environments.

Materials and Methods

Data Sources—Pathway Genes Identified

To assemble the 46 candidate genes presenting enzymes of the core pathways, we used FLYBase. Some genes possess multiple orthologs; one is the somatic member involved in core metabolism. We have used

several criteria to determine the most relevant member of each gene set. First is high homology to vertebrate members. Second, core enzymes have high somatic expression that is not specific to testis or ovaries. Central metabolic genes have exceptionally high codon biases. Third, mitochondrial proteins have high pIs (Hartmann et al. 1991; Dinur-Mills et al. 2008). In the case of genes with only a candidate gene CG definition, we have introduced abbreviations that are similar or identical with those used in mammalian names.

As our database for SNP identification, we used the 37 D. melanogaster genome sequences released in August 2010 by the DGRP (Mackay et al. 2012). These sequences had already been assembled and annotated to the FlyBase reference sequence (version 5.12). Nevertheless, final choice of useful SNPs was carried out gene-by-gene by further manual inspection. The coding regions of each gene were extracted and the quality score for each base assessed. SNPs where the minority allele quality score was less than 30 were reset to the majority allele. Most of these cases appear as singletons. Finally, focal SNPs for the cline were selected where the minority allele was more than 0.10 in the DGRP.

Selection of SNPs

Many SNPs could not be screened by pyrosequencing because flanking polymorphisms bias amplification in the bulk preparation or the pyrosequencing step failed. Many SNPs were also selected before the Ayroles et al. (2009) expression data were available. Amino acid polymorphisms were favored. We also favored SNPs that were diagnostic of important haplotypes seen in the DGRP collection. We only used sites with a minority allele that was more than 10% in the Raleigh data, and attempts were made to minimize LD among sites by spacing SNPs at distance of more than 500 bp when possible.

Population Collections

In 2009 and 2010, we collected samples from 18 local populations across the eastern United States. Adults were collected by sweep net and immediately separated by sex. Males were preserved in 95% ethanol and stored at -70 oC. Females were allowed to oviposit, preserved, and stored. We also included the 2005 population data from Raleigh, NC (Mackay et al. 2012); we obtained 34 inbred DGRP lines from the Bloomington Stock Center. We included two collections (pooled) from 2007 and 2008 from Sudbury, ON. In each isofemale line, the progeny from the F1 generation was preserved in EtOH for bulk DNA preparation. Two female progenies were collected from each preserved line and pooled with the collected males in the bulk sample for pyrosequencing. By sampling two progenies per line in the F1 generation, we are sampling two to four independent chromosomes from the population per line, with an average of 3. The expected number of chromosomes per bulked sample is therefore three times the number of female

lines plus twice the number of wild-collected males. The average number of independent genomes pooled per population sample was n = 114.7. The entire data base thus consisted of 2,524 genomes. The sources and sample sizes of the 20 population samples are provided in supplementary table S2, Supplementary Material online.

Pyrosequencing

Bulk DNA purification was performed with Puregene Core Kit A (Qiagen) using 42–100 flies per population (sample sizes in supplementary table S2, Supplementary Material online). We used pyrosequencing in the bulk DNA preps to estimate SNP frequency (Lavebratt and Sengul 2006; Doostzadeh et al. 2008). We checked the precision of the method by comparing the estimated frequency of each SNP to the expected frequency for the DGRP population based on the genome sequences (r = 0.99). Pyrosequencing was carried out using the PyroMark MD machine and peak heights scored to estimate SNP. Primers were designed using the PyroMark Assay Design software. A universal biotinylated primer was used in combination with two locus-specific primers (Guo and Milewicz 2003). In a few cases, performance of universal primer was poor, and direct-biotinylated locus-specific primers were used.

Clines

Allele frequency estimates were arc-sine transformed (Sokal and Rohlf 1981) and tested by linear regression against latitude. Individual probabilities for single tests are determined by random permutation of latitudes 10,000 times. We might expect a proportion of these cline discoveries to be false positives. The within-gene SNP tests are not independent, but we cannot capture the within population SNP correlation structure because the pyrosequencing provides mean estimates without individual genotypes. The entire set was tested for set-wide significance using q values (Storey and Tibshirani 2003), and a q value of 4% was assigned as a cutoff as support of a cline in each SNP.

Expression Variation and eQTNs

We downloaded the whole-adult Affymetrix Drosophila 2.0 array expression data (accession number E-MEXP-1594) reported in Ayroles et al. (2009) and the original 37 sequences of the DGRP first released in 2009 (Mackay et al. 2012). Probes with underlying SNPs were removed or masked (Benovoy et al. 2008; Chen et al. 2009). The sex effect for each gene was removed and the residuals rescaled to standardized deviates using the total sample variance. A nested ANOVA is carried out on standardized residuals to estimate SNP allele, nested line in SNP, and nested vial within line effects (Yijkl = μ, + Ai + Bij + Cijk + "ijkl, where Yijkl is an individual expression measure, Ai is the effect of the ith allele, Bij is the effect of line j within the ith SNP allele, and Cijk is the effect of kth vial within line j and the ith SNP allele and "ijkl

is the error term within vials) using the JMP program (JMP-SAS). We used the parameter, ai specifying the genotypic effect for each pair of alleles, to estimate population mean expression (see below).

Some of the within-gene SNP-specific expression effects are not independent. To test for gene-wide significance, we incorporate linkage disequilibrium among sites within each gene and carried out random permutations of expression across lines, keeping the haplotype structure intact. Sexes were treated separately (Massouras et al. 2012). For each 10,000 perturbation set, the highest F value was taken among the SNPs and the 5% tail of the distribution for the 10,000 permutations set as the FDR.

Population Mean Expression and PCA

For each gene, the SNP with the largest cis-acting effect (irrespective of significance) was used. The scaled allelic effect of the ith SNP was used with the SNP allele frequency, qi, to predict, assuming additivity, the mean expression, yi, expected for each collection locality by the simple equation $yi = a(1 - 2qi)$. Thus, the variance in gene expression across populations depends on the cis-acting effect estimate, a, and the variation in allele frequencies, qi. We subjected the variance–covariance matrix for all the genes and populations to a PCA and extracted the major components. The gene loadings on these factors can then be used to summarize the major sources of overall variance–covariance structure of changes in expression for different pathways.

Supplementary Material

Supplementary tables S1 and S2 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjour nals.org/).

Acknowledgments

The authors thank Thomas Merritt for supplying the lines from Sudbury, ON, John True, and Joe Lachance for additional collections from New York, and Frank Jiggins for suggestions on the manuscript. This work was supported by National Institutes of Health grant GM090094 to W.F.E. and John True and Collaborative National Foundation Science grants DEB0921372 to W.F.E. and DEB0542859 and DEB0921307 to P.S.S.

References

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

Alvarez-Ponce D, Guirao-Rico S, Orengo DJ, Segarra C, Rozas J, Aguade´

M. 2012. Molecular population genetics of the insulin/TOR signal transduction pathway: a network-level analysis in Drosophila melanogaster. Mol Biol Evol. 29:123–132.

Andres AM, Hubisz MJ, Indap A, Torgerson DG, Degenhardt JD, Boyko AR, Gutenkunst RN, White TJ, Green ED, Bustamante CD, et al. 2009. Targets of balancing selection in the human genome. Mol Biol Evol. 26:2755–2764.

Ayroles JF, Carbone MA, Stone EA, Jordan KW, Lyman RF, Magwire MM, Rollmann SM, Duncan LH, Lawrence F, Anholt RR, et al. 2009. Systems genetics of complex traits in Drosophila melanogaster. Nat Genet. 41:299–307.

Baker KD, Thummel CS. 2007. Diabetic larvae and obese flies—emerging studies of metabolism in Drosophila. Cell Metab. 6:257–266.

Baltzer C, Tiefenbo¨ck SK, Frei C. 2010. Mitochondria in response to nutrients and nutrient-sensitive pathways. Mitochondrion 10: 589–597.

Benovoy D, Kwan T, Majewski J. 2008. Effect of polymorphisms within probe-target sequences on olignonucleotide microarray experiments. Nucleic Acids Res. 36:4417–4423.

Brosnan JT. 2000. Glutamate, at the interface between amino acid and carbohydrate metabolism. J Nutr. 130:988S–990S.

Carmon A, MacIntyre R. 2010. The alpha-glycerophosphate cycle in Drosophila melanogaster VI. Structure and evolution of enzyme paralogs in the genus Drosophila. J Hered. 101:225–234.

Chen L, Page GP, Mehta T, Feng R, Cui XQ. 2009. Single nucleotide polymorphisms affect both cisand trans-eQTLs. Genomics 93: 501–508.

Cho J, Hur JH, Walker DW. 2011. The role of mitochondria in Drosophila

aging. Exp Gerontol. 46:331–334.

Cho YH, Yoo SD. 2011. Signaling role of fructose mediated by FINS1/FBP in Arabidopsis thaliana. PLoS Genet. 7:e1001263.

Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN, et al. 2007. Evolution of genes and genomes on the Drosophila phylogeny. Nature 450: 203–218.

Clark NL, Alani E, Aquadro CF. 2012. Evolutionary rate covariation reveals shared functionality and coexpression of genes. Genome Res. 22:714–720.

Cork JM, Purugganan MD. 2004. The evolution of molecular genetic pathways and networks. Bioessays 26:479–484.

Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, Fellner M, Gasser B, Kinsey K, Oppel S, Scheiblauer S, et al. 2007. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature 448:151–156.

DiMichele L, Powers DA. 1982. Physiological basis for swimming endurance differences between LDH-B genotypes of Fundulus heteroclitus. Science 216:1014–1016.

Dinur-Mills M, Tal M, Pines O. 2008. Dual targeted mitochondrial proteins are characterized by lower MTS parameters and total net charge. PLoS One 3:e2161.

Doostzadeh J, Shokralla S, Absalan F, Jalili R, Mohandessi S, Langston JW, Davis RW, Ronaghi M, Gharizadeh B. 2008. High throughput automated allele frequency estimation by pyrosequencing. PLoS One 3: e2693.

Duvernell DD, Eanes WF. 2000. Contrasting molecular population genetics of four hexokinases in Drosophila melanogaster, D. simulans and D. yakuba. Genetics 156:1191–1201.

Eanes WF. 2011. Molecular population genetics and selection in the glycolytic pathway. J Exp Biol. 214:165–171.

Eanes WF, Merritt TJ, Flowers JM, Kumagai S, Sezgin E, Zhu CT. 2006. Flux control and excess capacity in the enzymes of glycolysis and their relationship to flight metabolism in Drosophila melanogaster. Proc Natl Acad Sci U S A. 103:19413–19418.

Eanes WF, Merritt TJ, Flowers JM, Kumagai S, Zhu CT. 2008. Direct evidence that genetic variation in glycerol-3-phosphate and malate dehydrogenase genes (Gpdh and Mdh1) impacts adult ethanol tolerance in Drosophila melanogaster. Genetics 181:607–614.

Eto K, Suga S, Wakui M, Tsubamoto Y, Terauchi Y, Taka J, Aizawa S, Noda M, Kimura S, Kasai H, et al. 1999. NADH shuttle system regulates K(ATP) channel-dependent pathway and steps distal to cytosolic Ca(2 +) concentration elevation in glucose-induced insulin secretion. J Biol Chem. 274:25386–25392.

Fabian DK, Kapun M, Nolte V, Kofler R, Schmidt PS, Schlotterer C, Flatt

T. 2012. Genome-wide patterns of latitudinal differentiation among populations of Drosophila melanogaster from North America. Mol Ecol. 21:4748–4769.

Fell DA. 1997. Understanding the control of metabolism. London: Portland Press.

Flowers J, Sezgin E, Kumagai S, Duvernell D, Matzkin L, Schmidt P, Eanes

W. 2007. Adaptive evolution of metabolic pathways in Drosophila. Mol Biol Evol. 24:1347–1354.

Flowers JM, Hanzawa Y, Hall MC, Moore RC, Purugganan MD. 2009. Population genomics of the Arabidopsis thaliana flowering time gene network. Mol Biol Evol. 26:2475–2486.

Fontana L, Partridge L, Longo VD. 2010. Extending healthy life span— from yeast to humans. Science 328:321–326.

Fraser HB. 2013. Gene expression drives local adaptation in humans.

Genome Res. 23:1089–1096.

Fraser HB, Moses AM, Schadt EE. 2010. Evidence for widespread adaptive evolution of gene expression in budding yeast. Proc Natl Acad Sci U S A. 107:2977–2982.

Garrigan D, Lewontin R, Wakeley J. 2010. Measuring the sensitivity of single-locus "neutrality tests" using a direct perturbation approach. Mol Biol Evol. 27:73–89.

Gershman B, Puig O, Hang L, Peitzsch RM, Tatar M, Garofalo RS. 2007. High-resolution dynamics of the transcriptional response to nutrition in Drosophila: a key role for dFOXO. Physiol Genomics. 29: 24–34.

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

Greenberg AJ, Stockwell SR, Clark AG. 2008. Evolutionary constraint and adaptation in the metabolic network of Drosophila. Mol Biol Evol. 25: 2537–2546.

Guay C, Madiraju SRM, Aumais A, Joly E, Prentki M. 2007. A role for ATP-citrate lyase, malic enzyme, and pyruvate/citrate cycling in glucose-induced insulin secretion. J Biol Chem. 282:35657–35665

Guo DC, Milewicz DM. 2003. Methodology for using a universal primer to label amplified DNA segments for molecular analysis. Biotechnol Lett. 25:2079–2083.

Hall JG, Koehn RK. 1983. The evolution of enzyme catalytic efficiency and adaptive inference from steady-state data. Evol Biol. 16:53–69. Hartmann C, Christen P, Jaussi R. 1991. Mitochondrial protein charge.

Nature 352:762–763.

Heisel TJ, Li CY, Grey KM, Gibson SI. 2013. Mutations in histone acetyltransferase1 affect sugar response and gene expression in Arabidopsis. Front Plant Sci. 4:1–13.

Imai S. 2011. Dissecting systemic control of metabolism and aging in the NAD World: the importance of SIRT1 and NAMPT-mediated NAD biosynthesis. FEBS Lett. 585:1657–1662.

Imai S, Armstrong CM, Kaeberlein M, Guarente L. 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. Nature 403:795–800.

Ives PT. 1945. Genetic structure of American populations of Drosophila melanogaster. Genetics 30:167–196.

Ives PT. 1954. Genetic changes in American populations of Drosophila melanogaster. Proc Natl Acad Sci U S A. 40:87–92.

Jumbo-Lucioni P, Ayroles JF, Chambers MM, Jordan KW, Leips J, Mackay TF, De LM. 2010. Systems genetics analysis of body weight and energy metabolism traits in Drosophila melanogaster. BMC Genomics 11:297.

Karaca M, Frigerio F, Maechler P. 2011. From pancreatic islets to central nervous system, the importance of glutamate dehydrogenase for the control of energy homeostasis. Neurochem Int. 58:510–517.

Kim J, Bang H, Ko S, Jung I, Hong H, Kim-Ha J. 2008. Drosophila ia2 modulates secretion of insulin-like peptide. Comp Biochem Physiol A Mol Integr Physiol. 151:180–184.

Kim JW, Dang CV. 2005. Multifaceted roles of glycolytic enzymes. Trends Biochem Sci. 30:142–150.

Kolaczkowski B, Kern AD, Holloway AK, Begun DJ. 2011. Genomic differentiation between temperate and tropical Australian populations of Drosophila melanogaster. Genetics 187:245–260.

Kudla G, Murray AW, Tollervey D, Plotkin JB. 2009. Coding-sequence determinants of gene expression in Escherichia coli. Science 324: 255–258.

Lavebratt C, Sengul S. 2006. Single nucleotide polymorphism (SNP) allele frequency estimation in DNA pools using pyrosequencing. Nat Protoc. 1:2573–2582.

Lee G, Park JH. 2004. Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in Drosophila melanogaster. Genetics 167:311–323.

Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, Davey RP, Roberts IN, Burt A, Koufopanou V, et al. 2009. Population genomics of domestic and wild yeasts. Nature 458:337–341.

Lu Y, Rausher MD. 2003. Evolutionary rate variation in anthocyanin pathway genes. Mol Biol Evol. 20:1844–1853.

MacIntyre RJ, Davis MB. 1987. A genetic and molecular analysis of the alpha glycerophosphate cycle in Drosophila melanogaster. Isozymes Curr Top Biol Med Res. 14:195–227.

Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM, Cridland JM, et al. 2012. The Drosophila melanogaster genetic reference panel. Nature 482:173–178.

Marden JH. 2013. Nature's inordinate fondness for metabolic enzymes: why metabolic enzyme loci are so frequently targets of selection. Mol Ecol. 22:5743–5764.

Massouras A, Waszak SM, Albarca-Aguilera M, Hens K, Holcombe W, Ayroles JF, Dermitzakis ET, Stone EA, Jensen JD, Mackay TF, et al. 2012. Genomic variation and its impact on gene expression in Drosophila melanogaster. PLoS Genet. 8:e1003055.

Matschinsky FM. 2005. Glucokinase, glucose homeostasis, and diabetes mellitus. Curr Diab Rep. 5:171–176.

Matschinsky FM, Magnuson MA, Zelent D, Jetton TL, Doliba N, Han Y, Taub R, Grimsby J. 2006. The network of glucokinase-expressing cells in glucose homeostasis and the potential of glucokinase activators for diabetes therapy. Diabetes 55:1–12.

Merritt TJ, Duvernell D, Eanes WF. 2005. Natural and synthetic alleles provide complementary insights into the nature of selection acting on the Men polymorphism of Drosophila melanogaster. Genetics 171:1707–1718.

Moore B, Zhou L, Rolland F, Hall Q, Cheng WH, Liu YX, Hwang I, Jones T, Sheen J. 2003. Role of the Arabidopsis glucose sensor HXK1 in nutrient, light, and hormonal signaling. Science 300:332–336.

Nassel DR, Winther AME. 2010. Drosophila neuropeptides in regulation of physiology and behavior. Prog Neurobiol. 92:42–104.

O'Brien SJ, MacIntyre RJ. 1972. The a-glycerophosphate cycle in

Drosophila melanogaster. I. Biochemical and developmental aspects.

Biochem Genet. 7:141–161.

Obbard DJ, Welch JJ, Kim KW, Jiggins FM. 2009. Quantifying adaptive evolution in the Drosophila immune system. PLoS Genet. 5: e1000698.

Olson-Manning CF, Lee CR, Rausher MD, Mitchell-Olds T. 2013. Evolution of flux control in the glucosinolate pathway in Arabidopsis thaliana. Mol Biol Evol. 30:14–23.

Place AR, Powers DA. 1979. Genetic variation and relative catalytic efficencies: lactate dehtdrogenase-B allozymes of Fundulus heteroclitus. Proc Natl Acad Sci U S A. 76:2354–2358.

Ramsay H, Rieseberg LH, Ritland K. 2009. The correlation of evolutionary rate with pathway position in plant terpenoid biosynthesis. Mol Biol Evol. 26:1045–1053.

Rathmell JC, Newgard CB. 2009. A glucose-to-gene link. Science 324: 1021–1022.

Reaume CJ, Sokolowski MB. 2006. The nature of Drosophila melanogaster. Curr Biol. 16:R623–R628.

Savraj SG. 2009. Insulin/TOR signaling in growth and homeostasis: a view from the fly world. Int J Biochem Cell Biol. 41:1006–1010.

Schnorrer F, Schonbauer C, Langer CCH, Dietzl G, Novatchkova M, Schernhuber K, Fellner M, Azaryan A, Radolf M, Stark A, et al. 2010. Systematic genetic analysis of muscle morphogenesis and function in Drosophila. Nature 464:287–291.

Sezgin E, Duvernell DD, Matzkin LM, Duan Y, Zhu CT, 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.

Shapiro JA, Huang W, Zhang C, Hubisz MJ, Lu J, Turissini DA, Fang S, Wang HY, Hudson RR, Nielsen R, et al. 2007. Adaptive genic evolution in the Drosophila genomes. Proc Natl Acad Sci U S A. 104:2271–2276.

Shpak M, Wakeley J, Garrigan D, Lewontin RC. 2010. A structured coalescent process for seasonally fluctuating populations. Evolution 64: 1395–1409.

Smith NG, Eyre-Walker A. 2002. Adaptive protein evolution in

Drosophila. Nature 415:1022–1024.

Sokal RR, Rohlf FJ. 1981. Biometry. New York: Freeman.

Stanley CA. 2004. Hyperinsulinism/hyperammonemia syndrome: insights into the regulatory role of glutamate dehydrogenase in ammonia metabolism. Mol Gene Metabol. 81:45–51.

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

Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, et al. 2007. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315:848–853.

Taguchi A, White MF. 2008. Insulin-like signaling, nutrient homeostasis, and life span. Annu Rev Physiol. 70:191–212.

Toivonen JM, Partridge L. 2008. Endocrine regulation of aging and reproduction in Drosophila. Mol Cell Endocrinol. 299:39–50.

Vishnoi A, Sethupathy P, Simola D, Plotkin JB, Hannenhalli S. 2011. Genome-wide survey of natural selection on functional, structural, and network properties of polymorphic sites in Saccharomyces paradoxus. Mol Biol Evol. 28:2615–2627.

Wagner A, Fell DA. 2001. The small world inside large metabolic networks. Proc R Soc Lond B Biol Sci. 268:1803–1810.

Wellen KE, Hatzivassiliou G, Sachdeva UM, Bui TV, Cross JR, Thompson CB. 2009. ATP-citrate lyase lin Whitehead A, Crawford DL. 2006. Neutral and adaptive variation in gene expression. Proc Natl Acad Sci U S A. 103:5425–5430.

Wiederkehr A, Wollheim CB. 2006. Minireview: implication of mitochondria in insulin secretion and action. Endocrinology 147: 2643–2649.

Wright KM, Rausher MD. 2010. The evolution of control and distribution of adaptive mutations in a metabolic pathway. Genetics 184: 483–502.ks cellular metabolism to histone acetylation. Science 324:1076–1080.

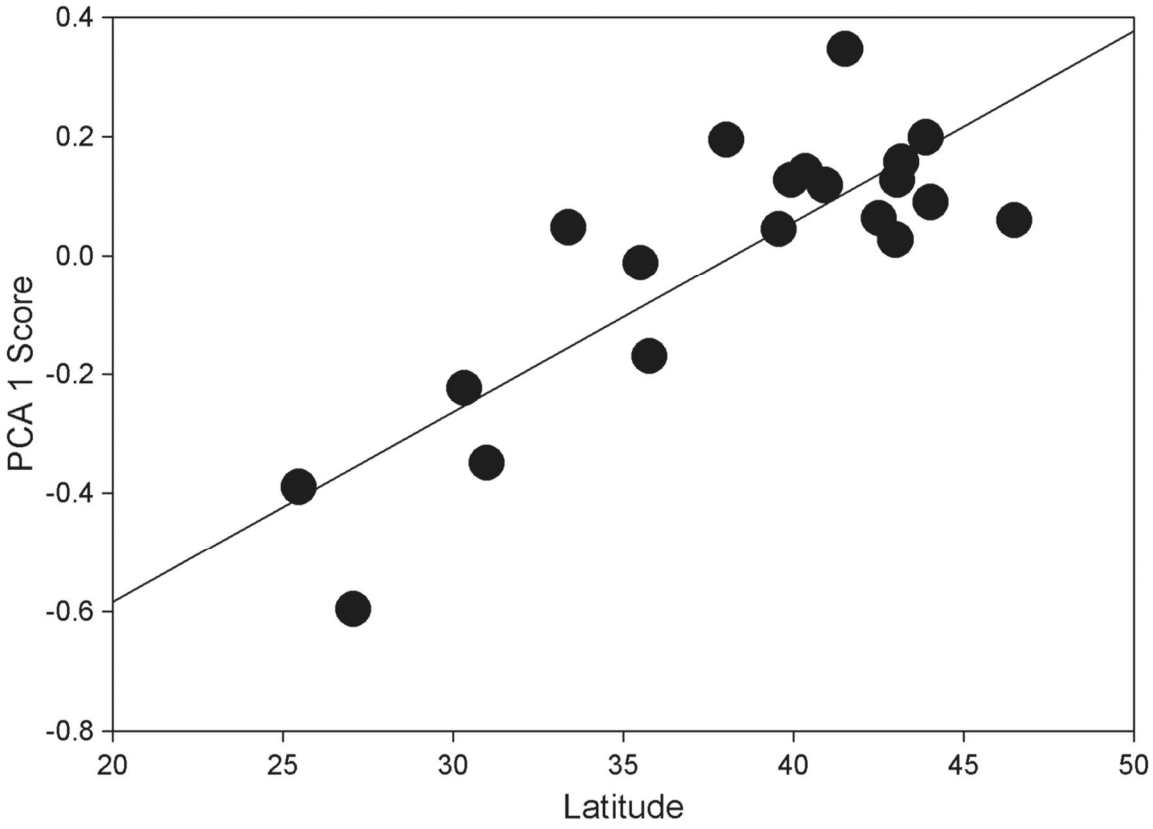


Figure 1. Relationship of population projections on the first axis and latitude.

