Genomic and Phenotypic Effects of Selection for Starvation Resistance in *Drosophila*

Running Head: Intensely-Selected Obese *Drosophila* Populations

James N. Kezos*[ah], Mark A. Phillips[ab], Misty D. Thomas[c], Akamu J. Ewunkem[b], Grant A. Rutledge[af], Thomas T. Barter[a], Marta A. Santos[ae], Brandon D. Wong[a], Kenneth R. Arnold[a], Laura A. Humphrey[a], Albert Yan[a], Chloe Nouzille[a], Isaias Sanchez[a], Larry G. Cabral[ag], Timothy J. Bradley[a], Laurence D. Mueller[a], Joseph L. Graves Jr[c], and Michael R. Rose[a]

[a] Department of Ecology and Evolutionary Biology,
School of Biological Sciences,
University of California, Irvine, CA 92697-2525

[b]Department of Integrative Biology,
College of Science,
Oregon State University, Corvallis, OR, 97331

[c] Department of Nanoengineering,
Joint School of Nanoscience and Nanoengineering
North Carolina A&T State University, Greensboro, NC 27401

[d]Department of Biology,
College of Science and Technology
North Carolina A&T State University, Greensboro, NC 27411

[e] cE3c – Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências da Universidade de Lisboa, Portugal

[f] USDA HNRCA at Tufts University
Boston, MA 02111

[g] Lyceum Pharmaceuticals, Inc.
Irvine, CA 92620

[h] Department of Development, Aging, and Regeneration
Sanford Burnham Prebys Medical Discovery Institute
La Jolla, CA 92037

*Corresponding author: jkezos@sbpdiscovery.org; 714-598-9922

James N. Kezos: jkezos@sbpdiscovery.org; 714-598-9922
Mark A. Phillips: mphillips6789@gmail.com; 727-637-6797
Misty D. Thomas: mthomas1@ncat.edu
Akamu J. Ewunkem: aewunkem@ncat.edu
Grant A. Rutledge: grant.rutledge@tufts.edu; (858) 243-4768

Thomas T. Barter: ttbarter317@gmail.com; 714-469-2509
Marta A. Santos: martasantos@fc.ul.pt
Brandon D. Wong: bdwong2@gmail.com; 415-760-8499
Kenneth R. Arnold: kenneth.arnold92@gmail.com; 626-646-9345
Laura A. Humphrey: lauraahumphrey@gmail.com; 562-896-5897
Albert Yan: ayanedhs@gmail.com; 559-827-1726
Chloe Nouzille: cmnouzille@gmail.com; 818-6346877
Isaias Sanchez: isaiass@uci.edu; 760-845-7418
Larry G. Cabral: lgcabral5@gmail.com
Timothy J. Bradley: tbradley@uci.edu
Laurence D. Mueller: ldmuelle@uci.edu; 949-824-4744
Joseph L. Graves Jr: gravesjl@ncat.edu
Michael R. Rose: mrrose@uci.edu; 949-824-81211

Key Words or Phrases:
Intense Selection, Starvation Resistance, Obesity, Heart Robustness, Metabolic Reserves, Mortality, Genome Sequencing

What Is Already Known:
Past studies have shown that intensely-selected populations are useful materials for analyzing the mechanistic foundations of adaptation in laboratory fruit flies. Starvation-resistant and obese *Drosophila* populations are a powerful tool not only for studying the evolution of starvation responses, but also for studying metabolic disorders and related cardiac dysfunction.

What This Study Adds:
This study applied an intense selection regime on a large, outbred, ten-fold replicated set of *Drosophila melanogas*ter populations for six years, tripling the average lipid contents and doubling the average cardiac arrest rates in respect to their ancestral populations. We identified an additional 131 genes of interest that the most notable *Drosophila* starvation-selection experiment to date (Hardy et al. 2015, Hardy et al. 2018) did not find in their populations. Our result of the complex age-dependent relationship between lipid levels and mortality rates in our obese fruit flies echoes some features of the relationship between obesity and mortality in human populations, once again supporting the importance of employing highly variable, outbred *Drosophila* populations for human metabolic and cardiac-related disorders.

I. ABSTRACT

In experimental evolution, we impose functional demands on laboratory populations of model organisms using selection. After enough generations of such selection, the resulting populations constitute excellent material for physiological research. An intense selection regime for increased starvation resistance was imposed on ten, large, outbred *Drosophila* populations. We observed the selection responses of starvation and desiccation resistance, metabolic reserves, and heart robustness via electrical pacing. Furthermore, we sequenced the pooled genomes of these populations. As expected, significant increases in starvation resistance and lipid content were found in our ten intensely-selected SCO populations. The selection regime also improved desiccation resistance, water content, and glycogen content among these populations. Additionally, the average rate of cardiac arrests in our ten obese SCO populations was double the rate of the ten ancestral CO populations. Age-specific mortality rates were increased at early adult ages by selection. Genomic analysis revealed a large number of SNPs across the genome that changed in frequency as a result of selection. These genomic results were similar to those obtained in our laboratory from less direct selection procedures. The combination of extensive genomic and phenotypic differentiation between these ten populations and their ancestors make them a powerful system for the analysis of the physiological underpinnings of starvation resistance.

II. INTRODUCTION

Stress affects both vertebrate and invertebrate organisms throughout their lifetime, and both groups experience and defend against similar stressors, sometimes but not always using the same or similar physiological machinery. Evolutionary physiologists take advantage of stress, using it as a tool either (1) to measure an organism's physical robustness, or (2) to create

differentiated populations with which to study adaptation. In the latter case, such differentiated populations provide excellent material for physiological analysis of organismal functions related to stress.

A common theme of *Drosophila* evolutionary physiology has been the study of the physiological basis for the experimental evolution of populations that have been selected for very different life-histories (reviewed in Burke and Rose 2009, as well as Rose et al. 2004). It is important to note that, though populations with different life-cycle timings are not specifically selected for different levels of stress resistance, they do nonetheless show distinctive enhancements in a variety of stress and performance characters. For example, early on it was found that a variety of stress resistances and performance characters were improved in experimentally evolved stocks with postponed senescence (vid. Service et al. 1985; Service et al. 1988; Graves and Rose 1990; Graves et al. 1992; Rose et al. 2004). Service et al. (1985) first analyzed the effects of evolutionarily postponed senescence on stress resistance, finding an increase in starvation resistance, desiccation resistance, and resistance to ethanol vapor (15%). In addition to desiccation resistance, Graves et al. (1988) as well as Graves and Rose (1990) showed that selection for postponed senescence also increased flight endurance. Physiological analysis of stress resistance in populations cultured with different life cycles has revealed evolutionary changes in multiple physiological characters, from lipid and glycogen content to water content and rates of water loss (Graves et al. 1992; Gibbs et al. 1997; Chippindale et al. 1998; Djawdan et al. 1998; Rose et al. 2004). Djawdan et al. (1998) found strong, positive correlations between energy content and starvation resistance.

Our lab and other colleagues have gone on to examine the response of life history and physiological traits to intense, or focal, selection regimes (e.g. selection for increased starvation

resistance). We have called this type of selection paradigm "culling selection" in the past (e.g. Rose et al. 1990), but it could also be described as *intense selection.* An intense selection regime can eventually achieve extreme levels of functional differentiation. Such extreme functional differentiation is achieved by the use of environments so inimical to survival that only a small percentage of each generation survives selection. Our past studies have shown that such stringently selected populations are useful material for analyzing the mechanistic foundations of adaptation in lab fruit flies. For example, Archer et al. (2007) produced flies that can survive complete desiccation for as much as ten times longer than unselected flies. Additionally, selection for increased resistance to starvation or desiccation in *Drosophila* can result in slower development, reduced larval or pupal viability, reduced early fecundity, and increased adult longevity (Rose et al. 1990; Rose et al. 1992; Chippindale et al. 1996; Bradley et al. 1999).

Here we present a well-replicated set of populations intensely selected for increased starvation resistance: ten "SCO" populations that have been undergoing selection since August 2010. This intense selection regime rapidly produced changes in body shape, body size, and stress resistance. We knew from previous experiments employing selection for starvation resistance that this selection regime would increase both starvation resistance and lipid content (Rose et al. 1992; Djawdan et al. 1998; Harshmann et al. 1999; Hardy et al. 2015). "Pseudocomparative" studies (vid. Burke and Rose, 2009) and reverse-selection studies (e.g. Passananti et al. 2004) have already found positive correlations among starvation resistance, lipid content, and longevity. Selection for moderate increases in these stress resistance capacities also increased mean life span (Rose et al. 1992), but this effect is reversed at still greater levels of stress resistance (Archer et al. 2003; Phelan et al. 2003; Archer et al. 2007). Increasing lipid content to an extreme level may be beneficial in surviving against starvation or even desiccation,

however, it could physiologically disrupt normal function of multiple organ systems, harming the overall health of the organism. Longevity begins to regress, implying that the relationship between starvation resistance and longevity becomes antagonistic at extreme levels.