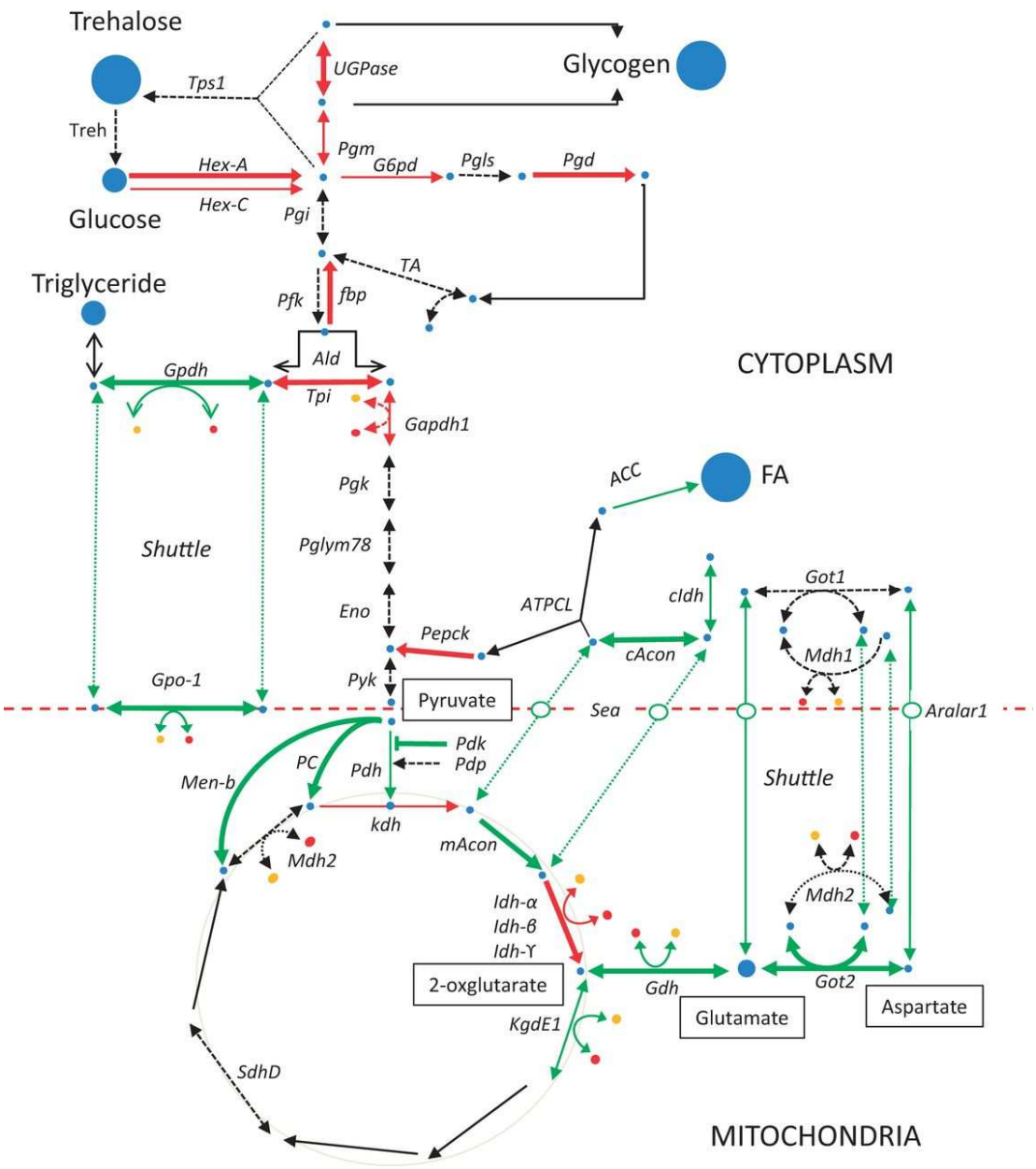


Figure 2. The central metabolic pathway and its immediate branches. The color reflects the direction of expression change with increasing latitude (green increasing, red decreasing). Thick arrows are genes with both significant clines and significant expression effects. They possess high PCA1 loadings. Thin colored lines have significant expression effects but no significant clines. Dotted lines have nonsignificant expression effects but may or may not have significant clines. The latter two groups are simply suggestive of change and direction. Black solid lines (ATPCL) are not available. The red and yellow dots represent NADH and NAD cofactors, respectively, and the large red dotted line is the cytoplasmic-mitochondrial boundary. The metabolities with highest connectivity are boxed. The results for *Men*, *Gpt*, and *Pfrx* are not included to prevent undue complexity.

Gene	I	II	III
Idh-β	−0.628[a]	0.083[a]	0.131[a]
Pgd	−0.244[a]	0.364[a]	−0.176[a]
Tpi	−0.113[a]	−0.101[a]	−0.044[a]
UGPase	−0.106[a]	−0.022[a]	−0.083[a]
Hex-A	−0.102[a]	0.102[a]	0.051[a]
Idh-ϒ	−0.075[b]	0.113[b]	−0.067[b]
G6pd	−0.062[b]	0.092[b]	0.040[b]
kdn	−0.061[b]	0.053[b]	−0.029[b]
Pepck	−0.041[a]	0.065[a]	−0.003[a]
Pglym78	−0.038	−0.085	−0.028
Gapdh1	−0.035[b]	0.016[b]	−0.041[b]
Pyk	−0.032[c]	0.028[c]	−0.015[c]
Pgi	−0.030	0.028	−0.035
Pgm	−0.025[b]	0.018[b]	0.023[b]
fbp	−0.025[a]	−0.008[a]	0.028[a]
SdhD	−0.013	−0.032	0.007
Pfrx	−0.013	−0.003	−0.039
Gpt	−0.013	0.000	0.020
Idh-α	−0.009	−0.003	0.009
Hex-C	−0.009[b]	−0.103[b]	−0.008[b]
TA	−0.006[c]	−0.016[c]	0.006[c]
Got-1	−0.002	−0.001	−0.018
Idh	0.003[b]	0.050[b]	−0.164[b]
Pgk	0.004[c]	−0.001[c]	0.004[c]
Men	0.008[a]	−0.063[a]	0.116[a]
Pgls	0.009	−0.008	0.001
Treh	0.009	0.001	0.002
Gpo-1	0.020[a]	0.014[a]	−0.034[a]
Pfk	0.024	0.006	0.043
Eno	0.032	−0.040	0.002
Pdp	0.037[c]	0.008[c]	0.020[c]
Tps1	0.039[c]	0.072[c]	0.006[c]
Aralar	0.043[b]	−0.035[b]	0.039[b]
KgdE1	0.046[b]	−0.101[b]	0.082[b]
Mdh2	0.066[c]	0.034[c]	0.058[c]
ACC	0.076[b]	0.148[b]	0.025[b]
Irp-1B	0.087[a]	−0.049[a]	0.073[a]
Pdh	0.087[b]	−0.225[b]	0.031[b]
Acon	0.096[a]	−0.003[a]	0.125[a]
PC	0.131[a]	0.050[a]	−0.034[a]
Pdk	0.175[a]	0.040[a]	−0.095[a]
Gdh	0.219[a]	−0.128[a]	0.644[a]
Got-2	0.302[a]	0.413[a]	−0.427[a]
Gpdh	0.337[a]	0.599[a]	0.365[a]
Men-b	0.382[a]	−0.372[a]	−0.328[a]

Table 1. Gene loadings for the first three factors. [a]Both significant cline and allelic effect. [b]Significant allelic effect only. [c]Significant cline in gene only.

Chapter 2- Population Genetics of metabolic enzymes in North American *Drosophila melanogaster*

Background

Experiments designed to find candidate genes for more detailed study are typically strictly forward genetics screen on a particular phenotype. Candidate genes are those that have a major effect on the phenotype among all genes tested, often most of the genome. How the candidates arise are a function of the experimental design, from the type and amount of genetic variation to the phenotypic measures. Knock-out lines of deletions covering the genome in an otherwise isogenic background work well to determine the structure and components of a network but ignore population variation. Quantitative trait mapping using inbred lines uses population variation but for two, and only two, lines.

Tests of genetic patterns of adaptation have been developed for use prior to the relatively new era of affordable high-throughput sequencing has been applied to new genomic data sets. Fay et al (2002) used the amino acid replacement to synonymous substitution ratio (A/S) and assumptions of the McDonald Kreitman tests on 45 genes in *Drosophila melanogaster* to determine whether positive selection is acting on particular loci as demographic events would have effected all loci similarly. Shapiro et al (SHAPIRO *et al.* 2007) took this further with 419 genes and Mackay et al (MACKAY *et al.* 2012) performed McDonald-Kreitman (MK) tests (MCDONALD and KREITMAN 1991) across the entire genome of *Drosophila melanogaster* for both coding and modified MK tests for non-coding regions (JENKINS *et al.* 1995; ANDOLFATTO 2005; EGEA *et al.* 2008). Studies using Fst outlier tests are quite common, as are software resource (BEAUMONT and NICHOLS 1996; ANTAO *et al.* 2008; FOLL and GAGGIOTTI 2008). Hohenlohe et al (2010)used both Fst scans between three populations of fresh water and marine three-spine sticklebacks. They found that several regions of significantly increased or decreased Fst overlapped when comparing the three freshwater populations and that some of these significant outlier regions overlapped with previously described QTL linkage groups.

Here, I consider common population genetics tests in the context of candidate gene selection and test whether any of these tests find genes similar to results from Lavington et al (2014), or each other. The tests used require only genetic information to detect genes that deviate from neutrality.

A major assumption is that gene networks, or at least a set of genes, are known and defined such that appropriate attention is paid to this defined set of genes rather than the whole genome. This risks leaving out genes that have not been identified as part of the network, however the main question remains for the network as it is defined: for a phenotype of interest, and a defined underlying genetic network for this phenotype, which gene(s) exhibit(s) patterns of selection?

The phenotype of interest here is metabolic flux and the genetic network is 61 genes in glycolysis, the TCA cycle, the glycerophosphate shuttle, malate-aspartate shuttle and transitions to other metabolic pathways. I performed McDonald-Kreitman tests (MCDONALD and KREITMAN 1991) on each gene and a Hudson-Kreitman-Aguade test (HUDSON *et al.* 1987) on all genes together using protein coding sequences of 61 metabolic network genes extracted from the Drosophila Population Genomics Project (Release 1.0) sequences (MACKAY *et al.* 2012). I also performed an Fst outlier test on SNPs described in Lavington et al (2014).

Methods and Materials

Data- Protein coding sequences were extracted from 37 Raleigh, North Carolina genome sequences of the Drosophila Population Genomics Project release 1.0 (DGRP). Protein coding sequences as determined by FlyBase release 5.22. Corresponding *Drosophila simulans* were either from FlyBase release 1.3, or a consensus file generated from alignments of whole genome sequences described in (BEGUN et al. 2007). Quality filtering of DGRP sequences and collection of SNP data of North American populations used was described in Lavington et al (2014). Only SNPs with reliable reads for all 20 populations were used for a total 131 SNPs.

Population polymorphism, divergence, and McDonald-Kreitman tests- Population polymorphism and divergence measures and MKA tests (MCDONALD and KREITMAN 1991) were performed in DNAsP 5 (LIBRADO and ROZAS 2009).

Hudson-Kreitman-Aguade test –The HKA test (HUDSON *et al.* 1987) was performed using the software package HKA (https://bio.cst.temple.edu/~hey/software/software.htm#HKA) with 10,000 iterations.

Fst outliers – Fst outlier test (BEAUMONT and BALDING 2004) was performed on SNP data from 20 populations described in software package BayeScan2.1 (FOLL and GAGGIOTTI 2008)

(http://cmpg.unibe.ch/software/BayeScan/) based on sizes of 5000, thinning interval of 10, 20 pilot runs of 5000 each and a burn in of 50000. Posterior Odds (PO) are used instead of Bayes Factor (BF) and Jeffrey's scale of evidence can still be used. BayeScan limits the value of Log10(PO) to a maximum of 1000 when tending to infinity.

Correlation between tests – Even when using the data same source, each tests uses different information and are expected to be independent. To test whether results from each test were independent, correlation of candidate gene calls was tested by filling a 2x2 contingency table between two tests with the cells filled by counts of significant/non-significant calls for each gene for each test. For example, *Pgls* SNPs were not Fst outliers, but the gene did deviate from neutrality in the McDonald-Kreitman test. Thus *Pgls* would add one to the significant MK/non-significant Fst outlier cell. Most tests were expected and observed to be non-significant. An over representation of genes significant for both tests, relative to genes independently significant for either of the tests, would suggest non-independence of the results of the two tests. Total counts varied between comparisons as not all tests were performed on all genes. Statistical analyses were performed in DNAsP 5.

Results and Discussion

Results of the McDonald-Kreitman tests presented here (Table 1) expands not only the genes tested, but also the pathway as compared to Flowers et al. (2008). The conclusion remains the same with the main pattern of adaptation within this set of genes clustered around the branch point of Glucose-6-phosphate.

The whole model, that is the sum of deviations of all genes in the test, Hudson-Kreitman-Aguade test (Tables 2 & 3) violated neutrality (χ^2=98.341, d.f.=60, p=0.0013) with the main contributions to the overall deviation from *sea*, *Idh* (cytosolic), and one of the mitochondrial aconitase genes (*CG9244*). *Gpdh*, *Pepck*, *Idh-β*, and *Irp-1B* also contributed to a lesser extent. This is a relatively small list, but certainly suggests an importance of citrate cycling between the mitochondrion and cytosol, with both mitochondrial and cytosolic aconitases, and IDH genes, as well as the mitochondrial citrate transporter *sea*.

Fst outliers are distributed throughout the pathway without a clear pattern (Table 4 & Fig 1). It is quite possible that many of the outliers are false positives for a variety of reasons (HOLSINGER and

WEIR 2009; BIERNE *et al.* 2013). Problems with drawing conclusions about adaptation from Fst outlier tests alone are understood and can arise from sample site selection, incorrect assumptions about the migration rates between subpopulations (LUIKART *et al.* 2003), and admixture (GOSSET and BIERNE 2013; KAO *et al.* 2015).

No significant non-independence of tests was found when a simple test was applied in pairwise comparisons (Table 5). This is not an unexpected result as each of these tests focuses on a different aspect of the data. There is a wide range of timescales considered between the polymorphism of North American flies (clinal variation and Fst outlier tests) and divergence from *D. simulans* (MK and HKA tests). Even within time scales, there is a difference between the Fst outlier test and the factor analysis from Lavington et al (2014) results and between the MK and HKA test results. Again, this is not altogether surprising as each of the tests examines a different aspect of the data even as data is shared between them. Taken together, we see only one candidate, *Irp-1B*, that violated neutrality for all three tests, and several that violate neutrality in two or three tests (Fig 2 & Table 6). There are gaps to fill in for the tests to cover all genes evenly, but the patterns of significant results described here support serious consideration of the glycerophosphate shuttle, malate-aspartate shuttle, and citrate cycling between the mitochondrion and cytoplasm for further experiments.

Some correlation between the factor analysis and Fst outlier test could have arisen due to the population sampling that was specifically designed to follow a potential gradient (FOURCADE *et al.* 2013). Extreme cases of demographic differences such as admixture or bottlenecks at any point along the cline, or sampling that happened to not be representative of that subpopulation, could have a significant effect but moderate differences, including isolation by distance should not (BEAUMONT and NICHOLS 1996; WILKINS 2004).

Multiple testing was of major concern, but it is important to note that the purpose of this study was to find candidate genes for use in experiments, and not to make statements about adaptation. The trends in the measures are of interest here, not necessarily statistical significance of each case made robust to multiple testing. That being said, each of these tests is reasonably conservative, in that it is difficult to violate the null hypothesis of neutrality and the differing results of each test highlight a problem with using only one type of data, let alone one type of test, for finding candidate genes.

These tests were a preliminary exploration of many different signals of adaptation and how results from each test of adaptation compared to all others. This has generated a set of candidate genes to be validated by experiments, as well as further analyses of the data used here. Ongoing work (Talbet et al, in revision) based on six candidate genes drawn from Lavington et al (2014) has found significant effects of enzyme activity knockdown of some of these candidates and a significant effect of all six tested candidates on starvation resistance. Of particular interest as candidates for future work are the two components of the glycerophosphate shuttle, *Gpdh* and *Gpo-1* for their role in cellular reduction-oxidation state (O'BRIEN and MACINTYRE 1972) and this work brings citrate cycling, particularly *Irp-1B* as obvious candidates for experimentation.

More work can also be done with the results used here. For instance, genes that fail to deviate from neutrality would be the best candidates for use in determining reasonable estimates of effective population size, Ne and migration rates, Nm. The history of these particular genes should reflect the demographic history. Multiple inferences of changing Ne could be made given that our lab has geographic sampling of 20 populations, temporal data for one of these populations, and public access to whole genome sequences of the DGRP (WAPLES 2005). Also, a more detailed look at the Fst results could determine which, if any, populations are generally differentiated from the rest and re-run the Fst outlier test without this population (BEAUMONT and NICHOLS 1996).

Also, important tests on data from one population for polymorphism data like Tajima's D (TAJIMA 1989), and polymorphism data using an outgroup Fay and u's H (FAY and WU 2000) and Fu and LI's D and F tests (FU and LI 1993) that are sensitive to demographic changes and selection were not considered here. The design of this study was simple and included two tests for each of two data types. The number and types of tests considered here were chosen somewhat arbitrarily. A more comprehensive future approach to testing correlation would include these measures.

This approach could easily be quantitative as well, disregarding the statistical cutoff and using quantitative, as opposed to the categorical significant/not significant, inputs. Correlation between tests and patterns of significance were naïve to the network structure. Connectivity of enzymes and substrates can also be quantified and included the in analysis. Excluding the network structure facilitated analysis and was considered conservative, but at the cost of confidence in the final inferences.

The classic tests considered here had been developed in a time of limited data but are finding traction in this population genomics era. However, the findings here caution against using only one test of selection. This may mean altering sampling design to include a geographic or temporal component.

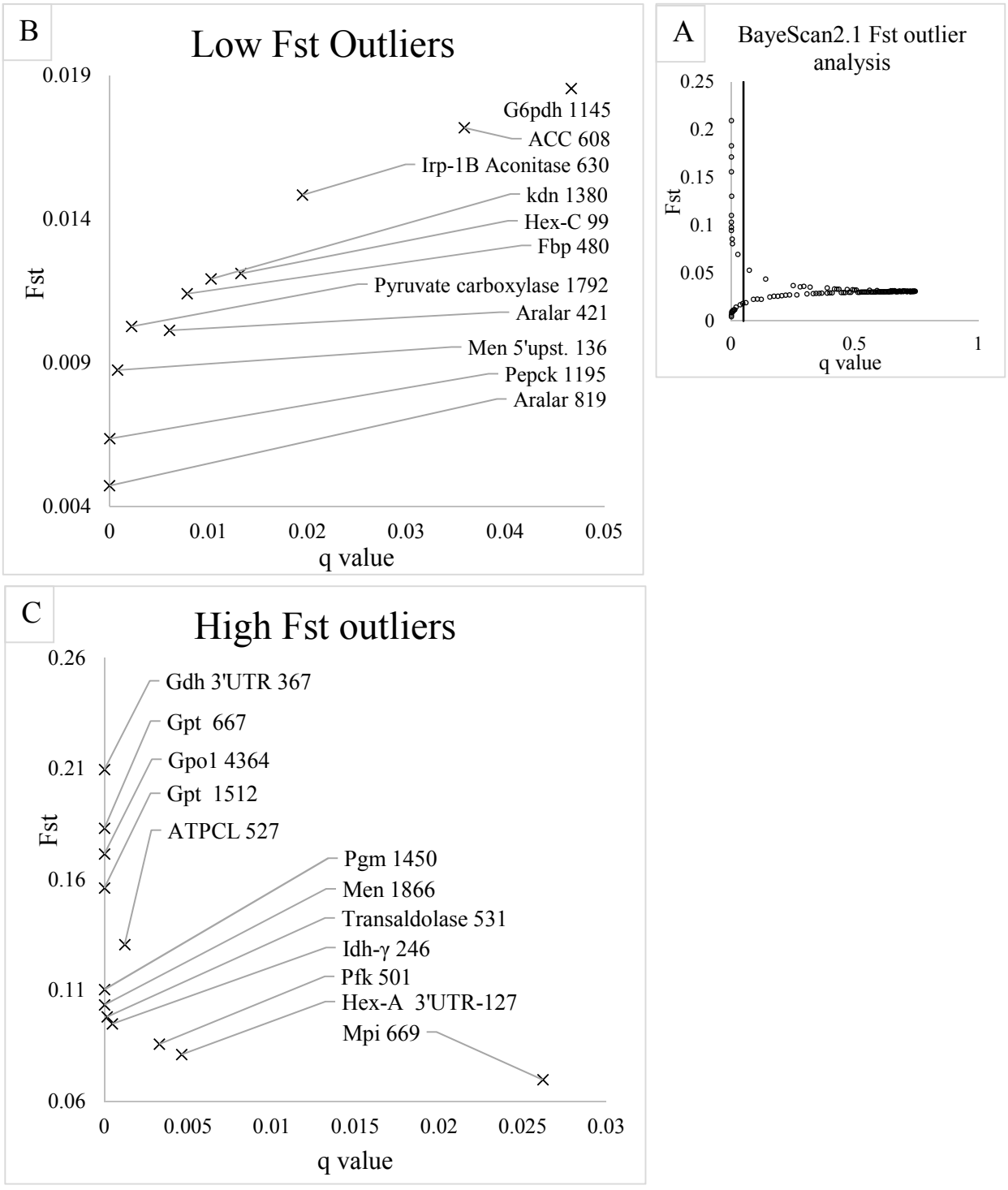


Figure 1. Results of Fst outlier test performed with BayeScan2.1. All results (A). Detail of BayeScan2.1 results for low Fst value outliers (B). Detail of BayeScan2.1 results for high Fst value outliers (C). Vertical lines represent q=0.05.

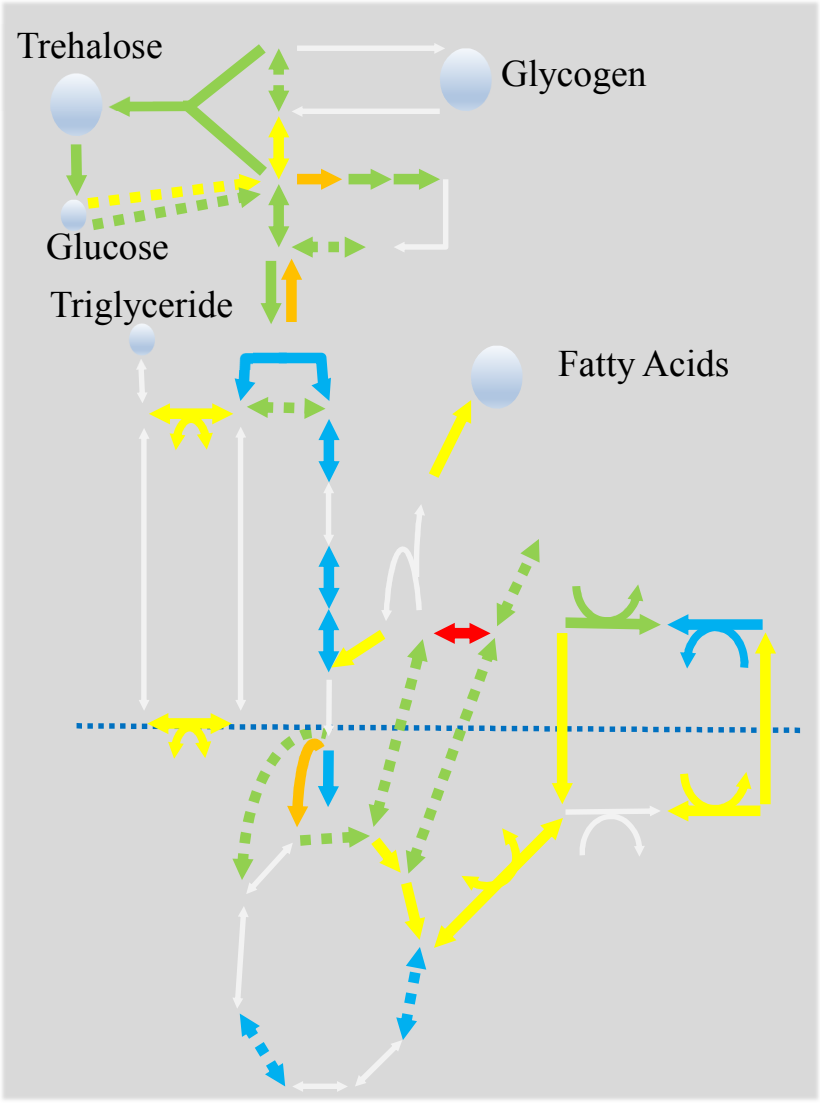


Fig 2. Number of significant tests by gene. Number of significant tests is represented as color: Blue = 0, Green = 1, Yellow = 2, Orange = 3, Red = 4. Dotted lines are genes where only three tests were completed. Grey lines are genes that were not tested. The dotted horizontal blue line

CG number	Gene	NI	Alpha	Fisher's exact test p-value	G-test G value	G-test p-value
CG11198	ACC	3.315	-2.315	*0.001239*	11.247	***0.0008***
CG3481	Adh	0.438	0.563	0.581538	0.559	0.45449
CG6058	Ald	0.346	0.654	0.564783	0.71	0.39936
CG2139	aralar1	10	-9	*0.02781*	6.977	*0.00826*
CG30499	Ribulose-phosphate 3 epimerase	0	1	1	NA	NA
CG4706	mito Acon	0.174	0.826	0.071357	4.306	*0.03798*
CG9244	mito Acon	10.444	-9.444	0.059935	4.083	*0.04331*
CG7176	cIdh	0	1	0.322474	NA	NA
CG17654	Eno	0.302	0.698	0.262105	2.184	0.13944
CG31692	fbp	9.333	-8.333	*0.046775*	4.7	*0.03015*
CG12055	Gapdh1	4.286	-3.286	0.310777	1.531	0.21595
CG5320	Gdh	2.796	-1.769	0.383117	1.246	0.2643
CG7254	GlyP	0.141	0.859	*0.005941*	8.767	*0.00307*
CG8430	Got1	0.295	0.705	0.125855	3.355	0.06699
CG4233	Got2	0.092	0.908	*0.003851*	10.149	*0.00144*
Cg9042	Gpdh	NA	NA	1	NA	NA
Cg8256	Gpo-1	1.585	-0.585	0.694226	0.278	0.59787
CG8094	Hex-C	2.583	-1.583	NA	NA	NA
CG32026	Idh-alpha	NA	NA	0.22222	NA	NA
CG6439	Idh-beta	0.531	0.469	1	0.189	0.664
CG5208	Idh-gamma	0.833	0.167	1	0.021	0.88565
CG6342	Irp-1B (cyto Acon)	15	-14	*0.002181*	10.37	*0.00128*
CG5214	Kdh	3.704	-2.704	0.180706	2.361	0.12437
CG7430	KdhE3	2.444	-1.444	0.57479	0.678	0.41041
CG11661	KgdE1	0.481	0.519	0.645043	0.422	0.51618
CG5362	Mdh1	12.75	-11.75	*0.026534*	6.775	*0.00924*
CG10120	Men	0.731	0.269	0.746149	0.195	0.65906
CG5889	Men-b	0	1	1	NA	NA

CG number	Gene	NI	Alpha	Fisher's exact test p-value	G-test G value	G-test p-value
CG8417	Mpi	3.15	-2.15	0.320129	1.537	0.21503
CG8808	Pdk	5.077	-4.077	0.28638	2.124	0.14503
CG17725	Pepck	0.49	0.51	0.308082	1.089	0.29668
CG4001	Pfk	1.833	-0.833	1	0.176	0.67522
CG8251	Pgi	16.667	-15.667	*0.015152*	7.074	*0.00782*
CG1721	Pglym78	NA	NA	NA	NA	NA
CG17645	Pglym87	2.667	-1.667	0.384123	1.319	0.25085
CG5165	Pgm	92.4	-91.4	**0**	30.224	**0**
CG7070	Pyk	1.444	-0.444	1	0.123	0.72571
CG7471	Rpd3	1.974	-0.974	0.691604	0.652	0.41923
CG1065	Scsalpha	NA	NA	0.507692	NA	NA
CG5718	Sdh-A	1.165	-0.165	0.801015	0.086	0.76962
CG6666	SdhC	1.179	-0.179	1	0.041	0.83895
CG6629	Sdh-C	NA	NA	0.307692	NA	NA
CG10219	SdhD	1.6	-0.6	1	0.116	0.73314
CG6782	sea	0	1	1	NA	NA
CG11963	skap	0.417	0.583	0.647844	0.659	0.41706
CG10622	SucB	4.483	-3.483	0.214292	2.27	0.13193
CG2827	TA	NA	NA	NA	NA	NA
CG2171	Tpi	NA	NA	0.507692	NA	NA
CG4104	Tps1	0.152	0.848	*0.00733*	8.47	*0.00361*
CG9364	Treh	0.223	0.777	*0.001619*	10.507	*0.00119*
CG9122	Trh	0.471	0.529	0.42846	0.871	0.35064
CG4347	UGP	NA	NA	0.242967	NA	NA
CG3724	6Pgd	1.762	-0.762	0.614615	0.361	0.54811
CG12529	G6PD	0.128	0.872	*0.00371*	10.113	*0.00147*
CG1640	Gpt	0.641	0.659	0.138318	2.625	0.1052
CG3001	Hex-A	NA	NA	NA	NA	NA
CG6861	kdn	0	1	1	NA	NA
CG7010	Pdh	1.714	-0.714	1	0.183	0.6689
CG12151	Pdp	0.342	0.658	0.138312	2.907	0.08819
CG3400	Pfrx	NA	NA	NA	NA	NA
CG17333	Pgls	7	-6	*0.022769*	5.668	*0.01727*

Table 1. Results of McDonald-Kreitman tests. Some could not be calculated due to the presence of random zeros in the contingency table. P-values for genes that significantly deviate from neutrality are in italics. P-values for genes that significantly deviate from neutrality after Bonferroni correction are in bold.

		Polymorphic sites within *D. melanogaster*			
Gene	**inheritance**	**observed**	**expected**	**variance**	**deviation**
ACC	Autosomal	80	92.81	892.23	0.184
Adh	Autosomal	18	11.57	24	1.721
Ald	Autosomal	14	10.18	19.81	0.735
aralar1	Autosomal	36	30.01	113.6	0.316
CG30499	Autosomal	6	8.69	15.7	0.461
CG4706 (mAcon)	Autosomal	23	41.92	205.05	1.746
CG9244 (mAcon)	Autosomal	5	19.85	56.44	3.909
cIdh	Autosomal	6	23.92	77.04	4.169
Eno	Autosomal	20	18.61	50.76	0.038
Fbp	Autosomal	16	17.92	47.73	0.077
Gapdh1	Autosomal	9	8.78	15.93	0.003
Gdh	Autosomal	17	15.06	36.1	0.104
GlyP	Autosomal	35	35.19	150.16	0
Got1	Autosomal	21	23.31	73.75	0.072
Got2	Autosomal	31	21.6	64.91	1.361
Gpdh	Autosomal	23	13.87	31.74	2.624
Gpo-1	Autosomal	56	41.95	205.3	0.961
Hex-C	Autosomal	14	18.98	52.42	0.473
Idh-alph	Autosomal	13	10.49	20.72	0.303
Idh-beta	Autosomal	31	21.01	61.98	1.61
Idh-gamma	Autosomal	13	14.81	35.15	0.093
Irp-1B (cAcon)	Autosomal	17	33.28	136.09	1.948
Kdh	Autosomal	13	15.53	37.93	0.169
KdhE3	Autosomal	11	13.02	28.75	0.142
KgdE1	Autosomal	28	28.23	102.22	0.001
Mdh1	Autosomal	14	13.22	29.45	0.02
Men	Autosomal	19	26.61	92.34	0.627
Men-b	Autosomal	28	24.74	81.54	0.131
Mpi	Autosomal	7	11.96	25.24	0.975
PCB	Autosomal	40	52.34	306.61	0.497
Pdk	Autosomal	16	15.29	36.99	0.014
Pepck	Autosomal	54	34.5	145.01	2.621
Pfk	Autosomal	25	27.05	94.95	0.044
Pgi	Autosomal	10	13.32	29.79	0.37
Pglym78	Autosomal	14	10.72	21.38	0.504
Pglym87	Autosomal	10	11.65	24.25	0.112
Pgm	Autosomal	16	21.99	66.87	0.537

		Polymorphic sites within *D. melanogaster*			
Gene	**inheritance**	**observed**	**expected**	**variance**	**deviation**
Rpd3	Autosomal	43	32.55	130.91	0.834
Scsalpha	Autosomal	16	11.5	23.76	0.854
Sdh-A	Autosomal	54	43.68	220.77	0.482
SdhC	Autosomal	4	4.97	7.26	0.13
Sdh-C	Autosomal	12	16.52	41.87	0.489
SdhD	Autosomal	7	5.14	7.58	0.458
Sea	Autosomal	1	14.38	33.57	5.332
Skap	Autosomal	7	13.05	28.85	1.268
SucB	Autosomal	33	26.75	93.15	0.42
Taldo	Autosomal	16	13.19	29.35	0.268
Tpi	Autosomal	16	10.4	20.43	1.536
Tps1	Autosomal	52	49.83	280.25	0.017
Treh	Autosomal	56	41.61	202.33	1.023
Trh	Autosomal	37	26.22	90.05	1.29
UGP	Autosomal	37	28.18	101.91	0.763
6Pgd	Sex-linked	8	16.3	40.95	1.681
G6PD	Sex-linked	23	27.39	97.03	0.199
Gpt	Sex-linked	27	24.32	79.21	0.091
Hex-A	Sex-linked	13	10.37	20.36	0.339
kdn	Sex-linked	8	10.87	21.85	0.378
Pdh	Sex-linked	16	10.05	19.41	1.827
Pdp	Sex-linked	22	18.8	51.63	0.198
Pfrx	Sex-linked	11	13.64	30.92	0.226
Pgls	Sex-linked	9	13.27	29.63	0.616

Table 2. HKA test results for polymorphic sites within *D. melanogaster*

Gene	inheritance	Divergent sites between species			
		observed	expected	variance	deviation
ACC	Autosomal	178.35	165.55	659.77	0.249
Adh	Autosomal	14.22	20.64	28.33	1.458
Ald	Autosomal	14.35	18.17	24.12	0.604
aralar1	Autosomal	47.54	53.53	105.21	0.341
CG30499	Autosomal	18.19	15.5	19.83	0.365
CG4706 (mAcon)	Autosomal	93.7	74.78	175.63	2.039
CG9244 (mAcon)	Autosomal	50.27	35.42	58.04	3.802
cIdh	Autosomal	60.59	42.67	75.51	4.254
Eno	Autosomal	31.81	33.2	53.08	0.036
Fbp	Autosomal	33.89	31.97	50.4	0.073
Gapdh1	Autosomal	15.43	15.66	20.08	0.002
Gdh	Autosomal	24.92	26.86	39.87	0.095
GlyP	Autosomal	62.97	62.78	133.85	0
Got1	Autosomal	43.89	41.58	72.76	0.073
Got2	Autosomal	29.14	38.53	65.31	1.352
Gpdh	Autosomal	15.62	24.75	35.79	2.327
Gpo-1	Autosomal	60.78	74.83	175.82	1.123
Hex-C	Autosomal	38.84	33.86	54.53	0.455
Idh-alph	Autosomal	16.22	18.72	25.04	0.251
Idh-beta	Autosomal	27.49	37.48	62.8	1.589
Idh-gamm	Autosomal	28.22	26.41	38.99	0.084
Irp-1B (cAcon)	Autosomal	75.65	59.37	122.93	2.156
Kdh	Autosomal	30.24	27.71	41.56	0.154
KdhE3	Autosomal	25.24	23.22	32.95	0.124
KgdE1	Autosomal	50.59	50.36	96.1	0.001
Mdh1	Autosomal	22.81	23.59	33.62	0.018
Men	Autosomal	55.08	47.47	88.11	0.658
Men-b	Autosomal	40.87	44.13	79.24	0.134
Mpi	Autosomal	26.3	21.34	29.55	0.833
PCB	Autosomal	105.7	93.36	250.56	0.608
Pdk	Autosomal	26.57	27.28	40.69	0.012
Pepck	Autosomal	42.05	61.55	129.87	2.927
Pfk	Autosomal	50.3	48.25	90.23	0.046
Pgi	Autosomal	27.08	23.76	33.94	0.325
Pglym78	Autosomal	15.84	19.12	25.71	0.419
Pglym87	Autosomal	22.43	20.78	28.57	0.095
Pgm	Autosomal	45.22	39.23	66.97	0.536
Rpd3	Autosomal	47.62	58.07	118.88	0.918

Gene	inheritance	Divergent sites between species			
		observed	expected	variance	deviation
Scsalpha	Autosomal	16	20.5	28.09	0.723
Sdh-A	Autosomal	67.59	77.92	187.4	0.568
SdhC	Autosomal	9.84	8.87	10.29	0.092
Sdh-C	Autosomal	34	29.48	45.14	0.453
SdhD	Autosomal	7.3	9.16	10.67	0.326
Sea	Autosomal	39.03	25.65	37.51	4.771
Skap	Autosomal	29.32	23.28	33.05	1.107
SucB	Autosomal	41.46	47.71	88.76	0.44
Taldo	Autosomal	20.73	23.54	33.53	0.235
Tpi	Autosomal	12.95	18.55	24.75	1.268
Tps1	Autosomal	86.7	88.88	231.33	0.02
Treh	Autosomal	59.84	74.23	173.58	1.193
Trh	Autosomal	36	46.78	86.24	1.347
UGP	Autosomal	41.46	50.27	95.86	0.811
6Pgd	Sex-linked	45.76	37.46	52.7	1.306
G6PD	Sex-linked	67.35	62.96	106.01	0.182
Gpt	Sex-linked	53.22	55.9	89.83	0.08
Hex-A	Sex-linked	21.22	23.84	30.02	0.23
kdn	Sex-linked	27.86	24.99	31.78	0.26
Pdh	Sex-linked	17.14	23.09	28.88	1.228
Pdp	Sex-linked	40.03	43.22	63.51	0.161
Pfrx	Sex-linked	34	31.36	42.04	0.166
Pgls	Sex-linked	34.78	30.51	40.62	0.45

Table 3. HKA test results for divergent sites between *D. melanogaster* and *D. simulans*

SNP	prob	log10(PO)	q value	alpha	Fst
6Pgd 616	0.18484	-0.64446	0.27861	0.11159	0.035828
ACC 2814	0.035607	-1.4327	0.73039	0.001305	0.031304
ACC 48	0.11482	-0.887	0.41696	0.060329	0.033686
ACC 608	0.76115	0.50335	**0.035862**	-0.7593	0.017181
Acon 1382	0.04761	-1.3011	0.65721	0.014312	0.031751
Adh 39	0.041608	-1.3624	0.69354	-0.00483	0.031131
Adh 578	0.085817	-1.0275	0.49753	0.039647	0.032814
Aralar 2014	0.052611	-1.2555	0.62232	-0.01696	0.030851
Aralar 421	0.97239	1.5468	**0.006051**	-1.3157	0.010125
Aralar 819	1	1000	**0**	-2.1032	0.004716
Aralar 843	0.036607	-1.4202	0.7226	-0.00609	0.031091
ATPCL 1364	0.039808	-1.3824	0.69842	-0.00074	0.031248
ATPCL 1370	0.041808	-1.3602	0.6859	-0.00742	0.031066
ATPCL 3138	0.069014	-1.13	0.5639	-0.03242	0.030513
ATPCL 527	0.9944	2.2493	**0.001217**	1.5202	0.13074
ATPCL 684	0.038608	-1.3962	0.70994	-0.00099	0.03124
cIdh 489	0.041808	-1.3602	0.6859	-0.00717	0.031064
cIdh 5'UTR-106	0.04781	-1.2992	0.65404	-0.01198	0.030955
cIdh 735	0.14243	-0.77967	0.34265	-0.0912	0.029289
Eno 231	0.073615	-1.0998	0.53023	-0.03317	0.030504
Eno 423	0.041208	-1.3667	0.696	0.000925	0.031294
Fbp 367	0.039608	-1.3847	0.7008	0.008879	0.031553
Fbp 480	0.96339	1.4202	**0.007849**	-1.1678	0.011398
G6pdh 1145	0.71514	0.39977	**0.046688**	-0.665	0.018544
G6pdh 1407	0.11822	-0.87265	0.39749	-0.07495	0.029657
G6pdh 738	0.04861	-1.2916	0.64749	-0.01552	0.030872
Gapdh 850	0.033407	-1.4614	0.7395	0.003307	0.031342
Gapdh1 825	0.054611	-1.2383	0.6061	-0.01974	0.030787
Gdh 1062	0.041608	-1.3624	0.69354	0.010477	0.031634
Gdh 216	0.070814	-1.118	0.55865	0.0333	0.032545
Gdh 3'UTR 367	1	1000	**0**	2.2145	0.20967
Gdh 5'UTR 21	0.042008	-1.358	0.68056	0.011672	0.031669
Gdh 84	0.079616	-1.063	0.50446	0.037091	0.032686
Gdh 986	0.066013	-1.1507	0.57402	-0.02664	0.030628
Got-1 1225	0.22725	-0.53155	0.26411	-0.17906	0.027607
Got-1 246	0.36147	-0.2471	0.15549	-0.30227	0.025242
Got-1 984	0.055411	-1.2316	0.60181	-0.02088	0.030763
Got-2 27	0.053611	-1.2468	0.6144	-0.0171	0.030829
Got-2 501	0.090218	-1.0036	0.49046	0.043958	0.032955
Got-2 882	0.037608	-1.4081	0.71214	0.005862	0.03145

SNP	prob	log10(PO)	q value	alpha	Fst
Gpdh 1011	0.31726	-0.33283	0.18819	-0.25532	0.026109
Gpdh 108	0.031406	-1.4891	0.74126	-0.0011	0.031211
Gpdh 330	0.036607	-1.4202	0.7226	0.002486	0.031359
Gpdh 741	0.034207	-1.4508	0.73772	-0.00074	0.031238
Gpdh 828	0.47289	-0.04713	0.10551	-0.39958	0.023295
Gpo1 162	0.061212	-1.1857	0.58836	0.023021	0.03211
Gpo1 3946	0.36267	-0.24485	0.13883	0.2869	0.044184
Gpo1 4194	0.05241	-1.2572	0.62615	-0.0147	0.030885
Gpo1 4364	1	1000	**0**	1.9047	0.17156
Gpo1 4833	0.45989	-0.06983	0.12103	-0.4407	0.02296
Gpt 1512	1	1000	**0**	1.7987	0.15618
Gpt 494	0.038608	-1.3962	0.70994	-0.00704	0.031062
Gpt 667	1	1000	**0**	2.0161	0.1832
Gpt 99	0.05021	-1.2768	0.64067	-0.01493	0.030906
Hex-A 3'UTR-127	0.9766	1.6204	**0.004614**	0.99983	0.081097
Hex-A 3'UTR-205	0.24705	-0.48398	0.23514	-0.18622	0.027441
Hex-A 3'UTR-293	0.11662	-0.87936	0.4074	-0.07444	0.029698
Hex-C 636	0.067413	-1.1409	0.56902	-0.03043	0.030549
Hex-C 99	0.93279	1.1423	**0.013266**	-1.1296	0.012105
Hext-1 728	0.078216	-1.0713	0.51769	-0.04014	0.030359
Idh-α 41	0.064613	-1.1607	0.5789	-0.02809	0.030591
Idh-α 576	0.073415	-1.1011	0.53623	-0.03669	0.030417
Idh-β 101	0.035407	-1.4353	0.73411	0.004299	0.031396
Idh-β 1047	0.094819	-0.97984	0.47575	-0.05064	0.030116
Idh-β 630	0.23085	-0.52269	0.24998	0.14349	0.037173
Idh-γ 246	0.9968	2.4934	**0.00046**	1.1836	0.09492
Idh-γ 852	0.036207	-1.4252	0.72459	-0.00314	0.031178
Irp-1B Aconitase 1259	0.05061	-1.2732	0.63716	-0.01686	0.030854
Irp-1B Aconitase 1521	0.039408	-1.387	0.70314	-0.00143	0.031217
Irp-1B Aconitase 630	0.86197	0.79553	**0.019504**	-0.90372	0.014838
kdn 1380	0.94859	1.266	**0.010269**	-1.1271	0.011912
kdn 63	0.036607	-1.4202	0.7226	-0.00574	0.0311
KgdE1 1401	0.25465	-0.46641	0.21991	-0.19974	0.0272
KgdE1 2997	0.11382	-0.89129	0.42616	0.058583	0.033575
Mdh-2 480	0.071414	-1.114	0.55328	-0.03439	0.030471
Mdh-2 963	0.043209	-1.3452	0.67782	-0.00062	0.031259
Men 1419	0.17984	-0.65902	0.30559	-0.12515	0.028619
Men 1866	1	1000	**0**	1.2998	0.10349
Men 5'upst. 136	0.9956	2.3545	**0.000818**	-1.446	0.008743
Men 881	0.10782	-0.91775	0.4437	-0.05961	0.029969