For this study, we functionally and genomically characterize our ten, intensely selected, starvation resistant populations. In particular, we compare them to both their ancestors (CO_{1-5}), which have not been subjected to such intense focal selection, as well as five other populations (nCO_{1-5}) that have converged on their ancestors with respect to both life history (Burke et al. 2016) and genomics (Graves et al. 2017). We examine functional characters that have long been studied in *Drosophila*, specifically stress resistance, metabolic reserves and longevity, as well as a functional character of more recent interest in our laboratory, heart robustness. We then compared patterns of genomic differentiation between our selected and control populations to identify candidate genes associated with adaptation to this selection regime. The present study shows how rapid and intense selection can induce extensive functional and genomic differentiation.

III. Methods

3.1 Populations Used

This study employed 20 of the large, outbred, and highly differentiated populations created by the Rose laboratory since 1980 (e.g. Rose 1984; Rose et al. 1992; Rose et al. 2004). All of the populations assayed here descend from a single *Drosophila melanogaster* population, called IV. The IV population originated in 1975 as a sample of *D. melanogaster* caught in Amherst, Massachusetts. After four and a half years of laboratory culture, the B_{1-5} populations (baseline) and O_{1-5} populations (70-day generation cycle) were derived from the single IV

population in 1980 (Rose 1984; Rose et al. 2004). The 20 populations studied here are referred to as $CO_{1\text{-}5}$, $nCO_{1\text{-}5}$, $SCO_{1\text{-}5A}$, and $SCO_{1\text{-}5B}$.

The $CO_{1\text{-}5}$ populations (28-day generation cycle) were derived from the $O_{1\text{-}5}$ populations in 1989 (Rose et al. 1992). For each CO population, the flies are maintained in two cages (an alpha cage and a beta cage). For example, the CO_1 population is split into two cages, $CO_{1\alpha}$ and $CO_{1\beta}$. To be clear, these two cages represent the same population, CO_1. The C-type selection regime and generation cycle is as follows. The egg laying process for the next generation is considered "Day 0". After a two-week development period, on day 14 from eggs laid, 1200-1500 adult flies (approximate even mix of male and females flies) are transferred from vials into a rectangular plexiglass cage. Adult flies are fed the standard banana medium in a petri dish, with food replaced every other day. On day 28 from eggs laid, a standard banana medium with added yeast paste is placed into the cage for egg collection for the ensuing generation. To ensure genetic similarity, a heavy mixing of flies from the two cages is conducted every 10 generations.

The $nCO_{1\text{-}5}$ populations (28-day generation cycle) were derived from the $O_{1\text{-}5}$ populations in 2009, with the numerical nomenclature representing the direct ancestry (nCO_1 population derived from O_1 population). Similar to the CO populations, each nCO population is maintained in an alpha cage and a beta cage. For example, the nCO_1 population is split into these two cages, $nCO_{1\alpha}$ and $nCO_{1\beta}$. Every ten generations, a mixing of flies between the alpha and beta cage for each population was conducted to ensure genetic similarity. The five CO and nCO populations follow the same selection regime, the C-type (Burke et al.2016). And despite the differences in generation number under C-type selection, the five CO and five nCO populations are phenotypically and genetically identical (Graves et al. 2017). The five CO and nCO populations

were used in order to have 10 C-type populations to compare to the 10 S-type populations, which allows for a statistically stronger genomic comparison (Graves et al. 2017).

The $SCO_{1\text{-}5A}$ and $SCO_{1\text{-}5B}$ populations (28-day generation cycle) are ten populations intensely selected for starvation resistance. The ten SCO populations were derived from the five CO populations in August 2010. From each CO population, two SCO populations were derived. For example, CO_1 is the ancestral population to SCO_{1A} and SCO_{1B}. From that point forward, SCO_{1A} and SCO_{1B} were treated as two unique populations with a shared ancestor. To maintain ample genetic variation and population size post selection process, each S-type population also had an alpha and a beta cage (i.e. $SCO_{1A\text{-}\alpha}$ and $SCO_{1A\text{-}\beta}$). The S-type selection regime and generation cycle is as follows. The egg laying process for the next generation is considered "Day 0". After a two-week development period, day 14 from eggs laid, 1200-1500 adult flies (approximate even mix of male and females flies) are transferred from vials into a rectangular plexiglass cage. Adult flies are fed the standard banana medium (in a petri dish) with added yeast paste for three days. On day 17 from egg, the standard food petri dish is removed from the cage and replaced with a nonnutritive agar petri dish to begin the starvation period. Each population is exposed to the nonnutritive agar until a 75-80% mortality threshold has been reached. When an approximate 600-700 of flies were observed laying motionless in the cage with no response to provocation, we would begin conducting mortality checks by collecting and counting the dead flies from each cage. The starvation period and subsequent mortality checks would continue until 75%, or at least 900 flies have been removed and counted from the cage. When the 75% mortality threshold was reached in both the alpha and beta cages within each population, flies from the two cages were condensed into one cage (e.g. $SCO_{1A\alpha}$ and $SCO_{1A\beta}$ into SCO_{1A}; approximate total of 600-700 flies). A three-day period of standard banana medium with added

yeast paste follows the starvation period with egg collection for the succeeding generation occurring on day 28 from egg. At the beginning of this experiment, the starvation period took just three days to achieve 75-80% mortality. Currently, the starvation period lasts for approximately 10 days. At the time of the experimental assays reported here, the SCO populations had experienced between 55 and 86 generations of intense selection. The group of populations subjected to this selection regime is referred to as the S-type treatment group.

These large, outbred populations were maintained on a banana medium (agarose, banana, light and dark corn syrup, barley malt, yeast, ethanol and water). with a 24L:0D light cycle. Each new population was founded from the same number replicate of their ancestral population (i.e. population SCO_{1A} was derived from CO_1, population SCO_{2A} was derived from CO_2). All 20 populations were kept at moderately large census populations sizes (N > 1,000) to avoid confounding inbreeding effects. All 20 populations experienced a 14-day developmental period in vials before being transferred to an acrylic cage for the remainder of their generation cycle.

3.2 Functional Assay Methods

Rearing protocols

Two run-in generations of 14-day life-cycles were used to remove any parental or grand-parental epigenetic effects not related to the selection regime that could potentially affect the measured physiological responses. The populations were cultured in banana medium from egg to adult, on a 24L:0D light schedule. A petri dish with the banana medium was placed into each respective cage. Adult flies were given 24 hours to lay eggs on the media. From the plate, eggs were collected at a density of 60 to 80 eggs and subsequently placed in a vial containing fresh media. At the end of each run-in generation (day 14 from egg), the populations were transferred

to an acrylic cage. Replicate populations of the same number were handled in parallel at all stages. On day 14 of the second run-in generation, the adults were assigned at random to one of the following eight assays.

At the beginning of the 14-day from egg desiccation resistance and starvation resistance assays, the CO and nCO populations had been under 325 and 72 generations of C-type selection. The SCO_A and SCO_B populations had completed 55 generations of intense selection. For the 14-day from egg cardiac arrest rate assay, the CO and nCO populations were under selection for 332 and 79 generations, respectively. The SCO_A and SCO_B populations had completed 62 generations of intense selection. At the beginning of this mortality assay, the CO and nCO populations had been under C-type selection for 341 and 88 generations. The SCO_A and SCO_B populations had completed 70 generations of intense selection. For the 14-day from egg glycogen content, lipid content, and water content assays, the CO and nCO populations had been under 352 and 98 generations of selection. The SCO_A and SCO_B populations had completed 81 generations of intense selection. For the age-specific cardiac arrest rate and lipid content assays, the CO and SCO_A populations had been under their respective selection regime for 357 and 86 generations, respectively.

Glycogen Content

For each population, six groups of 10 females (14 days old from egg) were anesthetized using ethyl ether and placed in 1.7 milliliter microcentrifuge tubes at room temperature. The next day, each group was placed in aluminum weighing boats and placed in an oven for one hour at 60 ºC to remove any water still remaining in the flies after this initial 24-hour storage period. Each group was transferred to their respective microcentrifuge tube where 700 microliters of deionized water was added to each tube. The flies were then ground using a hand-held battery-

operated grinder. Each tube was boiled in water for five minutes. Once complete, 500 microliters of the supernatant was carefully pipetted into a new microcentrifuge tube to minimize noise from precipitate contamination. Each tube was then vortexed to thoroughly mix the contents, then 100 microliters from each tube was transferred to a 13 x 100-millimeter test tube. Three milliliters of an anthrone reagent was added to each test tube. The anthrone reagent composition was 150 milligrams of anthrone per 100 milliliters of 72% sulfuric acid. Each test tube was then incubated in a water bath set to 90 ºC for 10 minutes. Two one-milliliter samples from each test tube were placed in their own cuvettes. Absorbance was then measured using a Perkin Elmer Lambda spectrophotometer at a wavelength of 620 nanometers. All measurements were taken within 10 minutes of being removed from the water bath. Five controls of known glycogen concentration underwent the same process.

Water Content

For each population, six groups of 10 females (14 days old from egg) were anesthetized using ethyl ether. Each group was placed in aluminum weighing boats, and had their wet mass measured. The flies were frozen overnight and placed in a drying oven the next day at 60 °C for 24 hours. Each group was then reweighed. The difference between the wet mass and dry mass represented the water content of that group.

Lipid Content

For each population, lipid content was measured for six groups of 10 females at age 14 days from egg. After the dry masses of the groups were recorded for the water content assay, each group was placed in their own Whatman thimble. The thimbles were placed in the extractor of a Soxhlet apparatus. Petroleum ether was used as the extraction solvent. The thimbles were in the Soxhlet apparatus for 24 hours, and upon completion, were removed and placed in a

drying oven at 60 °C for one hour. The post-extraction mass of each group was recorded using a microbalance scale. The difference between the dry mass and the post-extraction mass of each group represented the specific group's lipid content.

A second set of lipid content measurements were made for the five CO and five SCO_A populations at six different adult ages. These measurements were conducted at ages 14, 21, 28, 35, 42, and 49 days from egg. The protocol is the same as the previous paragraph.

Desiccation resistance assay

Thirty individual female flies (14 days old from egg) from each population were placed in their own desiccant straw. A piece of cheesecloth separated the fly from the pipet tip at the end of the straw that contained 0.75 grams of desiccant (anhydrous calcium sulfate). The pipette tip containing desiccant was sealed with a layer of Parafilm. Mortality was checked hourly, using lack of movement under provocation as a sign of death. Note that this was a materially different procedure than the one we have employed previously in our studies of desiccation resistance (e.g. Service et al. 1985; Graves et al. 1992; Djawdan et al. 1998), which used vials.

Starvation resistance assay

Thirty individual female flies (14 days old from egg) from each population were placed in their own starvation straw with agar. The agar plug provides adequate humidity, but no nutrients. Mortality was checked every four hours, using lack of movement under provocation as a sign of death.

Cardiac Arrest Rate Assay

Forty female flies (14 days old from egg) from each replicate per stock were chosen at random. The flies were anesthetized for three minutes using triethylamine, also known as FlyNap, and then placed on a microscope slide prepared with foil and two electrodes. FlyNap

was chosen as the anesthetic because of its minimal effect on heart function and heart physiology when administered for more than one minute, but less than the lethal time of five minutes (Chen and Hillyer 2013). The cold-shock method was not used as an anesthetic for the cardiac pacing assay, because the flies need to be fully anesthetized throughout the procedure. If the flies regain consciousness, the added stress and abdominal contractions while trying to escape would alter heart rate and function more than the side-effects of FlyNap's do. Paternostro et al. (2001) found that FlyNap does the least amount of cardiac disruption compared to the two other substances commonly used for *Drosophila* anesthesia, carbon dioxide and ether. Two electrodes were attached to a square-wave stimulator in order to produce electric pacing of heart contraction. Anesthetized flies were attached to the slide between the foil gaps using a conductive electrode jelly touching the two ends of the fly body, specifically the head and the posterior abdomen tip. The electrical shock settings for this assay were 40 volts, six Hertz, and 10 ms pulse duration. The settings of the electrical pacing assay were chosen in order to increase the contraction rate to such a high enough level that cardiac arrest would be consistently induced. Using an inverted compound microscope, an initial check of whether the heart was contracting or in arrest was made immediately after completion of the 30 second shock. A second check was made after a two-minute "recovery" period to determine if the heart was in cardiac arrest. The protocol for this assay was originally outlined in Wessells and Bodmer (2004).

A second set of cardiac arrest rate measurements were made for the CO and SCO_A populations at six different adult ages. These measurements were conducted at ages 14, 21, 28, 35, 42, and 49 days from egg. The protocol is the same as the above-mentioned 14-day from egg pacing experiment.

Adult Mortality

Flies from all populations of CO, nCO, and SCO were handled in parallel for two run-in generations in vials. After 14 days of development in vials, approximately 800 to 1200 adult flies from each of the populations of the CO, nCO, SCO_A and SCO_B were transferred into Plexiglass cages, with multiple cohort cages for each population. From then on, they were fed a standard banana diet, with the food being replaced every day. Dead flies were collected from cages, separated by sex, and counted at the same time every day, with necessary precautions taken to minimize amount of escapers. At the end of the mortality assay, the total number of dead flies (male and female) counted for each population becomes the initial starting number of flies in that population, once again, minimizing the effects of possible escapers. Age-specific mortality rates were obtained over all adult ages for the C-type and S-type populations. Each assayed cohort was initially maintained in five cages; to control population density over time, cages were consolidated as the number individuals in a cage reached 50% of the standardized density that we used for that volume of cage. Flies were briefly anesthetized using carbon dioxide during these consolidations.

3.3 Statistical Methods

Analysis for Physiological Characters

Each of the five measured phenotypes, glycogen content, lipid content, water content, starvation resistance, and desiccation resistance, were analyzed separately. We now outline the two linear mixed-effects models we used for starvation resistance.

The first model was used for our within selection regime comparisons (CO vs nCO and SCO_A vs SCO_B). Let y_{ijk} be the measured starvation resistance for selection treatment-i (i=1 (CO

or SCO_A) and 2 (nCO or SCO_B)), population-*j* (*j*= 1..10), and individual-*k* (*k*=1..n_l). Then the effects of the fixed and random effects can be modeled as,

$$y_{ijk} = \mu + \alpha\delta_i + \beta_j + \varepsilon_{ijk} \quad (1)$$

where δ_i=0 if *i*=1 and 1 otherwise.

The second model was used for our selection type comparisons (i.e. C-type selection vs. S-type selection). Let y_{ijk} be the measured starvation resistance for selection type-*i* (*i*=1 (C type) and 2 (S type)), population-*j* (*j*= 1..20), and individual-*k* (*k*=1..n_l). Then the effects of the fixed and random effects can be modeled as,

$$y_{ijk} = \mu + \alpha\delta_i + \beta_j + \varepsilon_{ijk} \quad (2)$$

where δ_i=0 if *i*=1 and 1 otherwise.

For these two models, the main effects of selection regime and replicate population were measured by α and β, respectively. The different populations contributed random effects to these measurements by genetically based differences that arise due to random genetic drift and were measured by β while individual random variation was measured by ε. Both sources of random variation were assumed to be independent normally distributed random variables with zero means. The model parameters were estimated with the R *lme* function (R Core Team 2015).

Cochran-Mantel-Haenszel (CMH) tests were used to analyze the rates of cardiac arrests between two different stocks (i.e. CO vs nCO). The CMH test is used when there are repeated tests of independence, or multiple 2x2 tables of independence. This is the equation for the CMH test statistic, with the continuity correction included, that we used for our statistical analyses:

$$X^2_{\mathrm{MH}} = \frac{\{|\Sigma\left[a_i - \frac{(a_i + b_i)(a_i + c_i)}{n_i}\right]| - 0.5\}^2}{\Sigma(a_i + b_i)(a_i + c_i)(b_i + d_i)(c_i + d_i)/(n_i^3 + n_i^2)}$$

We designated "a" and "b" as the number of cardiac arrests in population *i* of the first stock and population *i* of the second stock, respectively. We designated "c" and "d" as the number of contracting hearts in the two populations. The n_i represented the sum of a_i, b_i, c_i, and d_i. The subscript *i* (i = 1..5), represented one of the five replicate populations within each of the four stocks.

Analysis for Age-Specific Mortality Rates

We first tested for convergence within the selection treatments (i.e. CO vs. nCO, and SCO_A and SCO_B) for effects of selection on mortality rates over the adult lifespan. The observations consisted of mortality rates at a particular age (t) but within a small age interval (*k*-1, 2,…, m). These age intervals were chosen to span the ages, such that all comparison populations still had live flies. Within each interval, mortality rates were modeled by a straight line and allowing selection regime (*i*=1 (CO or SCO_A), *i*=2 (nCO or SCO_B)) to affect the intercept of that line but not the slope. Slopes were allowed to vary between intervals. Populations (*j*=1,…,10) were assumed to contribute random variation to these measures. Let y_{ijkt} be the measured mortality rate at age-*t*, interval-*k*, selection treatment-*i*, population-*j*. Then the effects of the fixed and random effects can be modeled as,

$$y_{ijkt} = \alpha + \beta_k + \delta_i\gamma_i + (\omega + \pi_k\delta_k)t + \delta_k\delta_i\mu_{ik} + c_j + \varepsilon_{ijkt} \quad (3)$$

where δ_s=0 if *s*=1, and 1 otherwise.

We then tested for divergence between the selection treatments (i.e. C-type selection vs. S-type selection). Let y_{ijkt} be the measured mortality rate at age-t, interval-k, selection treatment type-*i* (*i*=1 (C type) and 2 (S type)), population-*j* (*j*= 1..20). Then the effects of the fixed and random effects can be modeled as

$$y_{ijkt} = \alpha + \beta_k + \delta_i\gamma_i + (\omega + \pi_k\delta_k)t + \delta_k\delta_i\mu_{ik} + c_j + \varepsilon_{ijkt} \quad (4)$$

where δ_s=0 if s=1, and 1 otherwise.