SNP	prob	log10(PO)	q value	alpha	Fst
Men-b 1705	0.037207	-1.4129	0.7143	0.004206	0.03138
Men-b 45	0.11002	-0.9079	0.43508	0.058635	0.033602
Men-b 504	0.062212	-1.1782	0.58368	-0.02642	0.030635
Mpi 669	0.83997	0.72006	**0.026196**	0.79854	0.069828
Pdh 357	0.038608	-1.3962	0.70994	-0.0072	0.031072
Pdh 534	0.079216	-1.0653	0.51117	-0.03952	0.030347
Pdh 702	0.04921	-1.286	0.64412	0.013923	0.031757
Pdk 675	0.13143	-0.82012	0.38718	0.081477	0.034736
Pdk 870	0.044609	-1.3308	0.67503	-0.00648	0.031089
Pdp 1313	0.027606	-1.5468	0.74302	0.003322	0.031352
Pdp 1399	0.071614	-1.1127	0.54776	-0.03084	0.030555
Pdp 726	0.036607	-1.4202	0.7226	0.001682	0.031303
Pdp 78	0.035007	-1.4404	0.73593	-0.0029	0.031166
Pepck 1158	0.04841	-1.2935	0.6508	0.015502	0.031803
Pepck 1195	0.9998	3.6988	**2.50E-05**	-1.7692	0.006354
Pepck 1479	0.10642	-0.9241	0.45203	-0.06059	0.029921
Pfk 2016	0.035607	-1.4327	0.73039	-0.00234	0.031185
Pfk 501	0.983	1.762	**0.003272**	1.0642	0.085879
Pfrx 1548	0.045009	-1.3267	0.6722	-0.00738	0.031064
Pfrx 303	0.16923	-0.69099	0.33039	-0.11797	0.028785
Pgi 24	0.035407	-1.4353	0.73411	0.001107	0.031303
Pgi 985	0.66813	0.3039	0.05857	-0.61205	0.019466
Pgk 333	0.090418	-1.0026	0.48323	-0.04523	0.030249
Pgk 639	0.4985	-0.00261	0.089295	-0.42257	0.022846
Pgls 19	0.072014	-1.1101	0.54208	-0.03339	0.030509
Pgls 308	0.09902	-0.95899	0.4602	-0.06026	0.029987
Pgls 607	0.05121	-1.2678	0.63357	-0.01352	0.030947
Pglym78 144	0.053011	-1.252	0.61841	-0.01835	0.030805
Pglym78 255	0.18384	-0.64734	0.29239	0.12469	0.036617
Pgm 129	0.053811	-1.2451	0.6103	-0.01666	0.030843
Pgm 1393	0.13143	-0.82012	0.38718	-0.08378	0.029455
Pgm 1450	1	1000	**0**	1.3683	0.11046
Pgm 155	0.046609	-1.3108	0.66337	-0.01122	0.030974
Pyk 1376	0.09802	-0.96388	0.46809	0.050952	0.033285
Pyk 180	0.046209	-1.3147	0.66637	-0.01415	0.030892
Pyk 711	0.13243	-0.81633	0.36579	-0.0847	0.029437
Pyruvate carboxylase 1693	0.585517	0.150033	0.072807	0.48728	0.053284
Pyruvate carboxylase 1774	0.047409	-1.303	0.66032	0.015304	0.031801

SNP	prob	log10(PO)	q value	alpha	Fst
Pyruvate carboxylase 1792	0.9858	1.8414	**0.002216**	-1.2703	0.010257
Pyruvate carboxylase 2631	0.059012	-1.2026	0.59294	-0.02101	0.030757
SdhD 261	0.057812	-1.2121	0.59741	-0.01945	0.030786
Tpi 234	0.17784	-0.66494	0.31819	0.1093	0.035753
Tpi 399	0.32527	-0.3169	0.17224	-0.24564	0.026154
Tps1 2409	0.036007	-1.4277	0.72656	-0.003	0.031162
Tps1 303	0.05121	-1.2678	0.63357	-0.01439	0.030904
Transaldolase 129	0.076015	-1.0848	0.52404	-0.03489	0.03045
Transaldolase 531	0.9988	2.9202	**0.000156**	1.2265	0.098096
Treh 1071	0.29046	-0.38789	0.20399	-0.22503	0.026653
Treh 122	0.14103	-0.78467	0.35438	-0.09688	0.029241
UGPase 318	0.041608	-1.3624	0.69354	0.01014	0.031599
UGPase 45	0.045409	-1.3227	0.66931	-0.0098	0.030989

Table 4. Table of results of Fst outlier analysis of SNP data from 20 populations in BayeScan2.1

	Significant results				
	both	MK only	other only	neither	p value
MK vs Factor analysis	7	8	8	24	0.1844
MK vs Fst outliers	7	8	8	26	0.1772
MK vs HKA	2	14	2	31	0.5879
	both	HKA only	other only	neither	
HKA vs Fst outliers	2	5	17	40	1
HKA vs Factor analysis	3	4	12	27	0.6666
	both	Fst only	other only	neither	
Fst outliers versus Factor analysis	7	11	8	19	0.5381

Table 5. Values entered into 2x2 contingency table and two-tailed p value results of Fisher's exact test.

Gene	Factor analysis	Fst outlier	MK	HKA
Irp-1B (cAcon)	yes	yes	yes	yes
Fbp	yes	yes	yes	no
PC	yes	yes	yes	no
G6PD	yes	yes	yes	no
Gdh	yes	yes	no	no
Gpo-1	yes	yes	no	no
Hex-A	yes	yes	—	no
Got2	yes	no	yes	no
Idh-beta	yes	no	no	yes
6Pgd	yes	no	no	no
Pdk	yes	no	no	no
Gpdh	yes	no	—	yes
Men-b	yes	no	—	no
Tpi	yes	no	—	no
UGP	yes	no	—	no
aralar1	no	yes	yes	no
ACC	no	yes	yes	no
Pgm	no	yes	yes	no
Pepck	no	yes	no	yes
Idh-gamma	no	yes	no	no
Pfk	no	yes	no	no
Men	no	yes	no	no
Gpt	no	yes	no	no
Taldo	no	yes	—	no
Hex-C	no	yes	—	no
kdn	no	yes	—	no
Pgi	no	no	yes	no
mAcon (CG4706)	no	no	yes	yes
Pgls	no	no	yes	no
Mdh1	no	no	yes	no
Tps1	no	no	yes	no
Treh	no	no	yes	no
Eno	no	no	no	no
Got1	no	no	no	no
KgdE1	no	no	no	no
Pdp	no	no	no	no
Pglym87	no	no	no	no
Ald	no	no	no	no
Pdh	no	no	no	no
Gapdh1	no	no	no	no

Gene	Factor analysis	Fst outlier	MK	HKA
SdhD	no	no	no	no
Idh (cytosolic)	no	no	—	yes
Idh-alpha	no	no	—	no
sea	no	no	—	yes
Pfrx	no	no	—	no
Pglym78	no	no	—	no
Mpi	—	yes	no	no
mAcon (CG9244)	—	no	yes	no
Adh	—	no	no	no
Kdh	—	no	no	no
KdhE3	—	no	no	no
SucB	—	no	no	no
Rpd3	—	no	no	no
Sdh-A	—	no	no	no
SdhC	—	no	no	no
skap	—	no	no	no
Trh	—	no	no	no
CG30499	—	no	—	no
Sdh-C	—	no	—	no
Scsalpha	—	no	—	no
GlyP	—	—	yes	no

Table 6. Significant results of tests by gene. Results of each test for each gene are significant "yes", not significant "no" or were not performed "—".

Chapter 2 Supplementary Tables

CG number	Gene	linkage	Sites	S	η	π	k	Θ per site	Θ per gene
CG11198	ACC	Autosomal	7583	80	80	0.002349	17.8108	0.002527	19.1637
CG3481	Adh	Autosomal	771	18	18	0.005278	4.0691	0.005593	4.3118
CG6058	Ald	Autosomal	1092	14	14	0.002591	2.8288	0.003071	3.3536
CG2139	aralar1	Autosomal	2049	36	36	0.004652	9.5315	0.004209	8.6237
CG30499	Ribulose-phosphate 3 epimerase	Autosomal	666	6	6	0.002448	1.6306	0.002158	1.4373
CG4706	mito Acon	Autosomal	2352	23	23	0.002778	6.5345	0.002343	5.5096
CG9244	mito Acon	Autosomal	2052	55	55	0.001569	3.2192	0.006421	13.175
CG7176	cIdh	Autosomal	1251	6	6	0.001248	1.5616	0.001149	1.4373
CG17654	Eno	Autosomal	1503	20	20	0.004234	6.3634	0.003188	4.7909
CG31692	Fbp	Autosomal	1005	16	16	0.00456	4.5826	0.003814	3.8327
CG12055	Gapdh1	Autosomal	998	9	10	0.002907	2.9009	0.0024	2.3955
CG5320	Gdh	Autosomal	1650	17	17	0.002996	4.9429	0.002468	4.0723
CG7254	GlyP	Autosomal	2535	35	37	0.003774	9.5661	0.003496	8.8632
CG8430	Got1	Autosomal	1251	21	21	0.004647	5.8138	0.004021	5.0305
CG4233	Got2	Autosomal	1182	31	31	0.007812	9.2342	0.006283	7.4259
Cg9042	Gpdh	Autosomal	1083	23	23	0.006333	6.8589	0.005087	5.5096
Cg8256	Gpo-1	Autosomal	2175	56	57	0.005285	11.4955	0.006278	13.6541
CG8094	Hex-C	Autosomal	1365	14	14	0.003698	5.048	0.002457	3.3536
CG32026	Idh-alpha	Autosomal	1134	13	13	0.002317	2.6276	0.002746	3.1141
CG6439	Idh-beta	Autosomal	1113	31	33	0.008565	9.533	0.007102	7.905
CG5208	Idh-gamma	Autosomal	1209	13	13	0.00303	3.6637	0.002576	3.1141
CG6342	Irp-1B (cyto Acon)	Autosomal	2700	17	17	0.001039	2.8048	0.001508	4.0723
CG5214	Kdh	Autosomal	1407	13	13	0.001722	2.4234	0.002213	3.1141
CG7430	KdhE3	Autosomal	1515	11	11	0.001277	1.9339	0.001739	2.635
CG11661	KgdE1	Autosomal	3027	28	28	0.001619	4.9009	0.002216	6.7073
CG5362	Mdh1	Autosomal	1014	14	14	0.001558	1.5796	0.003307	3.3536
CG10120	Men	Autosomal	2218	19	19	0.002308	5.1201	0.002052	4.5514
CG5889	Men-b	Autosomal	1849	28	28	0.004202	7.7688	0.003628	6.7073
CG8417	Mpi	Autosomal	1191	7	7	0.000482	0.5736	0.001408	1.6768
CG1516	PC	Autosomal	3546	40	40	0.003092	10.964	0.002702	9.5818

Gene	G+C content	Tajima's D	Fu & Li's D*	Fu& Li's F*	Fu Fs	K (JC)	Fu & Li's D	Fu& Li's F
ACC	0.5508	-0.2595	0.4682	0.2552	-7.785	0.024	0.5484	0.3214
Adh	0.5781	-0.1864	-0.573	-0.5262	-1.851	0.019	-0.2814	-0.2993
Ald	0.6133	-0.5011	-1.1342	-1.0941	-5.892	0.013	-0.8075	-0.8443
aralar1	0.5664	0.372	-0.0281	0.1277	-5.142	0.024	-0.0876	0.0358
Ribulose-phosphate 3 epimerase	0.5704	0.3653	-0.424	-0.213	-1.621	0.028	-0.4829	-0.2579
mito Acon	0.5024	0.633	0.7979	0.8781	-1.065	0.041	1.1928	1.2017
mito Acon	0.5588	*-2.7364***	*-5.4115***	*-5.328***	1.552	0.001	*-6.4492***	*-6.1588***
cIdh	0.5562	0.2348	0.387	0.3976	-2.72	0.05	0.3753	0.392
Eno	0.5949	1.1003	0.3103	0.6744	-4.265	0.021	0.6094	1.0201
Fbp	0.5504	0.6382	-0.0244	0.226	0.001	0.035	-0.0746	0.205
Gapdh1	0.5998	0.6392	-0.8768	-0.4658	-6.955	0.016	0.2445	0.4465
Gdh	0.6078	0.7029	0.0709	0.3275	-24.008	0.015	-0.0746	0.2092
GlyP	0.5726	0.2807	0.0219	0.1302	-19.769	0.025	0.3369	0.3557
Got1	0.5547	0.5248	0.3778	0.5036	-9.899	0.036	0.3754	0.5168
Got2	0.5811	0.851	-0.3191	0.0898	-2.72	0.025	-0.4228	0.0486
Gpdh	0.5577	0.8333	0.7979	0.9568	-1.488	0.015	1.1928	1.2877
Gpo-1	0.5912	-0.5735	-0.2679	-0.4432	-12.854	0.028	0.0678	-0.2059
Hex-C	0.5513	1.6177	0.6427	1.128	-0.26	0.029	0.6639	1.1853
Idh-alpha	0.5981	-0.4946	-0.8455	-0.8618	-1.462	0.014	-0.9745	-0.9743
Idh-beta	0.5675	0.7232	0.6909	0.8311	-4.174	0.025	0.7418	0.8923
Idh-gamma	0.5498	0.5588	-0.3756	-0.0879	0.501	0.024	0.059	0.2726
Irp-1B (cyto Acon)	0.553	-1.0233	-1.0757	-1.2482	-10.158	0.029	-1.2486	-1.4093
Kdh	0.5829	-0.7023	0.0943	-0.1909	-2.636	0.022	0.059	-0.2368
KdhE3	0.5855	-0.8198	-0.1552	-0.4311	-2.752	0.017	-0.3738	-0.6474
KgdE1	0.5753	-0.9333	-1.0446	-1.1916	-11.902	0.017	-1.251	-1.3771
Mdh1	0.5801	-1.6937#	-1.5784	-1.9022	-1.736	0.023	-1.7884	-2.1078#
Men	0.5594	0.4165	-0.1127	0.0727	-3.872	0.025	-0.1762	0.037
Men-b	0.5647	0.5484	0.4817	0.5964	-3.804	0.022	0.4975	0.6251
Mpi	0.5338	*-1.8527**	-2.3967#	*-2.6071**	-4.578	0.022	*-2.5902**	*-2.795**
PCB	0.5957	0.5132	-0.4025	-0.1068	-1.779	0.03	-0.5915	-0.2594

CG number	Gene	linkage	Sites	S	η	π	k	Θ per site	Θ per gene
CG8808	Pdk	Autosomal	1240	16	16	0.002073	2.5706	0.003091	3.8327
CG17725	Pepck	Autosomal	1503	54	54	0.00834	12.5345	0.008606	12.9355
CG4001	Pfk	Autosomal	2367	25	25	0.00298	7.0541	0.00253	5.9887
CG8251	Pgi	Autosomal	1674	10	10	0.001058	1.7718	0.001431	2.3955
CG1721	Pglym78	Autosomal	768	14	14	0.006585	5.0571	0.004367	3.3536
CG17645	Pglym87	Autosomal	879	10	10	0.003601	3.1652	0.002725	2.3955
CG5165	Pgm	Autosomal	1672	16	16	0.001516	2.5345	0.002292	3.8327
CG7070	Pyk	Autosomal	1539	48	48	0.004802	7.3904	0.007471	11.4982
CG7471	Rpd3	Autosomal	1566	43	43	0.007527	11.7868	0.006578	10.3005
CG1065	Scsalpha	Autosomal	987	16	16	0.003329	3.2853	0.003883	3.8327
CG5718	Sdh-A	Autosomal	1956	54	54	0.006385	12.4895	0.006613	12.9355
CG6666	SdhC	Autosomal	786	12	12	0.004547	3.5736	0.003657	2.8746
CG6629	Sdh-C	Autosomal	459	4	4	0.001995	0.9159	0.002088	0.9582
CG10219	SdhD	Autosomal	549	7	7	0.002976	1.6336	0.003054	1.6768
CG6782	sea	Autosomal	954	1	1	0.000057	0.0541	0.000251	0.2395
CG11963	skap	Autosomal	1509	7	7	0.0013	1.961	0.001111	1.6768
CG10622	SucB	Autosomal	1251	33	34	0.007497	9.3784	0.00651	8.1446
CG2827	TA	Autosomal	996	16	17	0.006567	6.5405	0.004089	4.0723
CG2171	Tpi	Autosomal	744	16	16	0.005429	4.039	0.005152	3.8327
CG4104	Tps1	Autosomal	2430	52	52	0.004632	11.2553	0.005126	12.4564
CG9364	Treh	Autosomal	1791	56	59	0.007732	13.8483	0.007891	14.1332
CG9122	Trh	Autosomal	1668	37	38	0.006719	11.2072	0.005457	9.1028
CG4347	UGP	Autosomal	1542	37	37	0.006586	10.1562	0.005748	8.8632
CG3724	6Pgd	Sex-linked	1396	8	8	0.001183	1.6517	0.001373	1.9164
CG12529	G6PD	Sex-linked	1574	23	23	0.003688	5.8048	0.0035	5.5096
CG1640	Gpt	Sex-linked	1707	27	27	0.004975	8.4925	0.003789	6.4677
CG3001	Hex-A	Sex-linked	1626	13	13	0.000525	0.8529	0.001915	3.1141
CG6861	kdn	Sex-linked	1393	8	8	0.001278	1.7808	0.001376	1.9164
CG7010	Pdh	Sex-linked	1200	16	16	0.003721	4.4655	0.003194	3.8327
CG12151	Pdp	Sex-linked	1428	22	22	0.00498	7.1111	0.00369	5.27
CG3400	Pfrx	Sex-linked	1611	11	11	0.002462	3.967	0.001636	2.635
CG17333	Pgls	Sex-linked	732	9	9	0.00304	2.2252	0.002945	2.1559

Gene	G+C content	Tajima's D	Fu & Li's D*	Fu& Li's F*	Fu Fs	K (JC)	Fu & Li's D	Fu& Li's F
Pdk	0.5273	-1.0742	-0.4251	-0.7505	-7.982	0.022	-0.5203	-0.8547
Pepck	0.5874	-0.1121	-0.1009	-0.1243	-4.991	0.029	-0.1845	-0.2926
Pfk	0.5836	0.6103	0.3231	0.4954	-5.489	0.022	0.1732	0.3526
Pgi	0.5723	-0.7887	-0.8768	-0.9964	-3.155	0.016	-0.992	-1.1042
Pglym78	0.6203	1.6263	1.0869	1.485#	-7.216	0.021	1.1544	1.5776#
Pglym87	0.5778	0.9734	0.8334	1.0303	1.485	0.026	0.7663	0.8854
Pgm	0.5416	-1.1049	0.3764	-0.1257	-3.338	0.028	0.3712	-0.1638
Pyk	0.5971	-1.2848	*-3.0514**	*-2.9003**	0.079	0.003	*-3.6426***	*-3.3634***
Rpd3	0.5335	0.5157	-0.5989	-0.2581	-7.879	0.031	-0.1434	0.1234
Scsalpha	0.6108	-0.4659	-0.0244	-0.1984	-4.83	0.016	0.3712	0.1004
Sdh-A	0.5479	-0.1247	0.0422	-0.0182	-4.06	0.035	0.3267	0.1912
SdhC	0.4727	0.7603	0.975	1.0655	0.408	0.045	1.024	1.1203
Sdh-C	0.5878	-0.1073	-0.016	-0.0503	0.38	0.022	-0.3634	-0.4932
SdhD	0.5971	-0.0725	0.5291	0.4014	-0.762	0.013	0.5275	0.4001
sea	0.5639	-1.1309	-1.7498	-1.8171	-1.385	0.042	-1.7817	-1.8534
skap	0.5252	0.4772	0.5291	0.6002	-3.681	0.02	0.5275	0.6068
SucB	0.5314	0.5331	1.3729#	1.2889	1.793	0.034	1.5314#	1.4162
TA	0.5707	1.9928#	*1.5997***	*2.0379***	0.175	0.021	*1.7083**	*2.2165***
Tpi	0.6453	0.1756	-0.4251	-0.2702	-5.539	0.018	-0.5203	-0.3381
Tps1	0.5747	-0.3482	-0.6214	-0.6258	-2.319	0.037	-0.6179	-0.6297
Treh	0.5404	-0.0732	0.4654	0.3316	-14.162	0.034	0.6458	0.4563
Trh	0.563	0.8199	0.4608	0.6909	-5.921	0.022	0.5802	0.8486
UGP	0.5202	0.5164	0.2221	0.3818	-9.477	0.027	0.1483	0.375
6Pgd	0.6106	-0.4005	-0.6878	-0.7006	-2.485	0.034	-0.7757	-0.7789
G6PD	0.5948	0.1823	0.7979	0.7008	-0.951	0.044	0.7983	0.7158
Gpt	0.566	1.0814	*1.4812**	1.5935#	-8.602	0.032	*1.6399**	1.7431#
Hex-A	0.562	*-2.2992***	*-3.6647***	*-3.7944***	-1.866	0.013	-2.0079#	*-2.5245**
kdn	0.6183	-0.2051	-0.0203	-0.0913	-2.178	0.02	0.6559	0.451
Pdh	0.5759	0.5385	-0.0244	0.1876	-0.101	0.014	-0.0746	0.1638
Pdp	0.5637	1.1833	0.751	1.0568	-6.539	0.029	0.7983	1.1289
Pfrx	0.5998	1.5575	-0.1552	0.4587	0.109	0.021	0.3682	0.9113
Pgls	0.5642	0.0955	-1.0966	-0.8464	-4.048	0.049	-1.2226	-0.9462

Supplementary table 1. Population genetic measures and tests of the DGRP with a single *Drosophila simulans* reference sequence as an outgroup for tests and measures of divergence. Tables produced using DNAsP 5.0 (LIBRADO and ROZAS 2009).

Chapter 3- Starvation Experiments with RNAi Knockdown of Metabolic Enzymes in AKH and dILP Producing Cells of *Drosophila melanogaster*

Background

Metabolism is of central importance to life. Understanding the control of flux through metabolic pathway is of interest for a broad scope of study. It is believed that most enzymes in the pathway have little effect on flux when their activity is modified and therefore have little *control* of flux (Fell 1998; Olson-Manning *et al.* 2013). Observations of large whole metabolic shifts or individual metabolic enzyme activity associations with pathology have moved, with our increased understanding, from assayable effects to defining causes and targets of treatment (Martins *et al.* 2006; Kroemer and Pouyssegur 2008; Fontana *et al.* 2010, to name a few). For example, cancer metabolism includes the founding, and the subsequent modification, of the Warburg effect which was initially thought to be the down regulation of mitochondrial oxidative phosphorylation, but now is accepted as the up regulation of the glycolytic pathway (Bensinger and Christofk 2012). In addition, metabolic enzymes have become potential targets for drug therapies alongside classical signaling pathways (Wu and Zhao 2013; Losman and Kaelin 2013). Research in longevity has implicated metabolic enzymes indirectly as downstream targets of the Insulin/Insulin-like signaling pathway (Partridge *et al.* 2011) and lifespan studies in our lab with *Drosophila melanogaster* have also shown direct effects of central metabolic enzyme activity on lifespan (Talbert et al., in review). Consequently, a better understanding of metabolic flux control and the role of metabolic enzymes in life history traits promises to yield insights into how these genes may be used as targets to treat human disease.

Experimental studies to evaluate flux control through a metabolic network require a great deal of prior knowledge and complex experiments for each enzyme. The goal of such metabolic control analysis is to measure the "control coefficients" for each enzyme by perturbing enzyme activity and assessing flux or some proxy. The control exercised by a step enzyme is calculated as the change in defined output of the network (flux) given the either increased or reduced activity of that enzyme. To understand the relative control across an entire network, this must be repeated for a range of perturbations for each enzyme (Fell 1992).

In an example of estimating the flux control from sucrose to starch in a potato tuber (Geigenberger *et al.* 2004) used data collected from several studies to determine control based on an approximate control coefficient.. They calculated the control coefficients from each of the studied enzymes. The data from experiments using tubers or isolated discs of tuber tissue were all normalized to respective full activity ("wild-type") levels for comparison across all of the enzymes. The control coefficients were calculated as the slope of curves near full activity of the enzyme, with the expectation that most would be small. With only limited information about the structure of the network, one might expect flux from sucrose to starch to be controlled at a step closer to the branch points of sucrose or glucose-6-phosphate (G6P) and not the amyloplastidial ATP/ADP translocator (AATPT), yet, this is what they found. This counterintuitive conclusion would also be missed if one were to focus on only the direct enzymatic pathway from sucrose to starch, thus highlighting the importance of including as many steps involved in the measured flux as possible.