In equations 3 and 4, the c_j and ε_{ijkt} are independent standard normal random variables. The different populations contributed random effects to these measurements by genetically based differences that arise due to random genetic drift and were measured by c while individual random variation was measured by ε. Both sources of random variation were assumed to be independent normally distributed random variables with zero means. The model parameters were estimated with the R *lme* function (R Core Team 2015).

3.4 Pooled Genome Sequencing and Analysis

DNA extraction and sequencing

Genomic DNA was extracted from samples of 200 female flies collected from each of the 10 individual populations ($SCO_{1A,B}$-$SCO_{5A,B}$) using the Qiagen/Gentra Puregene kit, following the manufacturer's protocol for bulk DNA purification. The 10 gDNA pools were prepared as standard 290-300 bp fragment libraries using the Nextera Library Kit for Illumina sequencing, and constructed such that each population was given unique barcodes, normalized, and pooled together. Libraries were run across an Illumina MiSEQ at the genomics laboratory of Joint School for Nanoscience and Nanoengineering, administered by North Carolina A&T State University and the University of North Carolina, Greensboro. Resulting data were 200 bp paired-end reads on average. The raw fastq files containing DNA sequence data for the 10 SCO populations are available through the NCBI SRA repository (SRP159296, https://www.ncbi.nlm.nih.gov/sra/ SRP159296).

Read Mapping and SNP Identification

We mapped reads with BWA (version 0.7.8) (Li and Durbin 2009) against the *D. melanogaster* reference genome (version 6.14) using bwa mem with default settings. We filtered and sorted the resulting SAM files for reads mapped in proper pairs with a minimum mapping quality of 20 and converted them to the BAM using the view and sort commands in SAMtools (Li et al. 2009). Average coverage was above 40X or greater for all populations except SCO_{3a} and SCO_{5b} which were 29X and 26X respectively (see Supplementary Table S2, available online). Next, the steps were repeated using the raw fastq files for the C-type populations, $CO_{1\text{-}5}$ and $nCO_{1\text{-}5}$, used in our Graves et al. (2017) publication (files can be downloaded through NCBI SRA BioProject: PRJNA285429). This ultimately resulted in the creation of another 10 bam files. Bam files for all 20 populations were them combined into a single mpileup file using SAMtools. This mpileup file was then converted to a "synchronized" file using the PoPoolation2 software package (Kofler et al. 2011). This file displays allele counts for all positions in the reference genome for all populations in a succinct tab delimited format. RepeatMasker 4.0.3 (http://www.repeatmasker.org) was used to create a gff file containing simple sequence repeats found in the *D. melanogaster* genome version 6.14. These regions were then masked in the sync file.

SNP Variation

A SNP table was created using the sync file mentioned above. We only considered sites where coverage was between 15X and 200X, and for a site to be considered polymorphic we required a minimum minor allele frequency of 2% across all 20 populations. Based on these criteria, we identified ~1.19 million SNPs. All sites failing to meet these criteria were discarded. To assess broad patterns of SNP variation in the S-type and C-type populations, heterozygosity

was calculated and plotted over 150 kb non-overlapping windows directly from the major and minor counts in our SNP table. A t-test was also performed to compare mean heterozygosity between the two groups of populations. To assess how closely replicate populations resembled one another, F_{ST} estimates were also obtained using the formula: $F_{ST} = (H_T - H_S)/H_T$ where H_T is heterozygosity based on total population allele frequencies, and H_S is the average subpopulation heterozygosity in each of the replicate populations (Hedrick 2009). F_{ST} estimates were made at every polymorphic site in the data set for a given set of replicate populations.

SNP Differentiation and Gene ontology (GO) Analysis

The Cochran-Mantel-Haenzsel (CMH) test as implemented in the PoPoolation2 software package was used to compare SNP frequencies between the SCO and C-type populations. Comparisons were made at every site that met our previously described SNP calling criteria. When running the test, the $CO_{1\text{-}5}$ populations were paired with the $SCO_{1A\text{-}5A}$ populations with respect to ancestry (indicated by replicate number), while the SCO_{1B}-SCO_{5B} populations were paired with the $nCO_{1\text{-}5}$ populations. CMH tests between our two groups of populations were performed at each of these polymorphic sites. To correct for multiple comparisons, we used the permutation method featured in Graves et al. (2017). Briefly, populations were randomly assigned to one of two groups and the CMH test was then performed at each polymorphic position in the shuffled data set to generate null distributions of p-values. This was done 1000 times, and each time the smallest p-value generated was recorded. The quantile function in R was then used to define thresholds that define the genome-wide false-positive rate, per site, at 5%.

The Gowinda software package (Kofler and Schlotterer 2012) was used to identify enriched GO terms. Our complete SNP list based on our aforementioned SNP calling parameters was used as the background. A gene annotation file for the *D. melanogaster* reference genome

(6.14) was obtained from FlyBase. A gene set file for relevant GO terms was obtained from FuncAssociate3 (Berriz et al. 2009). With these files as inputs, Gowinda was set to run for 10^6 simulations with the gene-definition and mode parameters set to "gene". Setting gene-definition and mode to "gene" means we only considered SNPs within introns and exons associated with genes, and our analysis assumed all SNPs within a gene are linked. This list was then filtered so all GO categories containing less than 5 reference genes were discarded, and the resulting list was run through GO-Module to correct for hierarchical clustering (Yang et al. 2011.).

IV. Results

4.1 Starvation Resistance and Lipid Content

The S-type populations had a significantly higher average starvation survival time and average lipid content than the C-type populations (see Fig. 1 and Table 1; see Supplementary Table S1 for 20 population data, available online). First off, the average survival time during starvation, or starvation resistance, for the five SCO_A populations was 150.98 hours, and for the five SCO_B populations it was 152.2 hours. The starvation resistances of the SCO_A and SCO_B populations were statistically similar (p-value = 0.95). The average lipid content for the five SCO_A populations was 0.138 mg per fly, while the average lipid content for the five SCO_B populations at 14 days of age from egg was 0.143 mg per fly. The lipid contents between these two sets of S-type populations were also statistically similar (p-value = 0.754). The five CO populations had a much lower average starvation survival time of 66.84 hours, and an average lipid content of 0.058 mg per fly at 14 days of age. The recently-derived nCO populations had an average starvation survival time of 61.4 hours, and an average lipid content of 0.051 mg per fly. The CO and nCO populations were not significantly different in regards to starvation

resistance (p-value = 0.081) and lipid content (p-value = 0.367). After finding no significant differences within the two selection types, we compared all ten C-type populations against the ten S-type populations. The S-type populations had a significantly higher average lipid content than the C-type populations by 0.086 mg/fly (p-value < 0.0001). Approximately 8.6% of the S-type body mass is composed of lipid, whereas only 4.3% of the body mass of CO populations is composed of lipid. Additionally, the S-type populations had a significantly greater starvation survival time than the C-type populations, with a difference of 87.44 hours (p-value = 0.01).

In our first experiment looking at age-specific physiology, both the CO and SCO_A sets of populations increased in lipid content from age 14 to age 21 (see Fig. 2; Supplementary Table S3, available online). The five CO populations went from an average 0.068 mg/fly lipid content (5.56% lipid to mass) to 0.103 mg/fly in one week (8.67% lipid to mass; p-value = 0.019). However, the CO populations maintained a stable average lipid content (~ 0.10 mg/fly) from age 21 through age 49. At age 14 from egg, the SCO_A populations had an average lipid content of 0.129 mg/fly (8.57% lipid to mass), which significantly increased to 0.397 mg/fly (24.03% lipid to mass) by age 21 from egg (p-value < 0.0001). The average lipid content for the SCO_A populations continued to rise through age 35, reaching an average lipid content of 0.584 mg/fly (p-values < 0.0001). The lipid content to body mass of the S-type populations ranged from 29.55% to 32.75% during these later ages. The increase in lipid content to body mass ratio is coupled with a decrease in water content to body mass ratio in these populations (66.33% at age 14 days down to 50.57% at age 49 days). At all ages, the S-type populations had a significantly higher average lipid content than the C-type populations (p-value < 0.01).

4.2 Desiccation Resistance, Water Content, and Glycogen Content

Similar to starvation resistance and lipid content, we found significantly lower average water content and average glycogen contents in the C-type populations when compared to S-type populations (see Table 1; Supplementary Fig. S1 and Fig. S2, available online). The CO and nCO populations had similar average water contents of 0.919 mg/fly and 0.863 mg/fly (p-value = 0.114). The five SCO_A and five SCO_B populations had average water contents of 1.058 mg/fly and 1.094 mg/fly, respectively (p-value = 0.374). When we compared the 10 S-type populations with the 10 C-type populations, the average water content of the S-type populations was slightly significantly higher than the average water content of the C-type populations (p-value = 0.031). The average glycogen contents for the CO and nCO populations were 0.068 mg/fly and 0.076 mg/fly, respectively (p-value = 0.366). The average glycogen contents for the SCO_A and SCO_B populations were 0.151 mg/fly and 0.163 mg/fly, respectively (p-value = 0.473). When we compared the 10 C-type populations with the 10 S-type populations, we found that the S-type populations had a larger average glycogen content than the C-type populations by 0.084 mg/fly (p-value < 0.001, see Table 1).