Given the difficulty of experimentally quantifying flux control even for a single enzyme through a network as large and complex as central metabolism, it would be important to utilize other observations to produce a candidate list of enzymes that are most likely to control flux. We might assume that natural selection acts upon the flux of a pathway and the response to selection and associated evidence appears in the genes with the largest control coefficients (Eanes 2011). Again, we would expect that most enzymes have low control coefficients (Kacser and Burns 1981) and would not stand out when looking at molecular data. Dykhuizen et al. (Dykhuizen *et al.* 1987) first combined metabolic flux control theory with fitness measures (assayed as growth rate) to predict the control coefficients of two enzymes involved in lactose metabolism in *Escherichia coli,* β-galactosidase and β-galactoside permease. Using chemostats, with lactose as the limiting nutrient, they measured relative fitness of strains with known enzyme activities. Except for genetic variants with very low β-galactosidase, they were able to predict fitness of *E.coli* strains via flux control (most showed low control coefficients). Changes in β-galactoside permease, having a much larger control coefficient than β-galactosidase, resulted in large changes in fitness whereas even very large changes in β-galactosidase activity resulted in only small changes in fitness until activity levels were extremely low. Consequently, if the environment favored shifting flux in the pathway one might expect genetic variation in expression to appear at the permease step. With this association between fitness and metabolic flux control, we can connect flux control with molecular

signatures of selection within a network of interest. Given the central importance of metabolism, we could reasonably expect variation in glycolytic flux to be a target of selection for many traits, including starvation resistance, lifespan, fertility and fecundity. However, the presence or absence of flux control can be inferred in cases where the footprint of selection can be observed in molecular data (Eanes 2011). For example, when Flowers et al (2007) investigated patterns of adaptive evolution in glycolytic enzymes of *D.melanogaster* they found these signals clustered around the major branch point glucose-6-phosphate. While current high-throughput sequencing methods have made it relatively easy to scan the entire genomes of both model and non-model organisms for genetic variation (Stinchcombe and Hoekstra 2008; Fabian *et al.* 2012) such a wide view over networks as large and complex as metabolism would be expected to yield large numbers of non-interacting targets. We chose to focus on a particular portion of the metabolic pathway to find candidate genes that may have local control so as to generate fine-grain analyses.

With this focus on the glycolytic pathway, our lab has used data from high-throughput sequencing techniques and the population genetics to predict the enzymes that are under adaptive selection in a relatively recent evolutionary time scale in *Drosophila melanogaster*. This may identify loci that we suspect have significant flux control (Eanes 2011). The populations in our study were sampled along a latitudinal gradient on the eastern coast of North America, from the southern end of Florida to Maine, as well as from further inland in Sudbury, Ontario. This historically tropical species has adapted to temperate climates since its colonization in North America less than 200 years ago, and maintains stable local populations in the northern as well as the southern parts of its range (Reaume and Sokolowski 2006). Many of the genes in this pathway possess SNPs that vary in a latitude-dependent fashion (Sezgin *et al.* 2004) that posits a gradient in local tropical-temperate selection. Specifically, we used sequence data from the DGRP lines (Mackay *et al.* 2012) and calculated an allele effect on cis-acting whole-fly adult gene expression (Ayroles *et al.* 2009) for each SNP. We then used bulk pyrosequencing to estimate SNP allele frequencies across the 20 populations (Lavebratt and Sengul 2006; Doostzadeh *et al.* 2008). Results from a principal component analysis of the variance-covariance matrix of population SNP-associated expression variation across the north-south cline provided a multivariate description of patterns of expression change with latitude. It predicted a general down-regulation of enzymes at the entrance to the glycolytic pathway and an up-regulation of cofactor shuttles, *Gpdh* and *Gpo*-1, and well as on the

mitochondrial side of the glutamate-aspartate shuttle. The genes studied here, *Gpdh*, *Gpo-1*, and *HexA*, are significant players in that analysis.

Additionally, experiments in our lab (Talbert et al., in revision.) have shown that metabolic enzyme activity perturbation can have a significant effect on lifespan, oxidative stress resistance and starvation resistance. The experiments were performed with lines of *D. melanogaster* that compared 'full' activity against 'reduced' activity genotypes that were generated by P-element mutagenesis. For each gene, many mutant alleles were screened and 'full' activity and 'knockdown' or 'knockout' alleles were saved, assayed, and placed in identical genetic backgrounds using a controlled cross design(Merritt *et al.* 2006). Because dietary restriction is an important feature in life span extension (CHIPPINDALE *et al.* 1993; PARTRIDGE *et al.* 2005; TATAR 2011; PIPER *et al.* 2014), a diet effect was also tested in the lifespan experiments. Results show a significant effect of *Gpdh*, *HexA*, *Gdh*, and *Men* activity on lifespan for one or both diets. Reduction of *Idh* activity had no effect on lifespan, even though the expected dietary effect was still observed. *Mdh2* activity perturbation had a significant effect on low nutrient food, but not with high nutrient food. Also of note is that there was also no diet effect (wherein flies are expected to be longer lived on low nutrient food) (Partridge *et al* 1987; Chippindale *et al.* 1993), of *Gpdh* genotypes. The results of starvation resistance are all significant. The result of *HexA* hints at a possible trade-off between longevity and starvation resistance. It is of special interest that *Gpdh* showed a significant extension of life span in the genotype with lowered activity. *Gpdh* and *Gpo-1* represent the cytosolic and mitochondrial, respectively, sides of the glycerol-phosphate shuttle that is used to transfer NAD/NADH equivalents and thereby should play a direct role in the reduction-oxidation (redox) state of the cell (O'Brien and MacIntyre 1972).

This study investigates what effect, if any, perturbation of the metabolic pathway has in the context of adipokinetic hormone (AKH) signaling during starvation. AKH producing cells in the corpora cardiaca (CC) are required for starvation induced hyperactivity (Lee and Park 2004). Furthermore, metabolic state sensing has been linked to AKH release from the CC cells in response to starvation through AMPK (Braco *et al.* 2012). *HexA* uses ATP as a cofactor and is near the major branch point of glucose-6-phosphate (Flowers *et al.* 2007) and *Gpdh*, and *Gpo-1* are partners in the glycerophosphate shuttle and use NAD(P) and FAD as cofactors, respectively. Position in the overall pathway, the use of highly connected cofactors and predicted metabolic flux control makes

these genes attractive candidates to affect starvation response by control of metabolic flux and therefore metabolic state.

To test the effect of candidate genes in AKH signaling starvation response, I have used tissue specific RNAi knockdown using RNAi constructs from the Transgenic RNAi Project (TRiP) under the control of *GAL4* UAS and *GAL4* drivers expressed under the control of *Akh* (Lee and Park 2004) and *dILP2* (Rulifson *et al.* 2002) *Drosophila melanogaster* promoters. Insulin signaling is known to play a role in starvation resistance (Clancy *et al.* 2001; Ikeya *et al.* 2002; Broughton *et al.* 2005; Wang *et al.* 2008) although coupling of metabolic state with dILP release lies outside of the dILP producing cells (Geminard *et al.* 2009).

Methods and Materials

Drosophila Stocks: The following TRiP stocks were obtained from Bloomington Stock Center, Bloomington, IL.: *Gpdh* (CG9042) BDSC#51474, *Gpo-1* (CG8256)BDSC#55319, *HexA* (CG3001) BDSC#35155, *Men* (CG10120) BDSC#38256, *Mdh2* (CG7998) BDSC#36606, *Gdh* (CG5320) BDSC#51473, *Idh* (CG7176) BDSC#41708, *Pdk* (CG8808) BDSC#35142, *Cary2* control: *y*[1] *v*[1] ; P{*y*[+t7.7]=CaryP}attP2 (BDSC#36303), *Cary40* control: *y*[1]*v*[1];P{*y*[+t7.7]=CaryP}attP40 (BDSC#36304), *tub-GAL4* (BDSC#5138): y[1] w[*]; P{w[+mC]=tubP-GAL4}LL7/TM3, Sb[1] Ser[1]. Both the *dILP2-GAL4* and *Akh-GAL4* driver lines were derived from stocks with the driver cassette inserted in a P-element. The *Akh-Gal4* driver line was crossed with In(1)w[m4h]; Df(3L)Ly, sens[Ly-1]/TM3, ry[RK] Sb[1] Ser[1] P{ry[+t7.2]=Delta2-3}99B (BDSC#2030) and driver P-element retentions and excisions screened and placed in a w;CyO/6326;TM3, Sb/~ background. The *dILP2-GAL4* drier line was crossed with y[1] w[*]; CyO, H{w[+mC]=PDelta2-3}HoP2.1/PPO1[Bc] and driver P-element retentions and excisions screened and placed in a w;Cy/~;VT46 background . Akh-*GAL4* driver was derived from BDSC#25684. The *dILP2-GAL4* driver was derived from BDSC#52660 w[1118]; P{w[+mC]=Ilp3-GAL4.C}2/CyO . Long term stocks were maintained at 18ºC with 12:12 hours light:dark cycle. All crosses were performed at 25ºC. All flies were handled at room temperature with CO_2 anesthesia and held in incubators set to 12:12 hours light:dark cycle and 50-70% humidity. All food consisted of cornmeal-yeast flake-corn syrup diet stabilized with agar at concentrations per 100ml of 3.75g corn meal, 3.75g inactive yeast flake, 7.3g corn syrup solids,

1g agar, with final concentration of 2mg/ml of *p*-hydroxybenzoic acid methyl ester dissolved in ethanol (0.75% final concentration of ethanol). All vials and bottles were supplemented with active-dry yeast (Red Star Baker's Yeast) applied to the top of cooled food.

Crosses: Five *TRiP* males were crossed with 10 corresponding control line (*Cary2* or *Cary40*) 5-10 day old virgin females per vial to generate *TRiP*/*Cary* heterozygotes. Virgin females were collected twice daily and held at 18ºC for 5-7 days before mating. Fifty *TRiP*/*Cary* females were crossed with 25 *GAL4* males in each bottle. *TRiP* and control genotypes emerged together within each bottle and were separated by eye color. *GAL4* driver and excision control genoere crossed to *TRiP*/control females separately with two bottles per *GAL4* genotype (Supplemental Figs 1 & 2). Mating adults were transferred three times to new vials every two days or new bottles every 24 hours. '18ºC constant' cross bottles were transferred from 25ºC to 18ºC immediately after parental adults had been removed. Males were separated by bottle of origin (*GAL4* genotype) and eye color (*TRiP* genotype) and held at either 18ºC or 25ºC. All assays were performed on 5-7 day old males held in vials of 10-20 individuals (Supplemental Fig 3).

Starvation: Flies were held in standard vials with about ¼ of a rayon ball moistened with distilled water. Starvation vials were held in one 10x10 vial rack at 18ºC or 25ºC with 12:12 hours light:dark and 50-70% humidity. Dead flies were counted every 6 hours from start at room temperature. F2 males were placed in vials of 10 individuals each, 10 vials per *TRiP* genotype (*TRiP* or *Cary*) per bottle of origin for a total of 800 flies per experiment. F2 Males from replicate bottles of the same F1 cross parents were pooled and randomized by true bottle of origin across replicate vials. Genotypes and bottle of origin were randomized across an 8x10 grid of vials. Flies were transferred to new food vials at least 4 hours prior, or 18 hours for temperature shift experiments, to starvation start in the final randomization under CO_2 anesthesia. Flies were then transferred to starvation vials without anesthesia. Parametric Survival tests were performed in JMP 11 (SAS).

Locomotor activity: Locomotor activity was measured by a TriKinetics Drosophila Activity Monitor, by placing one fly in each 5x60mm glass tube with 1mm thick walls. Food or Distilled water and agar was placed at one end and capped with a 10mm long plastic cap. After placing a fly in the tube, the other end was capped with a 6mm cap with a small hole for air exchange. Fed flies were given standard diet to about 10mm and starved flies were given a 3% agar in distilled

water to about 10mm. Males flies 6-7 days old were used and had been acclimated to the light cycle while in vials of 20 individuals each at 18ºC. Vials of flies were transferred to 25ºC at least 18 hours prior to being transferred to locomotor activity tubes. Analysis of sleep and activity measures was performed in pySolo v1.0. Significant differences in variances between pairs were tested by calculating F. Means and standard deviations were exported and pairwise t-tests were for equal or unequal variances performed accordingly (Sokal and Rohlf 1995). Each experiment used 64 total flies: 32 fed and 32 starved, 8 flies from each of two bottles for each of the two *Akh-GAL4* versions. Flies were reared as described for starvation experiments but only individuals carrying the *TRiP* insert were used.

Individual fly data

Wet weight was taken on individual male flies 5-7 days old. Flies were transferred to -20ºC for one hour prior to weighing on a Cahn C32 microbalance. Flies were kept on ice except when being weighed. Analysis of Variance was performed in JMP 11 (SAS).

Results and Discussion

The general pattern observed is that there is no significant effect on starvation resistance of knockdown in IPCs (Fig 1 & Table 1). While statistical analyses considered each replicate bottle separately, the graphs combine data from the same genotypes coming from different bottles and allude to the fact that significant effects of the *GAL4* or *TRiP* elements are driven largely by one of the two replicate bottles. This is not an unreasonable expectation, given that the release of dILPs appears to be under the control of signals from the fat body (Geminard *et al.* 2009). The results of the *dILP2-GAL4* starvation experiments support Geminard et al's (2009) model of dILP signaling coupled to nutrient signaling outside of the dILP releasing cells. These also serve as a negative control to test the effects the *TRiP* inserts with a *GAL4* driver other than the *Akh-GAL4*.

The knockdown of any of the genes tested in AKH producing cells has no significant effect on starvation resistance (Fig 2 & Table 2) with the possible exception of *Gpo-1*. As with the *dILP2-GAL4* experiments, the statistical tests show whole-model significance, yet the *GAL4* and *TRiP* effects are driven by bottle and vial effects. The experiment with *Gpo-1* is quite interesting, but

the strength of the effect of the *TRiP* element alone raised concern. To address this in more detail, we replicated the starvation experiments at varied temperatures.

It is widely believed that the wild-type *Saccharomyces cerevisiae GAL4* used in *D. melanogaster* expression constructs, and under the control of the *dILP2* and *Akh* promoters here, are temperature sensitive. However, Mondal et al. (2007) have shown that this full length wild-type version of *GAL4* is not temperature sensitive. To rule out a temperature dependent effect of *Gpo-1 TRiP* RNAi, we replicated starvation experiments at different temperature regimes (Fig 3 & Table 3). Again, we saw a significant *Gpo-1 TRiP* effect even at 18ºC, the presumed non-permissive temperature. For comparison, I replicated starvation experiments for *HexA* and *Gpdh* using the *Akh-GAL4* driver at different temperatures (Fig 4 & Table 4). Again, for all experiments under these temperature regimens, the convincing result was for the *Gpo-1 TRiP* insert and the other significant effects seemed to be driven by bottle of origin effects. A general pattern of increased differences in survival by genotype did arise as temperature decreased.

The increased differences of survival by genotype is correlated by the temperature effect on body size for *Drosophila* (ALPATOV 1930; Imai 1933 and many others, also see Partridge *et al.* 1994). Wet weight, and not dry weight, was taken to facilitate metabolite measures with the same flies. Wet weights correlate well with dry weights (CHIPPINDALE *et al.* 1998; OKAMOTO *et al.* 2009) and leave the option of taking dry weight or metabolite measures at a later time. Weights by genotype reflect similar patterns seen in the starvation experiments (Fig 5 & Table 5). Notably, *Gpo-1 TRiP* flies are lighter regardless of the *GAL4-Akh* genotype. This is likely what is driving the starvation sensitivity differences with regard to the *Gpo-1 TRiP* experiments. A more detailed examination of temperature effect shows that temperature is the dominant effect, far above even the *Gpo-1 TRiP* genotype effect (Fig 6 & Table 6).

Analysis of Activity Index during the first full day of *Gpdh* flies, hours 11-34 of starvation (Fig 7), indicated no difference between starved flies by RNAi activity or between starved and fed flies of similar RNAi activity (Between *Akh-GAL4* Retention*:* $t = 0.734$, $d.f = 30$, $p = 0.4686$; Between *Akh-GAL4* Excision: $t = 0.91583$, $d.f. = 30$, $p = 0.3671$). There is also no significant difference between genotypes for fed or starved flies (Between fed: $t = 0.2908$, $d.f = 22.62$, $p = 0.7739$; Between starved: $t = 0.9618$, $d.f. = 19.29$, $p = 0.3482$). The lack of a difference between genotypes

within starved or fed states supports the lack of an effect of RNAi, but the lack of a difference between fed and starved flies suggests a lack of power due to small sample size. Inspection of the graphs shows a hint of an activity increase at 0900 as well as reduced sleep periods. More replicates would likely find a clearer difference between fed and starved flies, but this also leads to an issue with drawing conclusions about genotype differences.

Analysis of Activity Index during the first full day of *Gpo-1* flies, hours 11-34 of starvation (Fig 8), indicate a significant difference by genotype for both fed and starved (Between fed: $t = 3.29$, d.f. = 17.00, $p = 0.0043$; Between starved: $t = 3.11$, d.f. = 23.73, $p = 0.0049$). There was also a significant difference in Activity Index between fed and starved flies with active (Between *Akh-GAL4* Retention*:* $t = 2.16$, d.f =30, $p = 0.0388$; Between *Akh-GAL4* Excision: $t = 6.42$, d.f. = 21.02, $p = <0.0001$). This experiment suffers from the same problem of power as with the *Gpdh* experiment. More replicates would need to draw firm conclusions about genotypic differences.

Gpo-1 TRiP experiments suggest an effect on starvation resistance, though the cause is unclear. The large effect of the presence of the *Gpo-1 TRiP* construct alone, regardless of the presence or absence of the *Akh-GAL4* driver, would suggest a position effect. However, the *Cary40* control line contains an insert at the same genomic position and using the same source vector but without the *vermillion* gene or the *Gpo-1* specific RNAi sequence. Also, the *Gpdh*, *Gdh*, *Idh*, and *Men* TRiP lines used here also use the same *attp40* insertion site without such obvious effects. These comparisons rule out the position effect and a general presence of *TRiP* insert effect and point to an effect specific to the *Gpo-1 TRiP* construct.

It is possible that there may be low to moderate expression of *Gpo-1* dsDNA, perhaps ubiquitously. Viability tests of the TRiP constructs with a ubiquitous *tubulin-GAL4* driver suggest that *Gpo-1 TRiP* works, as this combination rare yields adults (Supplemental Table 1, $p < 0.0001$). Whole-fly enzyme activity of GPO-1 could be determined and tested between RNAi and control flies, as well as *Gpo-1* RNA content in the whole body by qPCR. Determination of *Gpo-1* RNAi efficiency in the AKH producing cells will be challenging due to the small number and location of the cells. Expression of the RNAi throughout the body is attractive because this would mean that there opens the possibility of a major effect of *Gpo-1* with the experiments here, but outside the context of AKH signaling, given the effect on wet weight and starvation resistance. We have not yet

developed a P-element excision series with *Gpo-1*, as has been done with the other genes tested here. The results here suggest that doing so could prove fruitful, particularly in testing the general effect of perturbation of the glycerophosphate shuttle, and not just *Gpdh*, the partner to *Gpo-1* in this shuttle.

It is also possible that the *Akh-GAL4* driver is not working properly or RNAi knockdown with the *TRiP* lines is inefficient in this context. Metabolic state sensing could be robust to metabolic flux perturbations. The model of metabolic state coupled to AKH signaling through AMPK within AKH producing cells (Lee and Park 2004; Braco *et al.* 2012) could be complicated if RNAi knockdown in this study is inefficient. That is, unless the sensing system is robust to general perturbation and this study has missed a specific effector. Remote sensing, similar to Insulin/Insulin-like signaling (Colombani *et al.* 2003; Geminard *et al.* 2009) is a possibility and preliminary experiments with knockdown and overexpression in the fat body would be in order.

Using starvation as an assay risks involving the Insulin/Insulin-like signaling pathway as well as possibly AKH signaling. Significant effects of metabolic enzyme perturbation in the fat body on starvation resistance are of great interest whether AKH or dILP signaling, or both, is involved. Release of AKH into the hemolymph in this case may be a clearer assay with respect to starvation induced AKH signaling as care should be taken to make sure that effects of Insulin/Insulin-like signaling and AKH signaling can be distinguished from each other.

Chapter 3 Figures and Tables

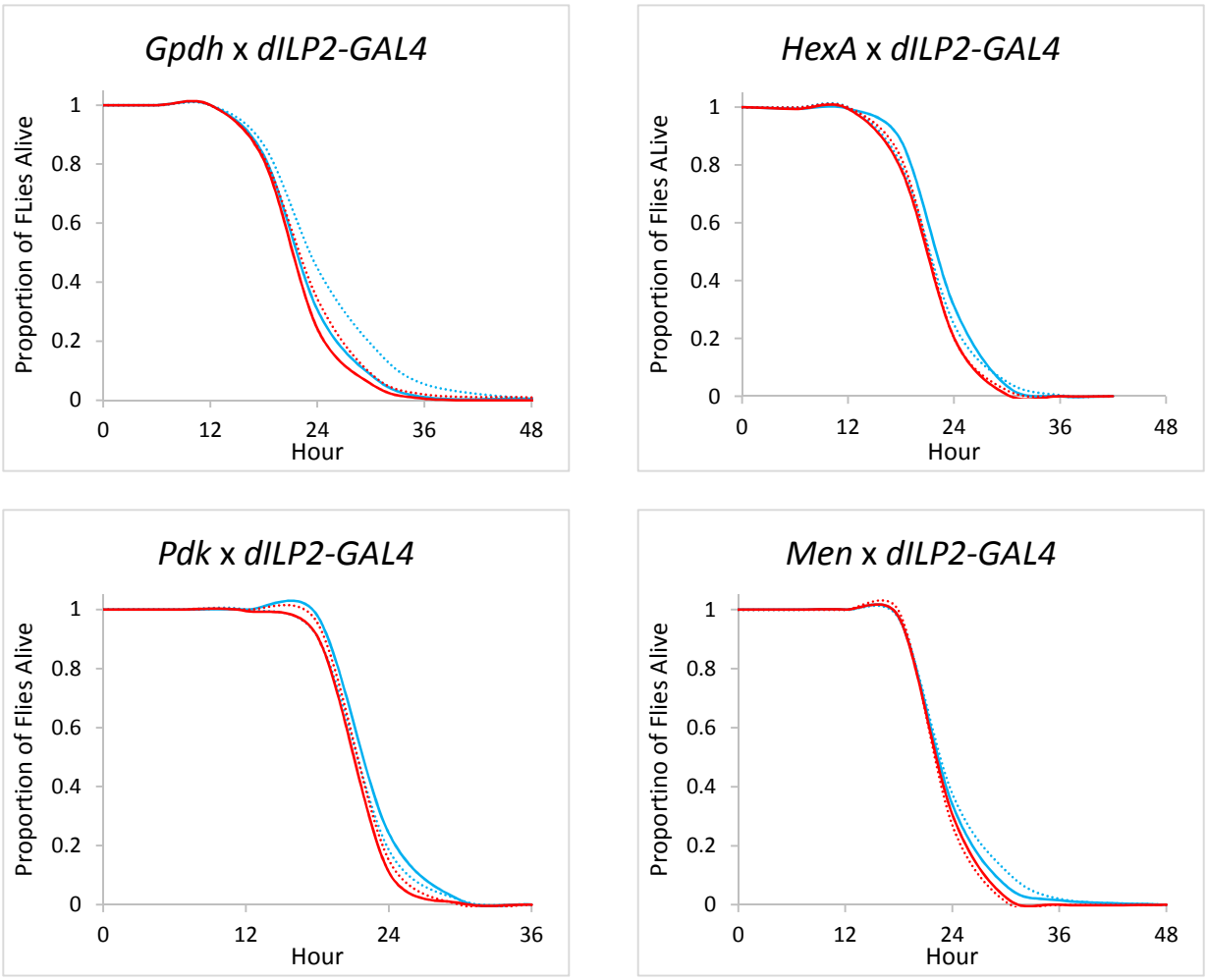
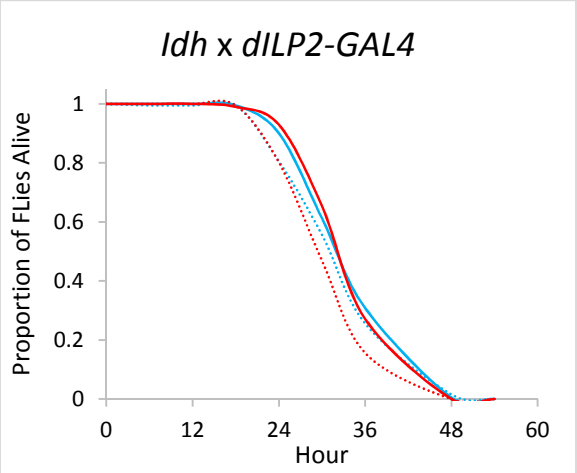
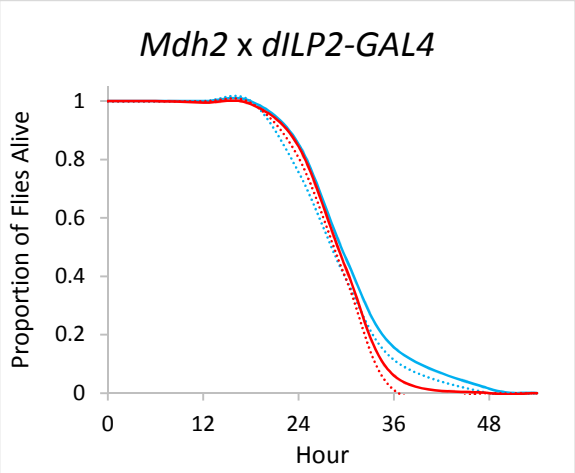


Fig 1. *TRiP* x *dILP2-GAL4* starvation experiments. All experiments were performed at 25ºC constant temperature. Red lines are flies with the *dILP2-GAL4* driver, blue lines are flies with that element excised. Solid lines are flies with the *TRiP* insert, dotted lines are flies with the *Cary* control insert.