With water content and glycogen content being major factors in desiccation resistance, it was not surprising to find parallel results for desiccation resistance among these populations (see Fig. S3, available online). The five CO and five nCO populations had average desiccation survival times of 14.71 hours and 14.63 hours, respectively. These two sets of C-type populations were not significantly different from each other (p-value: 0.935). The five SCO_A populations and the five SCO_B populations had average desiccation survival times of 17.29 hours and 16.34 hours, respectively. These two sets of S-type populations were not significantly different from each other (p-value = 0.278). When we compared the 10 S-type populations

against the 10 C-type populations, the S-type populations had a higher average desiccation survival time by 2.14 hours (see Table 1). However, our linear mixed effects analysis showed this difference was not statistically significant, though close to statistical significance (p-value = 0.0697).

4.3 Cardiac Arrest Rates

Our first set of cardiac experiments compared the C and S-type populations at age 14 days from egg. The S-type populations clearly had a dysfunctional cardiac system in comparison to the C-type populations (see Fig. 1 and Table 1). The five CO populations had an average arrest rate of 25.3%, whereas the five nCO populations had an average cardiac arrest rate of 27.3% (p-value = 0.643). The five SCO_A populations had an average arrest rate of 54%, whereas the five SCO_B populations had an average rate of 52.7% (p-value = 0.806). With no significant differences found within the two selection treatments, we then compared the cardiac arrest rates of the 10 S-type populations against the 10 C-type populations. The 10 C-type populations had an average cardiac arrest rate of 26.33%, whereas the 10 S-type populations had an average cardiac arrest rate of 53.33%. The average cardiac arrest rate of the S-type populations was significantly higher than the C-type populations (p-value = 2.37×10^{-21}).

In our second set of experiments, at ages 14, 21, and 28 days from egg, the five SCO_A populations had an average rate of cardiac arrests significantly higher than that of the five CO populations (p-value < 0.05; see Fig. 3; Supplementary Table S3, available online). It was not until age 35 days from egg where the cardiac arrest rates in the SCO_A populations (57.69%) were no longer significantly different than the CO populations (59.4%; p-value = 0.826). The lack of a significant difference in arrest rates was also observed at age 42 and age 49 from egg.

4.4 Age-Specific Mortality Rates

The age-specific mortality rates of the C-type populations and the S-type populations significantly differ starting at age 20 days from egg, continuing through age 40 days from egg (see Fig. 4). The mortality rates of the S-type populations dramatically diverge from the pre-aging plateau shown by the C-type populations. This rapid increase in S-type mortality rates was seen from age 14 to 25 days from egg and was then followed by a roughly stable mortality through age 40 days from egg. The mortality rates of the C-type populations converged on the rates of the S-type populations during the 41-43 day age-interval (p-value = 0.129). The mortality rates of the 10 S-type populations mirrored those of the 10 C-type populations for the remainder of the assay. After the aging phase ends, the twenty populations begin to show signs of a late-life plateau that has been documented before in our outbred *Drosophila* populations (e.g. Rose et al. 2002; Burke et al. 2016). Overall, the average lifespan of the S-type populations was 41.2 days from egg, about eight days fewer than the average C-type lifespan (49.05 days; see Table 1).

4.5 SNP Variation

Although there were some notable depressions in heterozygosity in both the SCO and C-type populations, we see relatively few regions where genetic variation had been completely expunged (see Fig. 5). These patterns were also robust to changes in window size (see Supplementary Fig. S4 and Fig. S5, available online). Many of the observed depressions in heterozygosity were consistent across replicate populations, which may be indicative of selection acting on the genome in parallel across replicates. Regions featuring depressed heterozygosity were also more common in the SCO populations than the C-type populations, which is reflected

in differences in mean heterozygosity across the genome between the two groups. We found that mean heterozygosity in the C-type populations ranges from 0.27 to 0.29, and 0.22 to 0.25 in the SCO populations (see Supplementary Table S4, available online). Based on a t-test comparing means between the two groups, we found that heterozygosity was significantly higher in the C-type populations (p-value = 3.38×10^{-10}). Lastly, in the C-type populations, we found the mean F_{ST} was 0.04 which indicates a high degree of similarity between replicates. However, in the SCO populations we observed a mean F_{ST} of 0.11, indicating reduced similarity between replicates when compared to the C-type populations.

4.6 SNP Differentiation and GO Analysis

Our CMH tests identified a total of 1,234 significantly differentiated SNPs across the major chromosome arms (see Fig. 6). Checking for overlap between candidate SNPs and known genetic regions via FlyBase genome browser, we found a total of 145 genes (see Supplementary Table S5, available online). Seventy-four of these genes are located on chromosome arm 3R, thirty-two are located chromosome arm 2R, twenty-three are located on chromosome X, thirteen are located on chromosome 2L, and the remaining three genes are located on chromosome 3L. The associated GO functions of these genes vary from cellular and organelle function to DNA and RNA expression elements. Many of these *Drosophila* genes have been linked to human orthologs associated with cardiovascular, renal, and metabolic disorders.

Subsequently, using the list of candidate SNPs and Gowinda, a tool that allows for analysis of gene set enrichment, we identified 53 significantly enriched GO terms with p-values < 0.05. (Kofler and Schlotterer 2012). Discarding terms with GO categories containing less than 5 genes and correcting for hierarchical clustering using GO-Module (Yang et al. 2011) reduced

this list to 34 enriched terms (see Table 1; see Supplementary Table S6 for associated genes, available online). Although these terms cover a variety of biological processes, we found that roughly a third of them were related to metabolic processes.

V. Discussion

5.1 Extreme Differentiation Between C-type and S-type Populations

With only 67 generations of intense selection for increased starvation resistance, the S-type populations have clearly differentiated from their ancestors in many of the physiological assays mentioned above. As seen in other work from our laboratory and our colleagues' laboratories, it was not surprising to find such extreme differentiation in starvation resistance and lipid content between the C-type and S-type populations (Rose et al. 1992; Hardy et al. 2015). The improved survival in a starvation environment is likely due to the changes in lipid content. In the S-type populations, lipid composes 8.6% of their average wet body mass (~1.64 mg/fly), whereas the lipid content in the C-type populations only contributes to 4.3% of their wet body mass (~1.263 mg/fly; see Table 1). Our results parallel what has been previously documented, that starvation resistance and lipid content tend to be highly correlated with each other (R^2 = 0.86; Djawdan et al. 1998; Hardy et al. 2015). Not only did we find a breakdown in the correlation between starvation resistance and average lifespan, we actually observed a negative correlation (R^2 = -0.73) between these traits, similarly described as in Phelan et al. (2003) with their intensely-selected starvation resistant SO populations. Although beneficial for a select few physiological characters, this obesity-driving selection regime is quite antagonistic on overall health.

Implementing this intense selection protocol also indirectly affected other physiological characters. Directly selecting for increased starvation resistance also increased water content, glycogen content, and desiccation resistance in the S-type populations (see Table 1, Supplementary Table S1 and Figures S1-S3). The average absolute water contents of the S-type population were significantly larger than the C-type populations, however in relation to the wet body mass, the ratio of water content to wet body mass actually decreased in the S-type populations (70.54% in C-type versus 65.61% in S-type populations). Using the wet body masses gathered from the water content and lipid content assays, the ratio of glycogen content to body mass nearly doubled in the S-type populations (5.7% in C-type versus 9.56% in S-type populations). Although insignificant, it is not surprising to observe an increased desiccation resistance in S-type populations following the observation of larger amounts of free water and glycogen content in these populations. It must be kept in consideration that the overall surface area to volume ratio of the S-type populations is potentially smaller, thus decreasing water loss rates. This would further contribute to the observed increased desiccation resistance. Desiccation resistance typically responds slower to selection than starvation resistance does (Passananti et al. 2004; Archer et al. 2007). It is likely that the continuation of our S-type selection regime would inevitably lead to a statistically significant increase in desiccation resistance. We found a strong, positive correlation between glycogen content and desiccation resistance within these 20 populations ($R^2 = 0.702$), as well as between water content and desiccation resistance ($R^2 = 0.632$).

Our results suggest that the increased lipid levels due to increased starvation resistance and subsequent alterations in lipid homeostasis weaken cardiac function, specifically the ability to recover from a fibrillation-like event. The stark contrast in rate of cardiac arrests is not hard to

comprehend, especially when you consider the nearly doubled lipid content to body mass ratio between the S-type and C-type populations. Lipid content increased in S-type populations through age 35 days, before stabilizing. A similar trend was seen in the cardiac arrest rates, with the exception at age 28 days in which the rate of cardiac arrests in both C-type and S-type populations was lower than at age 21 days. We found a strong, positive correlation between lipid content and cardiac arrest rates within these 20 populations ($R^2 = 0.789$). The S-type populations appear to display similar cardiac dysfunction as the experimental populations in studies previously referenced.