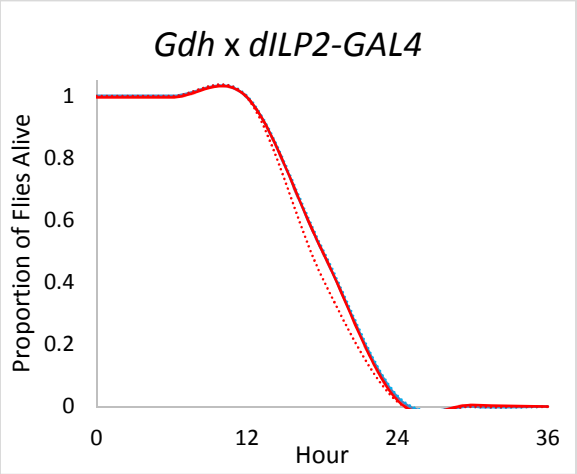


Fig 1 continued. *TRiP* x *dILP2-GAL4* starvation experiments. All experiments were performed at 25ºC constant temperature. Red lines are flies with the *dILP2-GAL4* driver, blue lines are flies with that element excised. Solid lines are flies with the *TRiP* insert, dotted lines are flies with the *Cary* control insert.

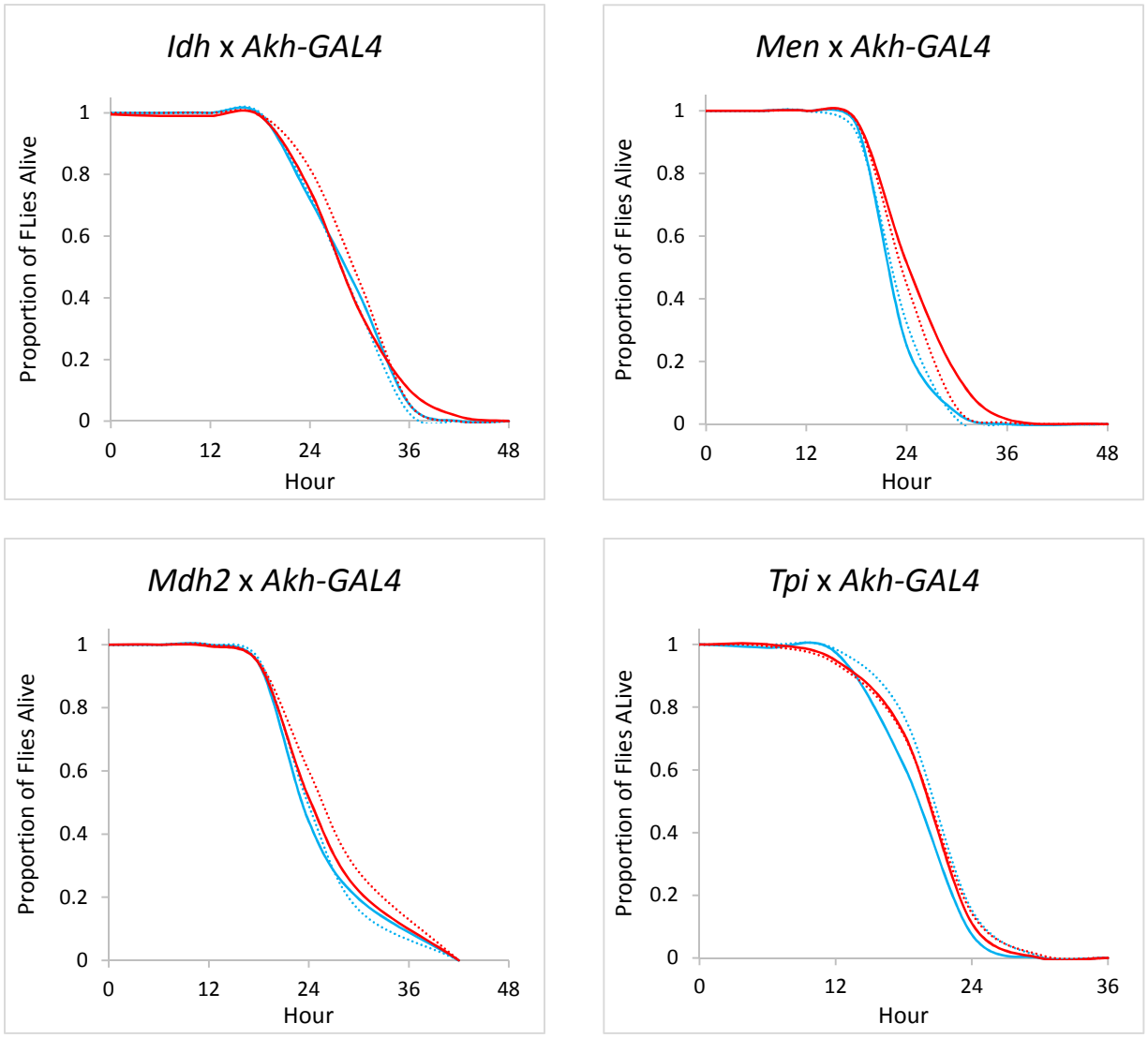


Fig 2. *TRiP* x *Akh-GAL4*starvation experiments. All experiments were performed at 25ºC constant. Red lines are flies with the *Akh-GAL4* driver, blue lines are flies with that element excised. Solid lines are flies with the *TRiP* insert, dotted lines are flies with the *Cary* control insert.

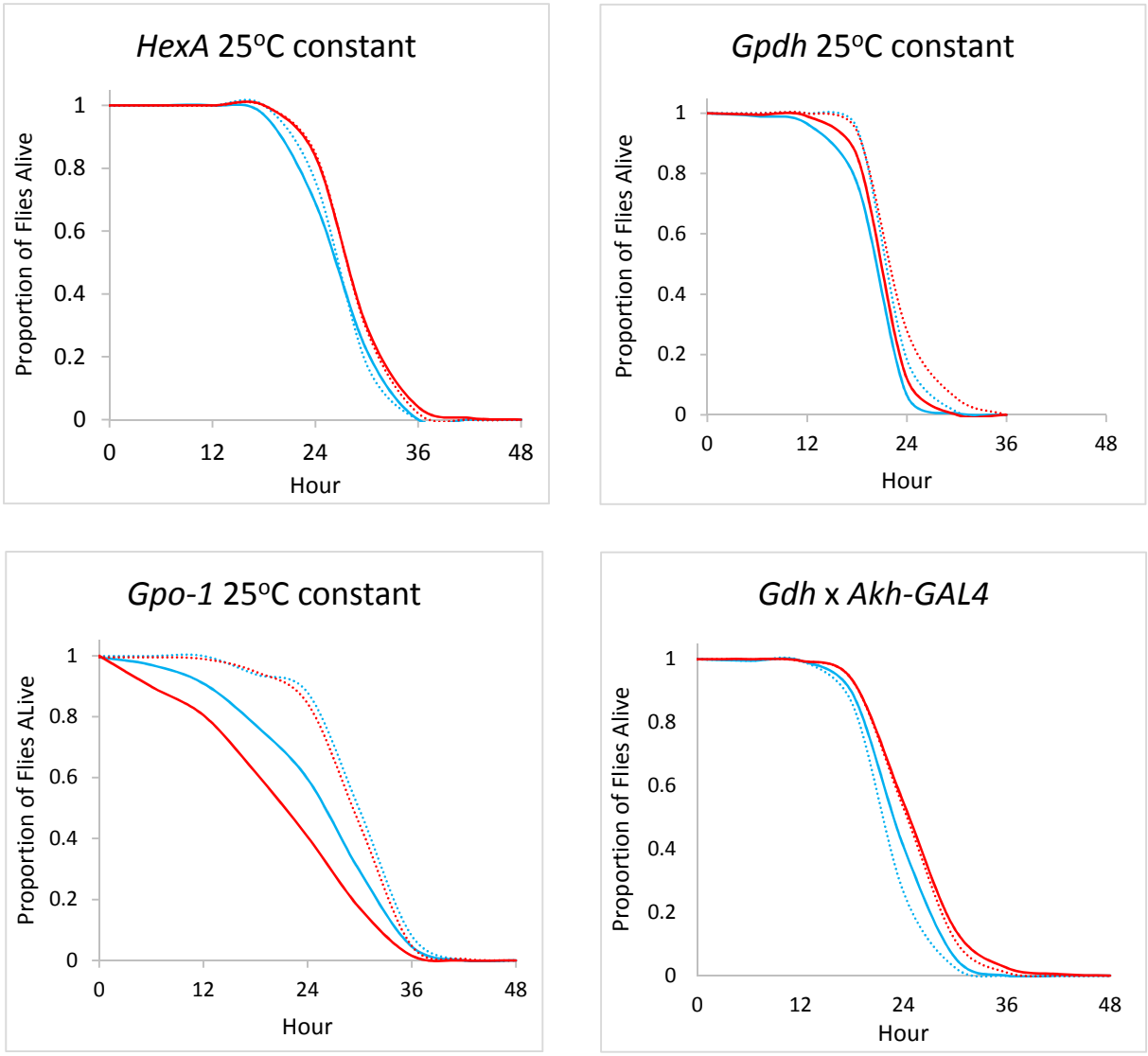


Fig 2 continued. *TRiP* x *Akh-GAL4*starvation experiments. All experiments were performed at 25°C constant. Red lines are flies with the *Akh-GAL4* driver, blue lines are flies with that element excised. Solid lines are flies with the *TRiP* insert, dotted lines are flies with the *Cary* control insert.

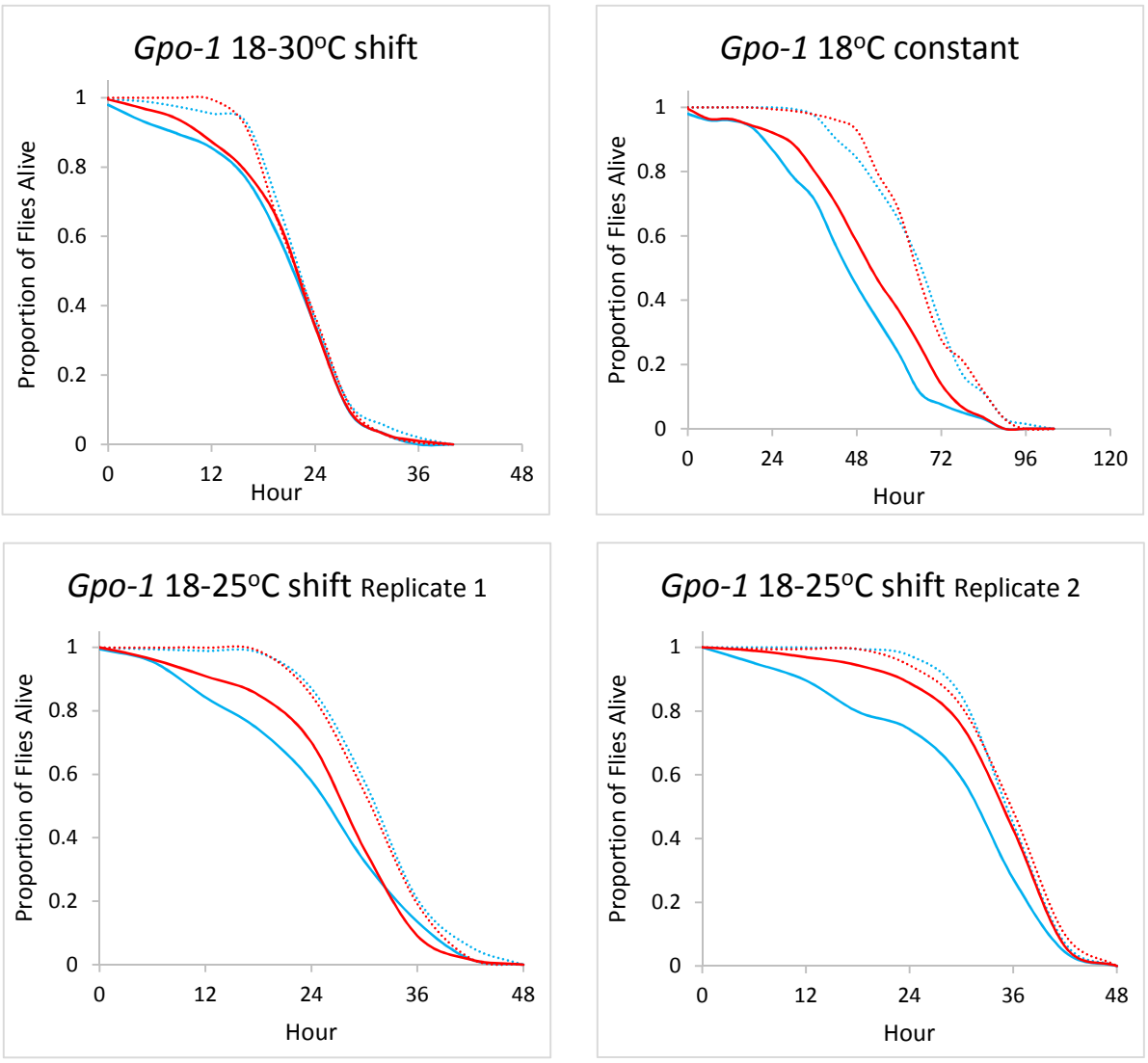


Fig 3. Results of starvation experiments using *Gpo-1 TRiP* with *Akh-GAL4*.Red lines are flies with the *Akh-GAL4* driver, blue lines are flies with that element excised. Solid lines are flies with the *TRiP* insert, dotted lines are flies with the *Cary* control insert. Two replicates are shown for the 18-25°C temperature shift.

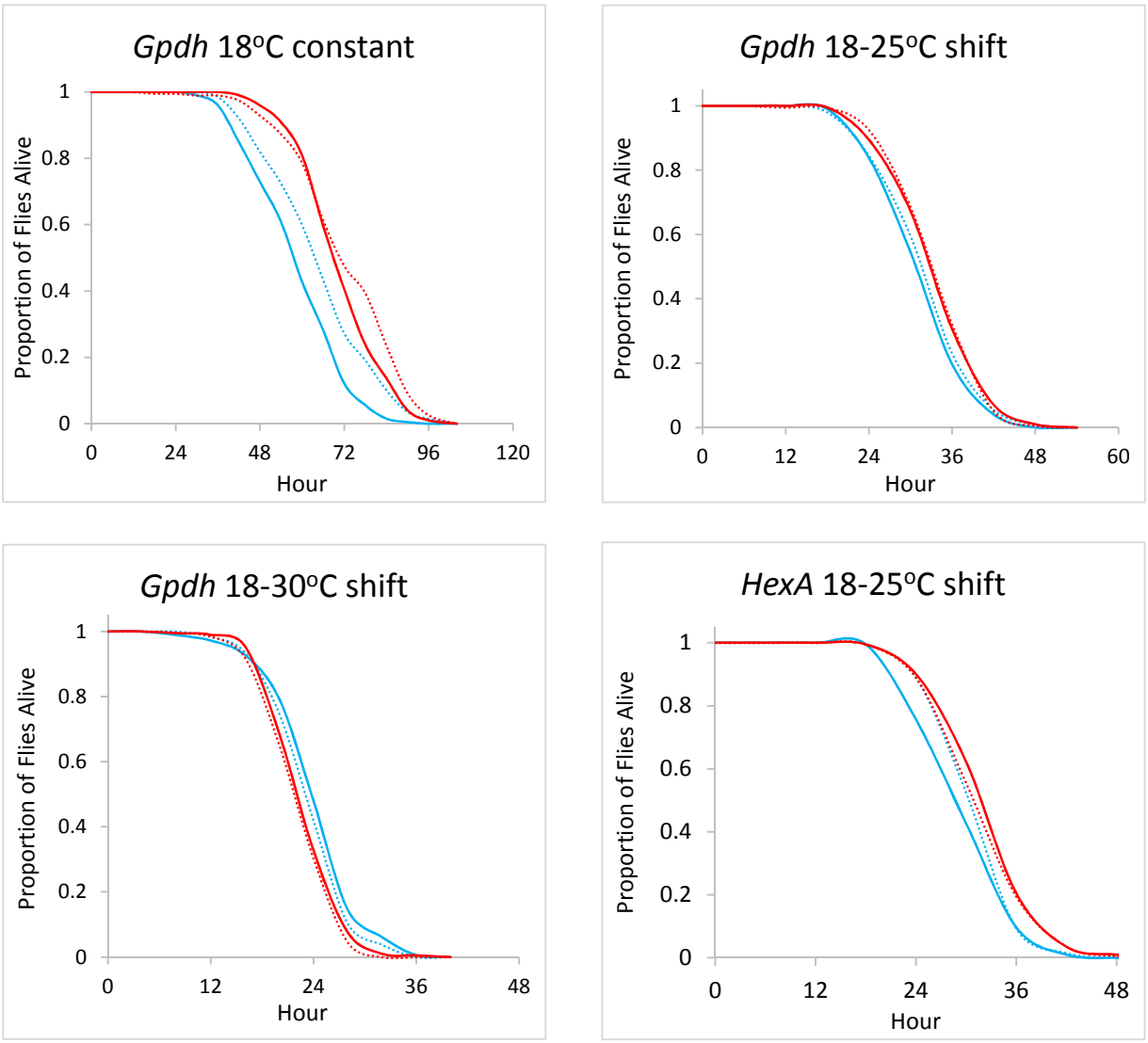


Fig 4. Results of starvation experiments using *Gpdh TRiP* and *HexA TRiP* with *Akh-GAL4*.Red lines are flies with the *Akh-GAL4* driver, blue lines are flies with that element excised. Solid lines are flies with the *TRiP* insert, dotted lines are flies with the *Cary* control insert.

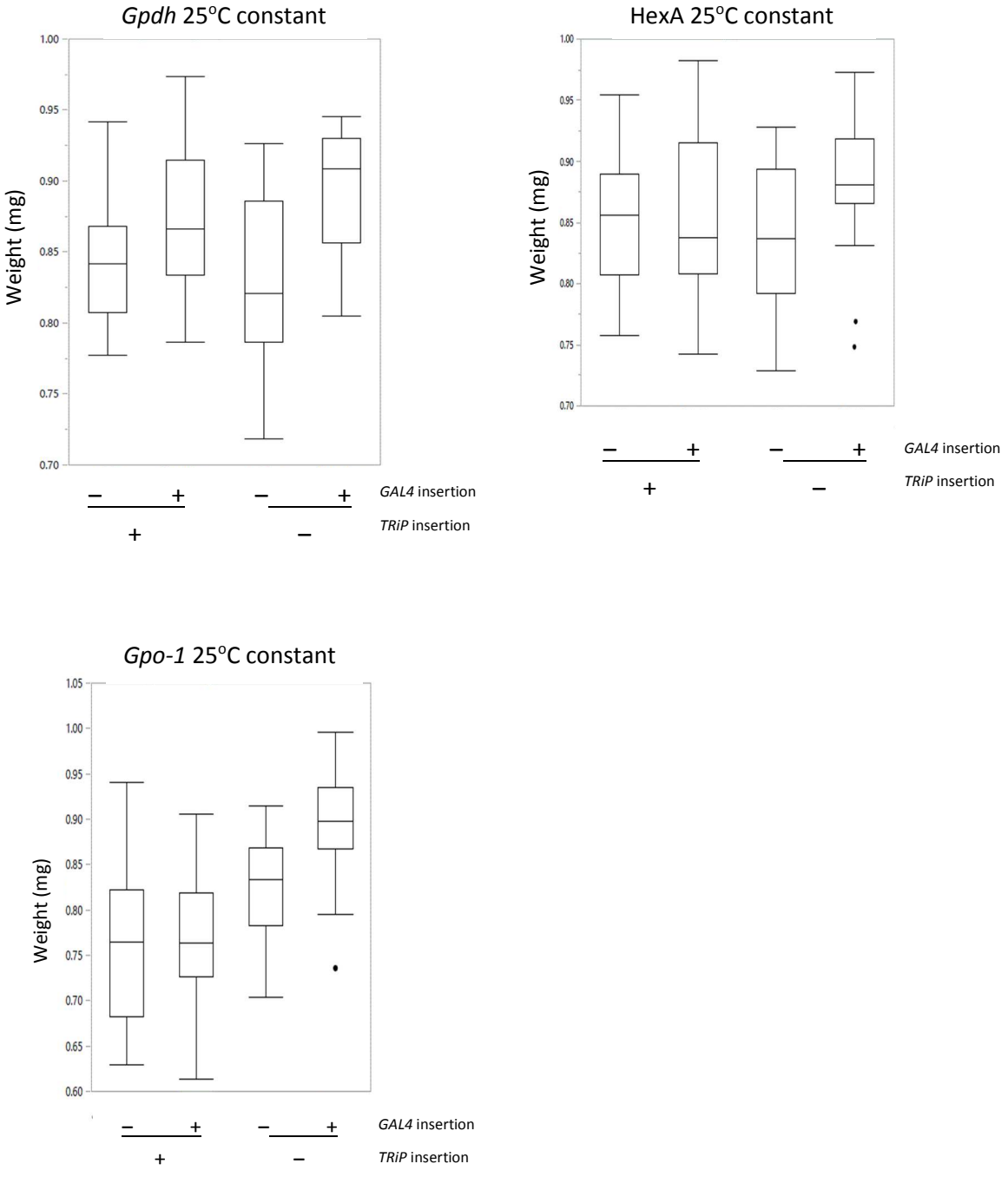


Fig 5. Weights of individual flies by genotype for *TRiP* x *AKh-GAL4* crosses. Presence and absence of insertions labeled as (+) and (–), respectively.