The cardiological differences may not be the sole factor behind the heightened mortality rates in the starvation-selected populations during early adulthood, but it is likely a major contributing factor. The association between the early age cardiac dysfunction and heightened early age mortality rates in the starvation-selected populations (see Figs. 3 and 4) suggests that the detrimental effects of obesity on health and lifespan could be age-dependent. On average, the lipid content to body mass ratio in S-type flies increased from 8.6% to 24.02% in just one week of adulthood. This first week of adulthood in S-type populations is also associated with the steep rise in mortality rates, as well as cardiac arrest rates (46% to 55.14%). While the mortality rate difference eventually goes away at later adult ages, it is also often found that the effect of obesity on human mortality rates shows a similar age dependence. Specifically, the optimal BMI levels for human mortality appear to increase with adult age (Flegel et al. 2013; Ford et al. 2014; Jackson et al. 2014). Even though the cardiac arrest rates remain rather high (55-60%) at later ages in the S-type populations, they do not significantly rise post age 35 day from egg. Additionally, the cardiac arrest rates are similar to the C-type populations, implying that the obesity of the S-type populations may not have as large of detrimental effect on *Drosophila*

physiology and longevity at these later ages in comparison to the earlier ages. Cardiac arrest rates in the C-type populations rose to comparable levels of the rates observed in the S-type populations, despite the lipid content to body mass ratio never surpassing 8.67%. Therefore, a potentially complex age-dependent relationship between lipid levels, cardiac function, and mortality rates is seen in obese fruit flies, echoing some features of the relationship between obesity and mortality in human populations.

5.2 Drosophila as a Model for Heart and Obesity-Related Disorders

Experimentally probing *D. melanogaster* physiology via selection experiments has demonstrated the value of this complex metazoan as an experimental model for the study of health and disease. Unlike most microbial models, the physiology of *Drosophila* is both complex and broadly analogous to some features of vertebrate physiology. *D. melanogaster* also have homologous genetic mechanisms with those that are thought to determine lifespan among vertebrates, including such genetic systems as TOR and insulin/insulin-like signaling (e.g. Partridge and Gems 2007). There are important differences between *Drosophila* and mammalian cardiovascular structure and function (e.g. open versus closed circulatory systems). However, there are similarities in (a) early heart development, (b) age-dependent decline in heart function, and (c) the genes associated with heart development, function, and diseases (Bodmer and Venkatesh 1998; Cripps and Olson 2002; Zaffran and Frasch 2002; Bier and Bodmer 2004; Ocorr et al. 2006; Nishimura et al. 2011; Diop and Bodmer 2012;). Additionally, *Drosophila* shares some of the genes that underlie its cardiac performance with those of human cardiac genetics, such as *tinman* (Bodmer 1993; Bodmer 2006) and *opa1* (Shahrestani et al. 2009).

Starvation-resistant and obese *Drosophila* populations are a powerful tool not only for studying the evolution of starvation responses, but also for studying metabolic disorders and related cardiac dysfunction (Birse et al. 2010; Diop and Bodmer 2012; Trinh and Boulianne 2013; Smith et al. 2014; Diop and Bodmer 2015; Hardy et al. 2015; Hardy et al. 2018). In a similar experimental setup as ours, Hardy et al. (2015) observed dilated hearts and reduced contractility in their three evolved obese populations after 65 generations of selection for starvation resistance. With comparable extreme phenotypic differentiation in our obese populations to that of the obese populations in Hardy et al. (2015), 14 of our 145 significantly altered genes matched the significant genes detected in Hardy et al. (2018), notwithstanding the much lower statistical power of the Hardy et al. study, due to its lower level of replication. These 14 genes entail a range of associated molecular functions and biological processes, from cholesterol transmembrane transport to muscle development and cytoskeleton organization. Observing 14 overlapping genes between our respective sets of obese *Drosophila* populations instills confidence in our experimental model system and the possibility to parse out underlying genetic mechanisms from our 145 gene set (see Supplementary Table S5, available online).

While it is difficult to see how all of our significantly differentiated genes could relate to this selection regime, especially with respect to mass lipid accumulation, we can conjecture why some genes have been affected by selection. For example, the gene *Gdh* (glutamate dehydrogenase) has been associated with glutamate metabolism and NADH oxidation. Glutamate dehydrogenase serves a major role in regulating amino acid-induced insulin secretion. Mutations in this gene lead to hyperinsulinemia, hypoglycemia, and hyperinsulinism-hyperammonemia syndrome (Stanley 2009; Stanley 2011; Smith and Smith 2016). These disorders have been linked with insulin resistance, diabetes, and obesity (Corkey 2012;

Templeman et al. 2017; Page and Johnson 2018). The differentiated gene *nau*, which encodes the protein nautilus, belongs to the basic helix-loop-helix ("bHLH") family of transcription factors. These bHLH transcription factors play an important role in muscle myogenesis, especially in the formation of the *Drosophila* visceral and heart muscles. Anomalously, there are no differentiated genes that have been specifically associated with lipid metabolism, at least not previously. Further exploration of the 145 identified genes and their potential involvement in lipid homeostasis and cardiac dysfunction is a necessary step to advance our current knowledge of obesity-related disorders.

Through intense selection for starvation resistance, we have created a set of ten outbred *Drosophila* populations that have dramatic increases in metabolic reserves and survival in stressful environments, but overall, experience a shorter lifespan by eight days. These visibly obese populations exhibit a decline in heart robustness in relation to their ancestral CO populations, a decline that is associated with increased mortality rates. As mentioned, we cannot attribute a specific cause behind the heightened mortality rates in our S-type populations. It could involve multiple types of organ dysfunction, such as cardiac, gastrointestinal and renal dysfunction. A weakened or stressed heart could lead to poor contractility, not allowing proper distribution of nutrients. Additionally, it could contribute to a disruption in adequate recycling and filtration via the pericardial nephrocytes. Different opinions have formed as to whether the heart frailty is due to the physical impingement of the fat body and lipid droplets on the heart tube, or if it is the increased viscosity of hemolymph due to increase lipids. While previous studies have shown that altered lipid homeostasis in obese flies is associated with disruption in physical activity, metabolic rates, sleep patterns and heart function, this study supports the idea of using a fly model to the explore the potential age-dependent relationship between obesity,

heart function, and longevity (Birse et al. 2010; Diop and Bodmer 2015; Hardy et al. 2015; Hardy et al. 2018).

Fruit flies with increased lipid content in other studies display hyperglycemia, insulin resistance, reduced cardiac contractility, and cardiac arrhythmias (Birse et al. 2010; Hoffmann et al. 2013; Na et al. 2013; Trinh and Boulianne 2013). These comparable characteristics to those of mammalian metabolic syndrome support the value of using *Drosophila* for heart studies. The types of heart disease that are prevalent among present-day human populations are unlikely to result from deleterious alleles of major effect, instead involving many genes and thus many biochemical pathways. It is therefore important for our understanding of human heart disease to study heart function in these large outbred populations of *D. melanogaster* that differ in allele frequencies at many loci, as shown here in our study. The use of multi-omic tools with these populations across multiple ages could help parse the genetic and molecular underpinnings of heart function, heart disease, and other obesity-related disorders.

VI. ACKNOWLEDGEMENTS

We thank Dr. James W. Hicks for the helpful discussions and comments on the experiments. We would also like to thank Dr. Donovan German, Dr. Michael Mulligan, Dr. Timothy Bradley and their graduate students for their assistance and sharing of equipment used for measuring metabolic reserves contents and extracting genomic content. We are also grateful to the many undergraduate research students who contributed to the stock maintenance and experimental assays conducted in Dr. Michael Rose's, Dr. Laurence Mueller's, and Dr. Timothy Bradley's laboratories.

VII. Literature Cited

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

Archer M.A., J.P. Phelan, K.A. Beckman, and M.R. Rose. 2003. Breakdown in correlations during laboratory evolution. II. Selection on stress resistance in *Drosophila* populations. Evolution 57: 536-543.

Bennett A.F., K.M. Dao, and R.E. Lenski. 1990. Rapid evolution in response to high temperature selection. Nature 346: 79-81

Berriz G.F., J.E. Beaver, C. Cenik, M. Tasan and F.P. Roth. 2009. Next generation software for functional trend analysis. Bioinformatics 25 (22):3043-3044.

Bier E. and R. Bodmer. 2004. *Drosophila,* an emerging model for cardiac disease. Gene 342: 1-11.

Birse R.T., J. Choi, K. Reardon, J. Rodriguez, S. Graham, S. Diop, K. Ocorr, R. Bodmer, and S. Oldham. 2010. High-fat-diet-induced obesity and heart dysfunction are regulated by the TOR Pathway in *Drosophila*. Cell Metab 12: 533–544.

Bodmer R. 1993. The gene *tinman* is required for specification of the heart and visceral muscles in *Drosophila*. Development 118: 719-729.

Bodmer R. 2006. Development of the Cardiac Musculature. In: Madame Curie Bioscience Database [Internet]. Texas: Landes Bioscience.

Bodmer R. and T.V. Venkatesh. 1998. Heart development in *Drosophila* and vertebrates: conservation of molecular mechanisms. Dev Genet 22: 181-186.

Bradley T.J., A.E. Williams, and M.R. Rose. 1999. Physiological responses to selection for desiccation resistance in *Drosophila melanogaster.* Amer Zool 39: 337-345.

Burke M.K. and M.R. Rose. 2009. Experimental evolution with *Drosophila*. Am J Physiol – Reg I 296: R1847-R1854.

Burke M.K., T.T. Barter, L.G. Cabral, J.N. Kezos, M.A. Phillips, G.A. Rutledge, K.H. Phung, R. H. Chen, H.D. Nguyen, L.D. Mueller, and M.R. Rose. 2016. Rapid divergence and convergence of life-history in experimentally evolved *Drosophila melanogaster.* Evolution 70 (9): 2085-2098. DOI10.1111/evo.13006.

Chen W. and J.F. Hillyer. 2013. FlyNap (trimethylamine) increases the heart rate of mosquitoes and eliminates the cardioacceleratory effect of neuropeptide CCAP. PLoS ONE 8(7): e70414.

Chippindale A.K., T.J.F. Chu, and M.R. Rose. 1996. Complex trade-offs and the evolution of starvation resistance in *Drosophila melanogaster*. Evolution 50: 753-766.

Chippindale A.K., A.G. Gibbs, M. Sheik, K.J. Yee, M. Djawdan, T.J. Bradley, and M.R. Rose. 1998. Resource acquisition and the evolution of stress resistance in *Drosophila melanogaster*. Evolution 52: 1342-1352.

Cooper V.S., D. Schneider, M. Blot, and R.E. Lenski. 2001. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. J Bacteriol 183: 2834-2841.

Corkey B.E. 2012. Banting lecture 2011: Hyperinsulinemia: Cause or consequence? Diabetes 61: 4-13.

Cripps R.M. and E.N. Olson. 2002. Control of cardiac development by an evolutionarily conserved transcriptional network. Dev Biol 246: 14-28.

Diop S.B. and R. Bodmer. 2012. *Drosophila* as a model to study the genetic mechanisms of obesity-associated heart dysfunction. J Cell Mol Med 16: 966-971.

Diop S.B. and R. Bodmer. 2015. Gaining insights into diabetic cardiomyopathy from *Drosophila*. Trends Endocrin Met 26 (11): 618-627.

Djawdan M., A.K. Chippindale, M.R. Rose, and T.J. Bradley. 1998. Metabolic reserves and evolved stress resistance in *Drosophila melanogaster.* Physiol Zool 71: 584-594.

Flegel K.M., B.K. Kit, H. Orpana, and B.I. Graubard. 2013. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. JAMA 309 (1): 71-82.

Ford D.W., T.J. Hartman, C. Still, C. Wood, D.C. Mitchell, P. Erickson, R. Bailey, H. Smiciklas-Wright, D.L. Coffman, and G.L. Jensen 2014. Body Mass Index, Poor Diet Quality, and Health-Related Quality of Life Are Associated With Mortality in Rural Older Adults. J Nutr Gerontol Geriatr 33 (1): 23-34.

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

Graves J.L. and M.R. Rose. 1990. Flight duration in *Drosophila melanogaster* selected for postponed senescence. In: Genetic Effects on Aging, II, edited by Harrison, D. West Caldwell: Telford Press, 57-63.

Graves J.L., S. Luckinbill, and A. Nichols. 1988. Flight duration and wing beat frequency in long and short lived *Drosophila melanogaster*. J Insect Physiol 34: 1021-1026.

Graves J.L., E. Toolson, C.M. Jeong, L.N. Vu, and M.R. Rose. 1992. Desiccation resistance, flight duration, glycogen and postponed senescence in *Drosophila melanogaster*. Physiol Zool 65: 268-286.

Graves J.L., K.L. Hertweck, M.A. Phillips, M.V. Han, L.G. Cabral, T.T. Barter, L.F. Greer, M.K. Burke, L.D. Mueller, and M.R. Rose. 2017. Genomics of parallel experimental evolution in *Drosophila*. Mol Biol Evol 34 (4): 831-842.

Hardy C.M., R.T. Birse, M.J. Wolf, L. Yu, R. Bodmer, and A.G. Gibbs. 2015. Obesity-associated cardiac dysfunction in starvation-selected *Drosophila melanogaster*. Am J Physiol – Reg I 309 (6): R658-R667. doi: 10.1152/ajpregu.00160.2015.

Hardy C.M., M.K. Burke, L.J. Everett, M.V. Han, K.M. Lantz and A.G. Gibbs. 2018. Genome-wide analysis of starvation-selected *Drosophila melanogaster* – A genetic model of obesity. Mol Biol Evol 35 (1): 50-65.

Harshman L.G., A.A. Hoffmann, and A.G Clark. 1999. Selection for starvation resistance in *Drosophila melanogaster*: physiological correlates, enzyme activities and multiple stress responses. J Evol Biol 12: 370-379.

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

Hoffmann J., R. Romey, C. Fink, and T. Roeder. 2013. *Drosophila* as a model to study metabolic disorders. Adv Biochem Eng Biot 135: 41-61.

Jackson C.L., H.C. Yeh, M. Szkio, F.B. Hu, N.Y. Wang, R. Dray-Spira, and F.L. Brancati. 2014. Body-mass index and all-cause mortality in US adults with and without diabetes. J Gen Intern Med 29 (1): 25-33

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

Kofler R. and C. Schlötterer. 2012. Gowinda: unbiased analysis of gene set enrichment for genome-wide association studies. Bioinformatics 28 (15): 2084–2085.

Li H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14): 1754-1760.

Li H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25 (16): 2078-2079.

Mongold J.A., A.F. Bennett, and R.E. Lenski. 1999. Evolutionary adaptation to temperature (VII): Extension of the upper thermal limit of *Escherichia coli.* Evolution 53 (2): 386-394.

Na J., L.P. Musselman, J. Pendse, T.J. Baranski, R. Bodmer, K. Ocorr, and R. Cagan. 2013. A *Drosophila* model of high sugar diet-induced cardiomyopathy. PLoS Genet 9: e1003175.

Nishimura M., K. Ocorr, R. Bodmer, and J. Cartry. 2011. *Drosophila* as a model to study cardiac aging. Exp Gerontol 46: 326-330.

Ocorr K., T. Akasaka, and R. Bodmer. 2006. Age-related cardiac disease model of *Drosophila*. Mech Ageing Dev 128: 112-116.

Page M.M. and J.D. Johnson. 2018. Mild suppression of hyperinsulinemia to treat obesity and insulin resistance. Trends Endocrin Met 29 (6): 389-399.

Passananti H.B., K.A. Beckman, and M.R. Rose. 2004. Relaxed stress selection in *Drosophila melanogaster*. In: Methuselah Flies: A Case Study in the Evolution of Aging, edited by Rose, M.R., Passananti, H.B., and Matos, M. Singapore: World Scientific Publishing, 323-352.

Partridge L. and D. Gems. 2007. Benchmarks for ageing studies. Nature 450: 165-167.

Paternostro G., C. Vignola, D.-U. Bartsch, J.H. Omens, A.D. McCulloch, and J.C. Reed. 2001. Age-associated cardiac dysfunction in *Drosophila melanogaster*. Circ Res 88: 1053-1058.

Phelan J. P., M.A. Archer, K.A. Beckman, A.K. Chippindale, T.J. Nusbaum, and M.R. Rose. 2003. Breakdown in correlations during laboratory evolution. I. Comparative analyses of *Drosophila* populations. Evolution 57: 527-535.

R Core Team. 2015. R: A Language and Environment For Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

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

Rose M.R., J.L. Graves, and E.W. Hutchinson. 1990. The use of selection to probe patterns of peliotrophy in fitness characters. In: Insect Life Cycles, edited by Gilbert, F. New York: Spring-Verlag, 29-42.

Rose M.R., L.N. Vu, S.U. Park, and J.L. Graves. 1992. Selection for stress resistance increases longevity in *Drosophila melanogaster*. Exp Gerontol 27: 241-250.

Rose M.R., M.D. Drapeau, P.G. Yazdi, K.H. Shah, D.B. Moise, R.R. Thakar, C.L. Rauser, and L.D. Mueller. 2002. Evolution of late-life mortality in *Drosophila melanogaster*. Evolution 56: 1982-1991.

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

Service P.M., E.W. Hutchinson, M.D. MacKinley, and M.R. Rose. 1985. Resistance to environmental stress in *Drosophila melanogaster* selected for postponed senescence. Physiol Zool 58: 380-389.

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

Shahrestani P., H. Leung, P.K. Le, W.L. Pak, S. Tse, K. Ocorr, and T. Huang. 2009. Heterozygous mutation of *Drosophila* Opa1 causes the development of multiple organ abnormalities in an age-dependent and organ-specific manner. PLoS One 4: e6867.7.

Smith W.W., J. Thomas, J. Liu, T. Li, and T.H. Moran. 2014. From fat fruit fly to human obesity. Physiol Behav 136: 15-21.

Smith H.Q. and T.J. Smith. 2016. Identification of a novel activator of mammalian glutamate dehydrogenase. Biochemistry 55: 6568-6576.