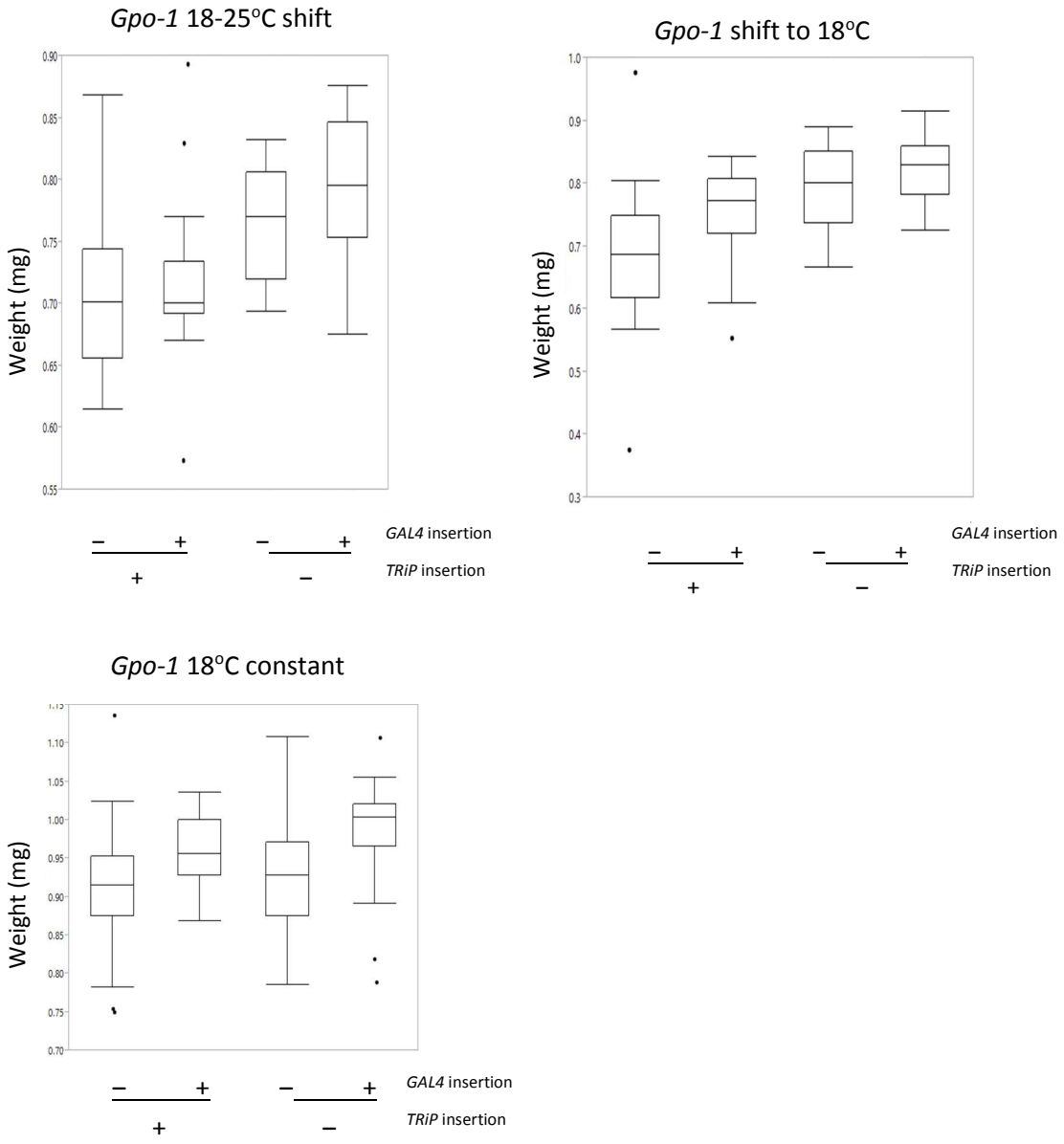


Fig 6. Temperature effect of weights of individual flies by genotype for *Gpo-1 TRiP* x *Akh-GAL4* crosses. Presence and absence of insertions labeled as (+) and (–), respectively. Flies were reared and treated similarly to flies used in starvation assays. Rearing conditions and temperature regimes were replicated so that flies weighed were representative of flies used in starvation experiments at the start of starvation.

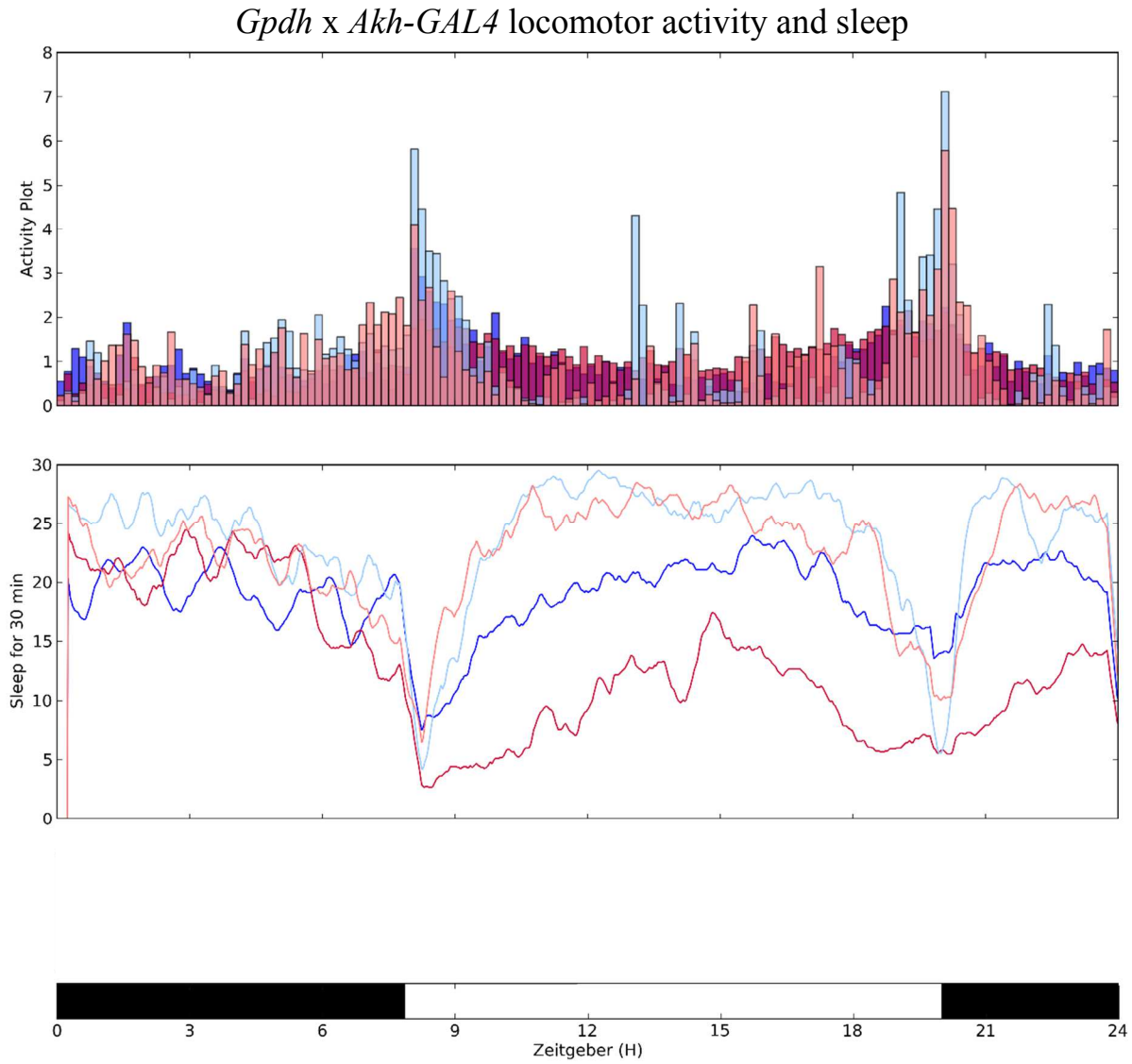


Fig 7. Locomotor activity and sleep analysis of *Gpdh* RNAi flies. All flies carry the *TRiP* insert. Blue represents flies carrying the *Akh-GAL4* insert, red carrying the *GAL4* excision chromosome. Dark colors represent starved flies, light colors fed flies. Zeitgeber depicts the light status (black=off, white=on) and hour of day (24h format). Flies were transferred to locomotor activity assay vials at about 14:00 on the day previous to data shown.

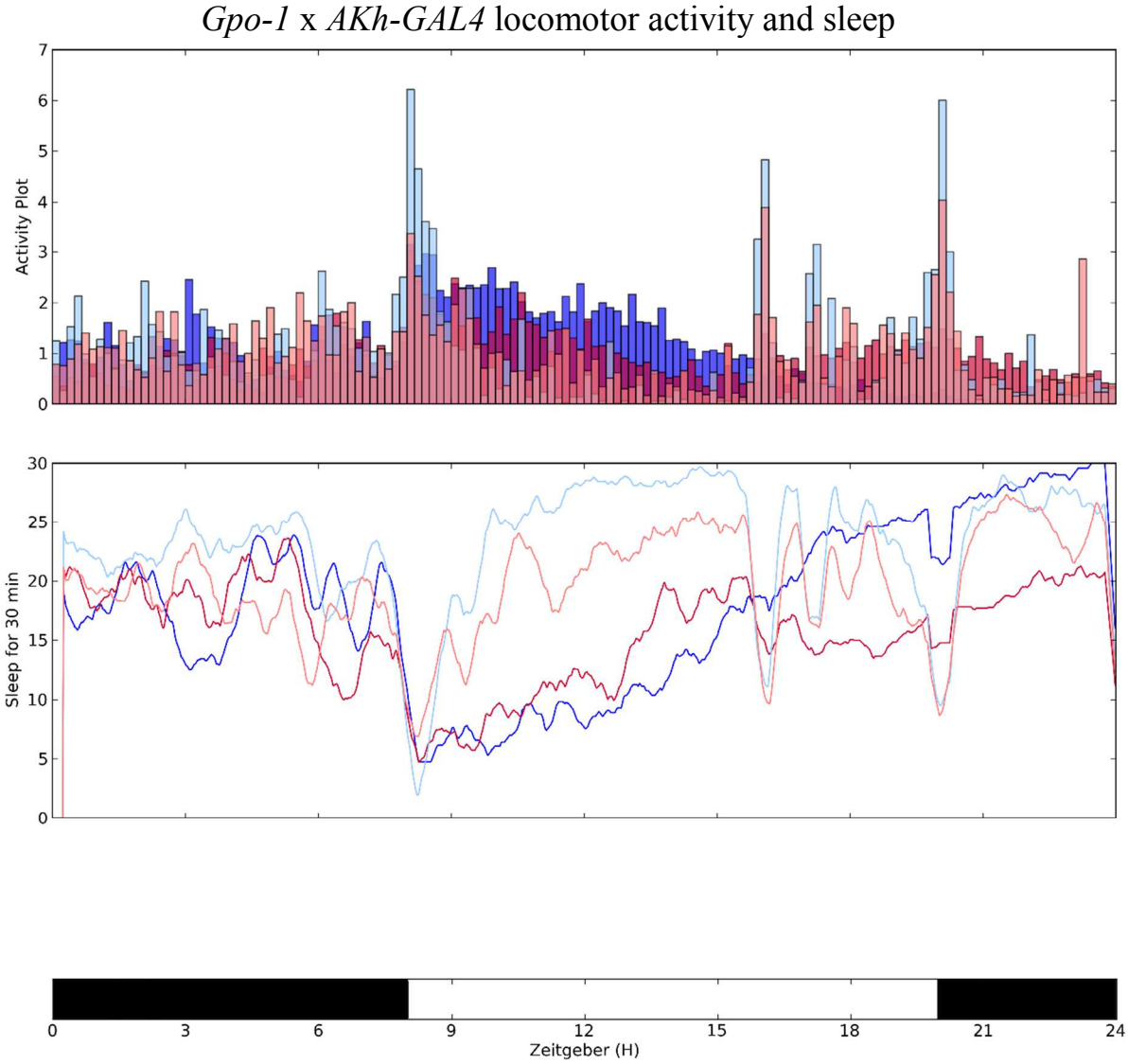


Fig 8. Locomotor activity and sleep analysis of *Gpo-1* RNAi flies. All flies carry the *TRiP* insert. Blue represents flies carrying the *Akh-GAL4* insert, red carrying the *GAL4* excision chromosome. Dark colors represent starved flies, light colors fed flies. Zeitgeber depicts the light status (black=off, white=on) and hour (24h format). Flies were transferred to locomotor activity assay vials at about 14:00 on the day previous to data shown.

Gpdh x *dILP2-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	1.087	0.2972
Bottle[Gal4]	2	2	3.599	0.1654
TRiP[Gal4,Bottle]	4	4	1.940	0.7469
Vial[Bottle,Gal4,TRiP]	72	72	8.444	1.0000
whole model		79	15.141	1.0000

Pdk x *dILP2-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	24.278	*<.0001*
Bottle[Gal4]	2	2	62.719	*<.0001*
TRiP[Gal4,Bottle]	4	4	37.638	*<.0001*
Vial[Bottle,Gal4,TRiP]	72	72	168.257	*<.0001*
whole model		79	262.556	*<.0001*

Idh x *dILP2-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	9.793	*0.0018*
Bottle[Gal4]	2	2	21.240	*<.0001*
TRiP[Gal4,Bottle]	4	4	24.379	*<.0001*
Vial[Bottle,Gal4,TRiP]	71	71	109.930	*0.0021*
whole model		78	147.439	*<.0001*

Gdh x *dILP2-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	6.366	*0.0116*
Bottle[Gal4]	2	2	9.327	*0.0094*
TRiP[Gal4,Bottle]	4	4	12.827	*0.0122*
Vial[Bottle,Gal4,TRiP]	72	72	123.735	*0.0001*
whole model		79	140.803	*<.0001*

HexA x *dILP2-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	19.074	*<.0001*
Bottle[Gal4]	2	2	14.678	*0.0006*
TRiP[Gal4,Bottle]	4	4	13.217	*0.0103*
Vial[Bottle,Gal4,TRiP]	70	70	126.545	*<.0001*
whole model		77	164.312	*<.0001*

Men x *dILP2-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	43.978	*<.0001*
Bottle[Gal4]	2	2	66.397	*<.0001*
TRiP[Gal4,Bottle]	4	4	9.256	0.055
Vial[Bottle,Gal4,TRiP]	72	72	138.328	*<.0001*
whole model		79	243.590	*<.0001*

Mdh2 x *dILP2-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	18.374	*<.0001*
Bottle[Gal4]	2	2	17.611	*0.0001*
TRiP[Gal4,Bottle]	4	4	16.813	*0.0021*
Vial[Bottle,Gal4,TRiP]	72	72	76.111	0.3477
whole model		79	122.559	*0.0012*

Table 1. Results of Parametric Survival tests for *TRiP* x *dILP2* starvation experiments. All experiments using the *dILP2-GAL4* driver were performed at a constant temperature of 25^{o}C.

Idh x *Akh-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	5.999	0.0143
Bottle[Gal4]	2	2	10.638	0.0049
TRiP[Gal4,Bottle]	4	4	10.329	0.0352
Vial[Bottle,Gal4,TRiP]	72	72	148.085	<.0001
whole model		79	171.342	<.0001

Mdh2 x *Akh-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	9.133	0.0025
Bottle[Gal4]	2	2	9.957	0.0069
TRiP[Gal4,Bottle]	4	4	8.535	0.0738
Vial[Bottle,Gal4,TRiP]	72	72	150.107	<.0001
whole model		79	168.389	<.0001

HexA x *Akh-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	33.1155	<.0001
Bottle[Gal4]	2	2	4.8513	0.0884
TRiP[Gal4,Bottle]	4	4	24.3126	<.0001
Vial[Bottle,Gal4,TRiP]	72	72	129.3311	<.0001
whole model		79	180.9590	<.0001

Gpo-1 x *Akh-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	11.3156	0.0008
Bottle[Gal4]	2	2	11.1605	0.0038
TRiP[Gal4,Bottle]	4	4	64.3339	<.0001
Vial[Bottle,Gal4,TRiP]	64	64	57.2158	0.7132
whole model		71	122.5475	0.0001

Men x Akh-GAL4

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	62.960	<.0001
Bottle[Gal4]	2	2	1.209	0.5464
TRiP[Gal4,Bottle]	4	4	33.141	<.0001
Vial[Bottle,Gal4,TRiP]	72	72	142.498	<.0001
whole model		79	238.876	<.0001

Tpi x Akh-GAL4

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	0.788	0.3747
Bottle[Gal4]	2	2	40.243	<.0001
TRiP[Gal4,Bottle]	4	4	32.434	<.0001
Vial[Bottle,Gal4,TRiP]	72	72	130.817	<.0001
whole model		79	188.083	<.0001

Gpdh x *Akh-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	18.79	<.0001
Bottle[Gal4]	2	2	4.80	0.0907
TRiP[Gal4,Bottle]	4	4	78.80	<.0001
Vial[Bottle,Gal4,TRiP]	72	72	173.15	<.0001
whole model		79	263.38	<.0001

Gdh x *Akh-GAL4*

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	72.248	<.0001
Bottle[Gal4]	2	2	77.568	<.0001
TRiP[Gal4,Bottle]	4	4	15.605	0.0036
Vial[Bottle,Gal4,TRiP]	72	72	127.576	<.0001
whole model	79		259.590	<.0001

Table 2. Results of Parametric Survival tests for 25^{o}C constant temperature starvation experiments with *TRiP* x *Akh-GAL4*.

Gpo-1 x *Akh-GAL4* 18-30ºC shift

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	0.112	0.7383
Bottle[Gal4]	2	2	2.305	0.3158
TRiP[Gal4,Bottle]	4	4	5.953	0.2027
Vial[Bottle,Gal4,TRiP]	72	72	91.444	0.0608
whole model		79	98.555	0.0674

Gpo-1 x *Akh-GAL4* 18ºC constant

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	6.512	0.0107
Bottle[Gal4]	2	2	12.382	0.002
TRiP[Gal4,Bottle]	4	4	137.349	<.0001
Vial[Bottle,Gal4,TRiP]	69	69	113.267	0.0006
whole model		76	234.715	<.0001

Gpo-1 x *Akh-GAL4*
Rep1 18-25ºC shift

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	8.004	0.0107
Bottle[Gal4]	2	2	0.923	0.002
TRiP[Gal4,Bottle]	4	4	23.616	<.0001
Vial[Bottle,Gal4,TRiP]	72	72	54.559	0.0006
whole model		79	78.405	<.0001

Gpo-1 x *Akh-GAL4*
Rep2 18-25ºC shift

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	0.295	0.5869
Bottle[Gal4]	6	6	4.817	0.5675
TRiP[Gal4,Bottle]	8	8	39.641	<.0001
Vial[Bottle,Gal4,TRiP]	64	64	44.296	0.9713
whole model		79	83.068	0.3553

Table 3. Results of Parametric Survival tests for *Gpo-1* x *Akh-GAL4* temperature regime experiments.

Gpdh x *Akh-GAL4* 18°C constant

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	106.725	<.0001
Bottle[Gal4]	2	2	30.528	<.0001
TRiP[Gal4,Bottle]	4	4	48.838	<.0001
Vial[Bottle,Gal4,TRiP]	72	72	101.501	0.0126
whole model		79	235.382	<.0001

Gpdh x *Akh-GAL4* 18-25°C shift

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	14.152	0.0002
Bottle[Gal4]	2	2	1.357	0.5073
TRiP[Gal4,Bottle]	4	4	13.245	0.0101
Vial[Bottle,Gal4,TRiP]	72	72	106.667	0.005
whole model		79	131.122	0.0002

Gpdh x *Akh-GAL4* 18-30°C shift

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	30.367	<.0001
Bottle[Gal4]	2	2	10.507	0.0052
TRiP[Gal4,Bottle]	4	4	7.028	0.1344
Vial[Gal4,Bottle,TRiP]	72	72	105.433	0.0063
whole model		79	148.188	<.0001

HexA x *Akh-GAL4* 18-25°C shift

Source	Nparm	d.f.	Two-tailed χ^2	Prob > χ^2
Gal4	1	1	39.349	<.0001
Bottle[Gal4]	2	2	7.735	0.0209
TRiP[Gal4,Bottle]	4	4	4.421	0.352
Vial[Bottle,Gal4,TRiP]	72	72	152.603	<.0001
whole model		79	190.294	<.0001

Table 4. Results of Parametric Survival tests for *Gpdh* x *Akh-GAL4* and *Gpdh* x *Akh-GAL4* temperature regime experiments.

Gpdh x *Akh-GAL4*

Source	Nparm	d.f	Sum of Squares	F Ratio	Prob > F
GAL4	1	1	0.050467	21.501	*<.0001*
Bottle[GAL4]	2	2	0.001341	0.286	0.7523
TRiP[GAL4,Bottle]	4	4	0.036131	3.848	*0.0063*
whole model		7	0.087939	5.352	*<.0001*

HexA x *Akh-GAL4*

Source	Nparm	d.f	Sum of Squares	F Ratio	Prob > F
GAL4	1	1	0.011517	3.638	0.0597
Bottle[GAL4]	2	2	0.015066	2.380	0.0985
TRiP[GAL4,Bottle]	4	4	0.014108	1.114	0.3551
whole model		7	0.040691	1.836	0.0902

Gpo-1 x *Akh-GAL4*

Source	Nparm	d.f	Sum of Squares	F Ratio	Prob > F
GAL4	1	1	0.041003	8.823	*0.0038*
Bottle[GAL4]	2	2	0.000908	0.098	0.9071
TRiP[GAL4,Bottle]	4	4	0.277933	14.951	*<.0001*
whole model		7	0.318013	9.776	*<.0001*

Table 5. Results of Analysis of Variance of *TRiP* and *Akh-GAL4* effects on weight. Flies were reared and held at 25^{o}C constant temperature.

Gpo-1 x *Akh-GAL4* 18-25ºC shift

Source	Nparm	d.f.	Sum of Squares	F Ratio	Prob > F
GAL4	1	1	0.009445	3.173	0.0783
Bottle[GAL4]	2	2	0.005727	0.962	0.3861
TRiP[GAL4,Bottle]	4	4	0.139112	11.683	*<.0001*
whole model		7	0.154284	7.404	*<.0001*

Gpo-1 x *Akh-GAL4* 18ºC shift

Source	Nparm	d.f.	Sum of Squares	F Ratio	Prob > F
GAL4	1	1	0.055014	10.422	*0.0018*
Bottle[GAL4]	2	2	0.066365	6.286	*0.0028*
TRiP[GAL4,Bottle]	4	4	0.216436	10.251	*<.0001*
whole model		7	0.328891	8.901	*<.0001*

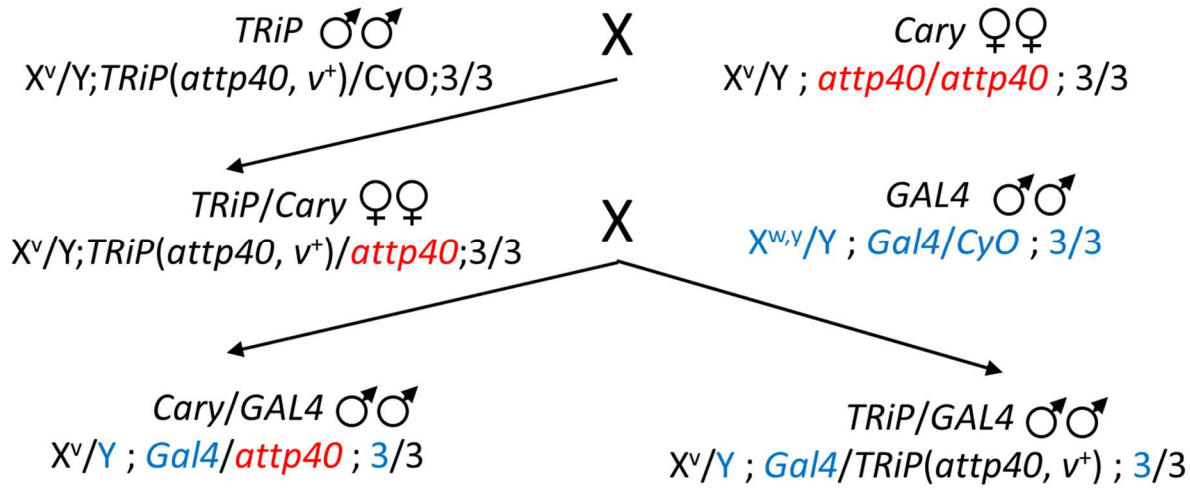
Gpo-1 x *Akh-GAL4* 18ºC constant

Source	Nparm	d.f.	Sum of Squares	F Ratio	Prob > F
GAL4	1	1	0.111505	10.134	*0.002*
Bottle[GAL4]	2	2	0.122663	5.574	*0.0053*
TRiP[GAL4,Bottle]	4	4	0.087998	2.000	0.1015
whole model		7	0.322166	4.183	*0.0005*

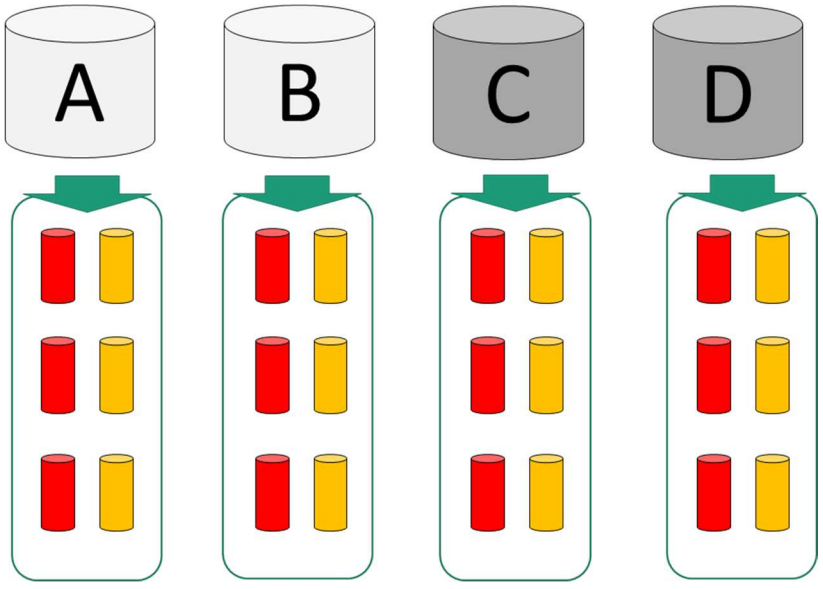
Gpo-1 x *Akh-GAL4* Temperature effect

Source	Nparm	d.f.	Sum of Squares	F Ratio	Prob > F
Temp regimen	2	2	1.50586	107.359	*<.0001*
GAL4[Temp regimen]	3	3	0.20444	9.717	*<.0001*
Bottle[Temp regimen,GAL4]	6	6	0.18992	4.513	*0.0002*
TRiP[GAL4,Temp regimen,Bottle]	12	12	0.57539	6.837	*<.0001*
whole model		23	2.44585	15.163	*<.0001*

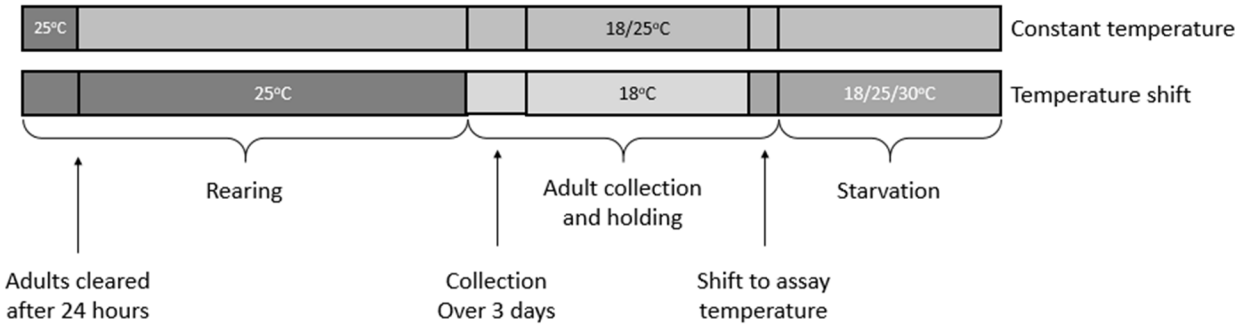
Table 6. Results of Analysis of Variance of *Gpo-1* x *Akh-GAL4* effects on individual fly weight at different temperature regimes. Data from the three temperature regimens shown here were combined for the Temperature effect analysis.



Supplementary Fig 1. Cross design to generate experimental flies. This set of crosses is an example of the design used. Specific crosses will differ from this but the overall design is the same. *TRiP* constructs inserted at *attP2* are on the third chromosome. The *dILP2-GAL4* construct is inserted in the second chromosome, as depicted here, and the *Akh-GAL4* construct is inserted in the third chromosome.



Supplementary Fig 2. Collection design of experimental flies. For a given experiment, two replicate bottles of *TRiP*/*Cary* females are crossed with *GAL4* retention males (bottles A & B) and *GAL4* excision males (bottles C & D). Males from each bottle are collected every 24 hours and separated into vials according to the bottle of origin and the *TRiP* genotype as determined by eye color (*TRiP* v^+ or *Cary v*). Males are held in vials of 10-20 individuals for 5-7 days prior to the start of the experiment.



Supplemental Fig 3. Temperature regime for Constant and Temperature Shift experiments. For all experiments, parents of experimental flies were allowed to lay eggs at 25ºC for 24 hours. For all temperature shift experiments, flies were held at 18ºC from collection until shift to assay temperature. Temperature shift occurred 18-24 hours prior to transfer to starvation vials.

					t-test between TRiP and Cary p-values	
Target gene	**TRiP;*tub-GAL4***	**Cary;*tub-GAL4***	**TRiP;TM3**	**Cary;TM3**	***tub-GAL4***	**TM3**
HexA	*2*	56	34	52	*0.0081*	0.3519
Gpdh	64	77	76	64	0.4815	0.5133
Gpo-1	*6*	140	97	140	*<0.0001*	0.0551
Pdk	42	47	38	39	0.6730	0.9337
Men	129	90	114	97	0.2407	0.5982
Idh	124	121	74	58	0.9391	0.5839
Mdh2	*0*	80	71	58	*0.0005*	0.4118
Gdh	50	74	80	42	0.2443	0.0561

Supplemental Table 1. Viability assay of TRiP lines crossed with *tubulin-GAL4* driver. The *tubulin-GAL4* driver line used segregates TM3, Sb[1], Ser[1] and identified by the Sb phenotype. The TRiP insert was identified by red (wild-type) eyes and the Cary control insert was identified by vermillion (v) eye color. Two replicate vials of 10 TRiP/Cary heterozygote females crossed with 5 *tub-GAL4*/TM3 males were transferred to new food every two days three or four times. Progeney were collected and males scored once daily. Data here are sums of counts across all vials for the given genotype. Two-tailed Student's t-test p-values are presented between the TRiP and Cary control inserts with either the *tub-GAL4* or TM3 (no GAL4 driver) chromosome. Viability is expected to be impacted by deletion of many of these genes.

Bibliography

ALPATOV W. W., 1930 PHENOTYPICAL VARIATION IN BODY AND CELL SIZE OF DROSOPHILA MELANOGASTER. Biol. Bull. **58**: 85–103.

ANDOLFATTO P., 2005 Adaptive evolution of non-coding DNA in Drosophila. Nature **437**: 1149–52.

ANTAO T., LOPES A., LOPES R. J., BEJA-PEREIRA A., LUIKART G., 2008 LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics **9**: 323.

AYROLES J., CARBONE M., STONE E., 2009 Systems genetics of complex traits in Drosophila melanogaster. … Genet. **41**: 299–307.

BEAUMONT M., BALDING D., 2004 Identifying adaptive genetic divergence among populations from genome scans. Mol. Ecol.

BEAUMONT M. A., NICHOLS R. A., 1996 Evaluating Loci for Use in the Genetic Analysis of Population Structure. Proc. R. Soc. London B Biol. Sci. **263**: 1619–1626.

BEGUN D. J., HOLLOWAY A. K., STEVENS K., HILLIER L. W., POH Y.-P., HAHN M. W., NISTA P. M., JONES C. D., KERN A. D., DEWEY C. N., PACHTER L., MYERS E., LANGLEY C. H., 2007 Population genomics: whole-genome analysis of polymorphism and divergence in Drosophila simulans. PLoS Biol. **5**: e310.

BENSINGER S. J., CHRISTOFK H. R., 2012 New aspects of the Warburg effect in cancer cell biology. Semin. Cell Dev. Biol. **23**: 352–61.

BHARUCHA K. N., TARR P., ZIPURSKY S. L., 2008 A glucagon-like endocrine pathway in Drosophila modulates both lipid and carbohydrate homeostasis. J. Exp. Biol. **211**: 3103–10.

BIERNE N., ROZE D., WELCH J. J., 2013 Pervasive selection or is it…? why are F ST outliers sometimes so frequent? Mol. Ecol. **22**: 2061–2064.

BRACO J. T., GILLESPIE E. L., ALBERTO G. E., BRENMAN J. E., JOHNSON E. C., 2012 Energy-dependent modulation of glucagon-like signaling in Drosophila via the AMP-activated protein kinase. Genetics **192**: 457–66.

BROUGHTON S. J., PIPER M. D. W., IKEYA T., BASS T. M., JACOBSON J., DRIEGE Y., MARTINEZ P., HAFEN E., WITHERS D. J., LEEVERS S. J., PARTRIDGE L., 2005 Longer lifespan, altered metabolism, and stress resistance in Drosophila from ablation of cells making insulin-like ligands. Proc. Natl. Acad. Sci. U. S. A. **102**: 3105–10.

CHIPPINDALE A., GIBBS A., SHEIK M., 1998 Resource acquisition and the evolution of stress resistance in Drosophila melanogaster. Evolution (N. Y).

CHIPPINDALE A. K., LEROI M., KIM S. B., ROSE M. R., 1993 Phenotypic plasticity and selection in Drosophila life-history evolution . I . Nutrition and the cost of reproduction. **93**.

CLANCY D. J., GEMS D., HARSHMAN L. G., OLDHAM S., STOCKER H., HAFEN E., LEEVERS S. J., PARTRIDGE L., 2001 Extension of life-span by loss of CHICO, a Drosophila insulin receptor substrate protein. Science **292**: 104–6.

COLOMBANI J., RAISIN S., PANTALACCI S., RADIMERSKI T., MONTAGNE J., LÉOPOLD P., 2003 A nutrient sensor mechanism controls Drosophila growth. Cell **114**: 739–49.

DOOSTZADEH J., SHOKRALLA S., ABSALAN F., JALILI R., MOHANDESSI S., LANGSTON J. W., DAVIS R. W., RONAGHI M., GHARIZADEH B., 2008 High throughput automated allele frequency estimation by pyrosequencing. PLoS One **3**: e2693.

DYKHUIZEN D. E., DEAN a M., HARTL D. L., 1987 Metabolic flux and fitness. Genetics **115**: 25–31.

EANES W. F., 2011 Molecular population genetics and selection in the glycolytic pathway. J. Exp. Biol. **214**: 165–171.

EGEA R., CASILLAS S., BARBADILLA A., 2008 Standard and generalized McDonald-Kreitman test: a website to detect selection by comparing different classes of DNA sites. Nucleic Acids Res. **36**: W157–62.

FABIAN D. K., KAPUN M., NOLTE V., KOFLER R., SCHMIDT P. S., SCHLÖTTERER C., FLATT T., 2012 Genome-wide patterns of latitudinal differentiation among populations of Drosophila melanogaster from North America. Mol. Ecol. **21**: 4748–69.

FAY J. C., WU C. I., 2000 Hitchhiking under positive Darwinian selection. Genetics **155**: 1405–1413.

FAY J. C., WYCKOFF G. J., WU C.-I., 2002 Testing the neutral theory of molecular evolution with genomic data from Drosophila. Nature **415**: 1024–6.

FELL D. A., 1992 Metabolic Control Analysis : experimental development of its theoretical and. **330**: 313–330.

FELL D., 1998 Increasing the flux in metabolic pathways: A metabolic control analysis perspective. Biotechnol. Bioeng. **58**: 121–4.

FLOWERS J. M., SEZGIN E., KUMAGAI S., DUVERNELL D. D., MATZKIN L. M., SCHMIDT P. S., EANES W. F., 2007 Adaptive evolution of metabolic pathways in Drosophila. Mol. Biol. Evol. **24**: 1347–54.

FOLL M., GAGGIOTTI O., 2008 A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics **180**: 977–93.

FONTANA L., PARTRIDGE L., LONGO V. D., 2010 Extending healthy life span--from yeast to humans. Science **328**: 321–6.

FOURCADE Y., CHAPUT-BARDY A., SECONDI J., FLEURANT C., LEMAIRE C., 2013 Is local selection so widespread in river organisms? Fractal geometry of river networks leads to high bias in outlier detection. Mol. Ecol. **22**: 2065–73.

FU Y. X., LI W. H., 1993 Statistical tests of neutrality of mutations. Genetics **133**: 693–709.

GEIGENBERGER P., STITT M., FERNIE A. R., 2004 Metabolic control analysis and regulation of the conversion of sucrose to starch in growing potato tubers. Plant, Cell Environ. **27**: 655–673.

GEMINARD C., RULIFSON E. J., LEOPOLD P., 2009 Remote control of insulin secretion by fat cells in Drosophila. Cell Metab. **10**: 199–207.

GOSSET C. C., BIERNE N., 2013 Differential introgression from a sister species explains high F(ST) outlier loci within a mussel species. J. Evol. Biol. **26**: 14–26.

HARBISON S. T., 2004 Quantitative Trait Loci Affecting Starvation Resistance in Drosophila melanogaster. Genetics **166**: 1807–1823.

HOFFMANN A. A., HARSHMAN L. G., 1999 Desiccation and starvation resistance in Drosophila: patterns of variation at the species, population and intrapopulation levels. Heredity (Edinb). **83**: 637–643.

HOFFMANN A. A., 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. Funct. Ecol. **19**: 222–227.

HOHENLOHE P. A., BASSHAM S., ETTER P. D., STIFFLER N., JOHNSON E. A., CRESKO W. A., 2010 Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet. **6**: e1000862.

HOLSINGER K. E., WEIR B. S., 2009 Genetics in geographically structured populations: defining, estimating and interpreting F(ST). Nat. Rev. Genet. **10**: 639–50.

HUDSON R. R., KREITMAN M., AGUADÉ M., 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116**: 153–159.

IKEYA T., GALIC M., BELAWAT P., NAIRZ K., HAFEN E., 2002 Nutrient-Dependent Expression of Insulin-like Peptides from Neuroendocrine Cells in the CNS Contributes to Growth Regulation in Drosophila. Curr. Biol. **12**: 1293–1300.

IMAI T., 1933 The influence of temperature on variation and inheritance of bodily dimensions inDrosophila melanogaster. Wilhelm Roux' Arch. f□r Entwicklungsmechanik der Org. **128**: 634–660.

JENKINS D. L., ORTORI C. a, BROOKFIELD J. F., 1995 A test for adaptive change in DNA sequences controlling transcription. Proc. Biol. Sci. **261**: 203–7.

KACSER H., BURNS J. A., 1981 The Molecular Basis of Dominance. Genetics **97**: 639–666.

KAO J. Y., ZUBAIR A., SALOMON M. P., NUZHDIN S. V, CAMPO D., 2015 Population genomic analysis uncovers African and European admixture in Drosophila melanogaster populations from the south-eastern United States and Caribbean Islands. Mol. Ecol. **24**: 1499–509.

KARAN D., DAHIYA N., MUNJAL A. K., GIBERT P., MORETEAU B., PARKASH R., DAVID J. R., 1998 Desiccation and Starvation Tolerance of Adult Drosophila: Opposite Latitudinal Clines in Natural Populations of Three Different Species. Evolution (N. Y). **52**: 825–831.

KREBS H. A., 1967 The redox state of nicotinamide adenine dinucleotide in the cytoplasm and mitochondria of rat liver. Adv. Enzyme Regul. **5**: 409–434.

KROEMER G., POUYSSEGUR J., 2008 Tumor cell metabolism: cancer's Achilles' heel. Cancer Cell **13**: 472–82.

LAVEBRATT C., SENGUL S., 2006 Single nucleotide polymorphism (SNP) allele frequency estimation in DNA pools using Pyrosequencing. Nat. Protoc. **1**: 2573–82.

LAVINGTON E., COGNI R., KUCZYNSKI C., KOURY S., BEHRMAN E. L., O'BRIEN K. R., SCHMIDT P. S., EANES W. F., 2014 A small system--high-resolution study of metabolic adaptation in the central metabolic pathway to temperate climates in Drosophila melanogaster. Mol. Biol. Evol. **31**: 2032–41.

LEE G., PARK J. H., 2004 Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in Drosophila melanogaster. Genetics **167**: 311–23.

LIBRADO P., ROZAS J., 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25**: 1451–1452.

LOSMAN J.-A., KAELIN W. G., 2013 What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. Genes Dev. **27**: 836–52.

LUIKART G., ENGLAND P. R., TALLMON D., JORDAN S., TABERLET P., 2003 The power and promise of population genomics: from genotyping to genome typing. Nat. Rev. Genet. **4**: 981–94.

MACKAY T. F. C., RICHARDS S., STONE E. A., BARBADILLA A., AYROLES J. F., ZHU D., CASILLAS S., HAN Y., MAGWIRE M. M., CRIDLAND J. M., RICHARDSON M. F., ANHOLT R. R. H., BARRÓN M., BESS C., BLANKENBURG K. P., CARBONE M. A., CASTELLANO D., CHABOUB L., DUNCAN L., HARRIS Z., JAVAID M., JAYASEELAN J. C., JHANGIANI S. N., JORDAN K. W., LARA F., LAWRENCE F., LEE S. L., LIBRADO P., LINHEIRO R. S., LYMAN R. F., MACKEY A. J., MUNIDASA M., MUZNY D. M., NAZARETH L., NEWSHAM I., PERALES L., PU L.-L., QU C., RÀMIA M., REID J. G., ROLLMANN S. M., ROZAS J., SAADA N., TURLAPATI L., WORLEY K. C., WU Y.-Q., YAMAMOTO A., ZHU Y., BERGMAN C. M., THORNTON K. R., MITTELMAN D., GIBBS R. A., 2012 The Drosophila melanogaster Genetic Reference Panel. Nature **482**: 173–178.

MARTINS I. J., HONE E., FOSTER J. K., SÜNRAM-LEA S. I., GNJEC A., FULLER S. J., NOLAN D., GANDY S. E., MARTINS R. N., 2006 Apolipoprotein E, cholesterol metabolism, diabetes, and the convergence of risk factors for Alzheimer's disease and cardiovascular disease. Mol. Psychiatry **11**: 721–36.

MCDONALD J. H., KREITMAN M., 1991 Adaptive protein evolution at the Adh locus in Drosophila. Nature **351**: 652–654.

MERRITT T. J. S., SEZGIN E., ZHU C.-T., EANES W. F., 2006 Triglyceride pools, flight and activity variation at the Gpdh locus in Drosophila melanogaster. Genetics **172**: 293–304.

MONDAL K., DASTIDAR A. G., SINGH G., MADHUSUDHANAN S., GANDE S. L., VIJAYRAGHAVAN K., VARADARAJAN R., 2007 Design and isolation of temperature-sensitive mutants of Gal4 in yeast and Drosophila. J. Mol. Biol. **370**: 939–50.

O'BRIEN S. J., MACINTYRE R. J., 1972 THE alpha-GLYCEROPHOSPHATE IN DROSOPHILA MELANOGASTER II. GENETIC ASPECTS. Genetics **71**: 127–138.

O'BRIEN S. J., MACINTYRE R. J., 1972 The α-glycerophosphate cycle inDrosophila melanogaster. I. Biochemical and developmental aspects. Biochem. Genet. **7**: 141–161.

OKAMOTO N., YAMANAKA N., YAGI Y., NISHIDA Y., KATAOKA H., O'CONNOR M. B., MIZOGUCHI A., 2009 A fat body-derived IGF-like peptide regulates postfeeding growth in Drosophila. Dev. Cell **17**: 885–91.

OLSON-MANNING C. F., LEE C.-R., RAUSHER M. D., MITCHELL-OLDS T., 2013 Evolution of flux control in the glucosinolate pathway in Arabidopsis thaliana. Mol. Biol. Evol. **30**: 14–23.

PARTRIDGE L., ALIC N., BJEDOV I., PIPER M. D. W., 2011 Ageing in Drosophila: the role of the insulin/Igf and TOR signalling network. Exp. Gerontol. **46**: 376–381.

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 on JSTOR. Evolution (N. Y). **48**: 1269–1276.

PARTRIDGE L., GREEN A., FOWLER K., 1987 Effects of egg-production and of exposure to males on female survival in Drosophila melanogaster. J. Insect Physiol. **33**: 745–749.

PARTRIDGE L., PIPER M. D. W., MAIR W., 2005 Dietary restriction in Drosophila. Mech. Ageing Dev. **126**: 938–50.

PIPER M. D. W., BLANC E., LEITÃO-GONÇALVES R., YANG M., HE X., LINFORD N. J., HODDINOTT M. P., HOPFEN C., SOULTOUKIS G. A., NIEMEYER C., KERR F., PLETCHER S. D., RIBEIRO C., PARTRIDGE L., 2014 A holidic medium for Drosophila melanogaster. Nat. Methods **11**: 100–5.

REAUME C. J., SOKOLOWSKI M. B., 2006 The nature of Drosophila melanogaster. Curr. Biol. **16**: R623–8.

ROBINSON S. J. W., ZWAAN B., PARTRIDGE L., 2000 STARVATION RESISTANCE AND ADULT BODY COMPOSITION IN A LATITUDINAL CLINE OF DROSOPHILA MELANOGASTER. Evolution (N. Y). **54**: 1819.

ROSE M., VU L., PARK S., GRAVES J., 1992 Selection on stress resistance increases longevity in Drosophila melanogaster. Exp. Gerontol.

RULIFSON E. J., KIM S. K., NUSSE R., 2002 Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. Science **296**: 1118–20.

SERVICE P. M., ROSE M. R., 1985 Genetic Covariation Among Life-History Components: The Effect of Novel Environments. Evolution (N. Y). **39**: 943–945 CR – Copyright © 1985 Society for th.

SEZGIN E., DUVERNELL D. D., MATZKIN L. M., DUAN Y., ZHU C.-T., VERRELLI B. C., EANES W. F., 2004 Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in Drosophila melanogaster. Genetics **168**: 923–31.

SHAPIRO J. a, HUANG W., ZHANG C., HUBISZ M. J., LU J., TURISSINI D. a, FANG S., WANG H.-Y., HUDSON R. R., NIELSEN R., CHEN Z., WU C.-I., 2007 Adaptive genic evolution in the Drosophila genomes. Proc. Natl. Acad. Sci. U. S. A. **104**: 2271–6.

SOKAL R. R., ROHLF F. J., 1995 Biometry: the principles andpractice of statistics in biological research. WH. Free. Co., San Fr. SokalBiometry Princ. Pract. Stat. Biol. Res.

STINCHCOMBE J. R., HOEKSTRA H. E., 2008 Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity (Edinb). **100**: 158–70.

TAJIMA F., 1989 The effect of change in population size on DNA polymorphism. Genetics **123**: 597–601.

TATAR M., 2011 The plate half-full: status of research on the mechanisms of dietary restriction in Drosophila melanogaster. Exp. Gerontol. **46**: 363–8.

WANG B., GOODE J., BEST J., MELTZER J., SCHILMAN P. E., CHEN J., GARZA D., THOMAS J. B., MONTMINY M., 2008 The insulin-regulated CREB coactivator TORC promotes stress resistance in Drosophila. Cell Metab. **7**: 434–444.

WAPLES R. S., 2005 Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? Mol. Ecol. **14**: 3335–52.

WILKINS J. F., 2004 A separation-of-timescales approach to the coalescent in a continuous population. Genetics **168**: 2227–44.

WU W., ZHAO S., 2013 Metabolic changes in cancer : beyond the Warburg effect. **45**: 18–26.

ZHU C.-T., RAND D. M., 2012 A hydrazine coupled cycling assay validates the decrease in redox ratio under starvation in Drosophila. (LC Lyons, Ed.). PLoS One **7**: e47584.