Stanley C.A. 2009. Regulation of glutamate metabolism and insulin secretion by glutamate dehydrogenase in hypoglycemic children. Am J Clin Nutr 90: 862S-866S.

Stanley C.A. 2011. Two genetic forms of hyperinsulinemic hypoglycemia caused by dysregulation of glutamate dehydrogenase. Neurochem Int 59 (4): 465-472.

Templeman N.M., S. Skovsø, M.M. Page, G.E. Lim, and J.D. Johnson. 2017. A causal role for hyperinsulinemia in obesity. J Endocrin 232: R173-R183.

Tenaillon O., A. Rodriguez-Verdugo, R.L. Gaut, P. McDonald, A.F. Bennet, A.D. Long, and B.S. Gaut. 2012. The molecular diversity of adaptive convergence. Science 335: 457-461.

Trinh I. and G.L. Boulianne. 2013. Modeling obesity and its associated disorders in *Drosophila*. Physiology 28: 117-124.

Wessells R.J. and R. Bodmer. 2004. Screening assays for heart function mutants in *Drosophila.* BioTechniques 37: 58-66.

Yang X., J. Li, Y. Lee and Y.A. Lussier. 2011. GO-*Module*: Functional synthesis and improved interpretation of Gene Ontology patterns. Bioinformatics 27 (10): 1444-1446.

Zaffran S. and M. Frasch. 2002. Early signals in cardiac development. Circ Res 91: 457-469.

VIII. TABLES

Table 1. Phenotypic Characterization of C-type and S-type Populations

Table 1
Phenotypic Characterization of C-type and S-type Populations

	C-type Average	S-type Average	P-value
Cardiac Arrest Rate (%)	26.33 (±3.5)	53.33 (±2.9)	2.37×10^{-21}
Lifespan (Days)	49.05 (±0.99)	41.2 (±1.16)	-
Survival during Starvation (h)	64.14 (±1.23)	151.67 (±4.23)	0.01
Survival during Desiccation (h)	14.67 (±0.24)	16.82 (±0.3)	0.0697
Lipid Content (mg/fly)	0.055 (±0.003)	0.141 (±0.006)	< 0.0001
Glycogen Content (mg/fly)	0.072 (±0.005)	0.157 (±0.009)	< 0.001
Water Content (mg/fly)	0.891 (±0.01)	1.076 (±0.012)	0.031
Body Mass (mg/fly)	1.263 (±0.032)	1.643 (±0.035)	< 0.001

Note – All data was collected at age 14 days from egg with the exception of lifespan. Data represents phenotypic characterization of female flies from the 10 C-type populations and 10 S-type populations.

Table 2. Enriched Gene Ontology Terms

GO ID	P-VALUE	GO TERMS
GO:0044765	0.000493	single-organism transport
GO:0051234	0.001618	establishment of localization
GO:0006810	0.003443	transport
GO:0090304	0.003974	nucleic acid metabolic process
GO:0016070	0.004889	RNA metabolic process
GO:0044237	0.006299	cellular metabolic process
GO:0035639	0.008713	purine ribonucleoside triphosphate binding
GO:0032555	0.008765	purine ribonucleotide binding
GO:0032550	0.008765	purine ribonucleoside binding
GO:0044248	0.011031	cellular catabolic process
GO:0042221	0.013156	response to chemical
GO:0034641	0.014051	cellular nitrogen compound metabolic process
GO:0032940	0.015651	secretion by cell
GO:0046483	0.016596	heterocycle metabolic process
GO:0042742	0.018441	defense response to bacterium
GO:0046903	0.018929	secretion
GO:0006139	0.019267	nucleobase-containing compound metabolic process
GO:0008324	0.019351	cation transmembrane transporter activity
GO:0019752	0.019677	carboxylic acid metabolic process
GO:0016192	0.019858	vesicle-mediated transport
GO:0005524	0.021219	ATP binding
GO:0032559	0.021378	adenyl ribonucleotide binding
GO:0051649	0.02157	establishment of localization in cell
GO:0044265	0.021589	cellular macromolecule catabolic process
GO:0006725	0.023259	cellular aromatic compound metabolic process
GO:0022891	0.033601	substrate-specific transmembrane transporter activity
GO:0044260	0.03366	cellular macromolecule metabolic process
GO:0006909	0.0347	phagocytosis
GO:0032991	0.035485	macromolecular complex
GO:0005737	0.035698	cytoplasm
GO:0045216	0.038275	cell-cell junction organization
GO:0006082	0.044803	organic acid metabolic process
GO:0043436	0.044803	oxoacid metabolic process
GO:0017111	0.048551	nucleoside-triphosphatase activity

IX. Figure Legends

Figure 1. Physiological differentiation between the ten intensely selected starvation resistant S-type populations (SCO_A and SCO_B) and the ten control C-type populations (CO and nCO) at age 14 days from egg. (TOP) The average rate of cardiac arrests after electrical pacing of female fruit flies (mean ± 1 SEM). The S-type populations had higher average rates of cardiac arrests than the C-type populations (p-value = 2.37×10^{-21}). (MIDDLE) The average survival time of female fruit flies in a starvation environment (mean ± 1 SEM). S-type populations had significantly longer survival times than the C-type populations (p-value = 0.01). (BOTTOM) The average lipid content of female fruit flies (mean ± 1 SEM). S-type populations had significantly larger average lipid contents than the C-type populations (p-value < 0.0001).

Figure 2. The age-specific average lipid contents of female fruit flies from the five CO populations and five SCO-a populations (mean ± 1 SEM). The SCO populations had statistically significantly, larger average lipid contents than the CO populations at all six ages.

Figure 3. The age-specific average rate of cardiac arrests of female fruit flies from the five CO and five SCO-a populations (mean ± 1 SEM). The SCO populations had significantly higher average rates of cardiac arrests than the CO populations at ages 14, 21, and 28 days from egg. At ages 35, 42, and 49 days from egg, the average rate of cardiac arrests are no longer statistically significantly different.

Figure 4. (A) The age-specific mortality rates for each of ten C-type and ten S-type populations. (B) The average age-specific mortality rate of the C-type populations and the S-type populations. The S-type populations have a significantly higher age-specific mortality rate than the C-type populations from ages 20 to 40 days from egg (p-values < 0.05). The mortality rates of the C-type populations converge on the rates of the S-type populations during the 41-43 day age-interval (p-value = 0.129).

Figure 5. Heterozygosity in the CO *(A),* nCO *(B),* SCO_A *(C),* and SCO_B (*D)* populations plotted over 150-kb windows across all major chromosome arms. All replicates are shown for each population.

Figure 6. Results from CMH tests comparing SNP frequencies between the SCO and C-type populations plotted as –log(p-values). A total of 1,234 significantly differentiated SNPs across the major chromosome arms were detected, with 145 genes identified within the significantly differentiated regions. Permutation derived significance threshold is shown in red.

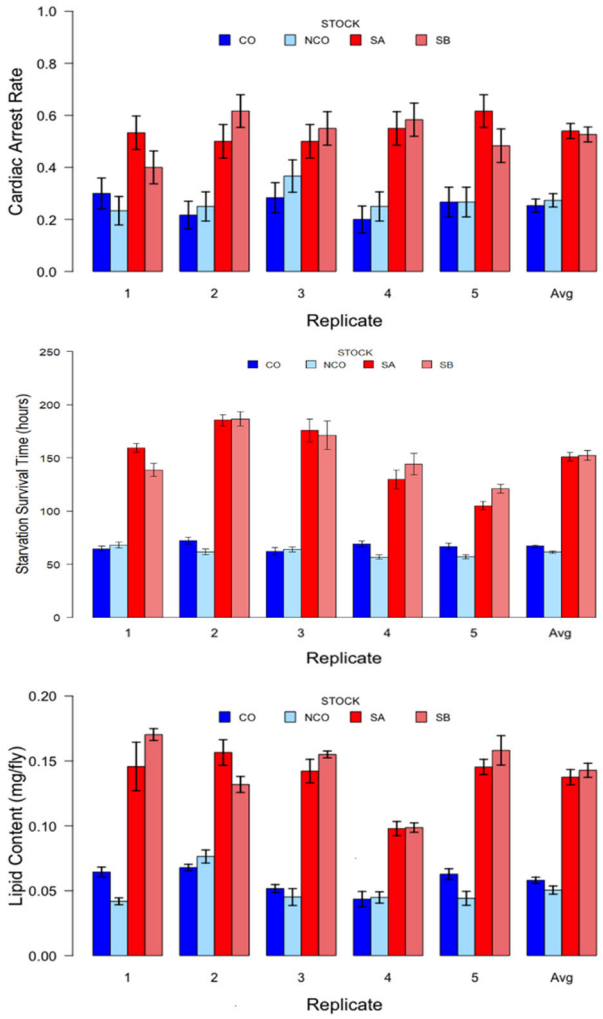
STOCK
CO
NCO
SA
SB
Cardiac Arrest Rate
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
Avg
Replicate
Starvation Survival Time (hours)
250
200
150
100
50
0
Lipid Content (mg/fly)
0.20
0.15
0.10
0.05
0.00

Figure 1

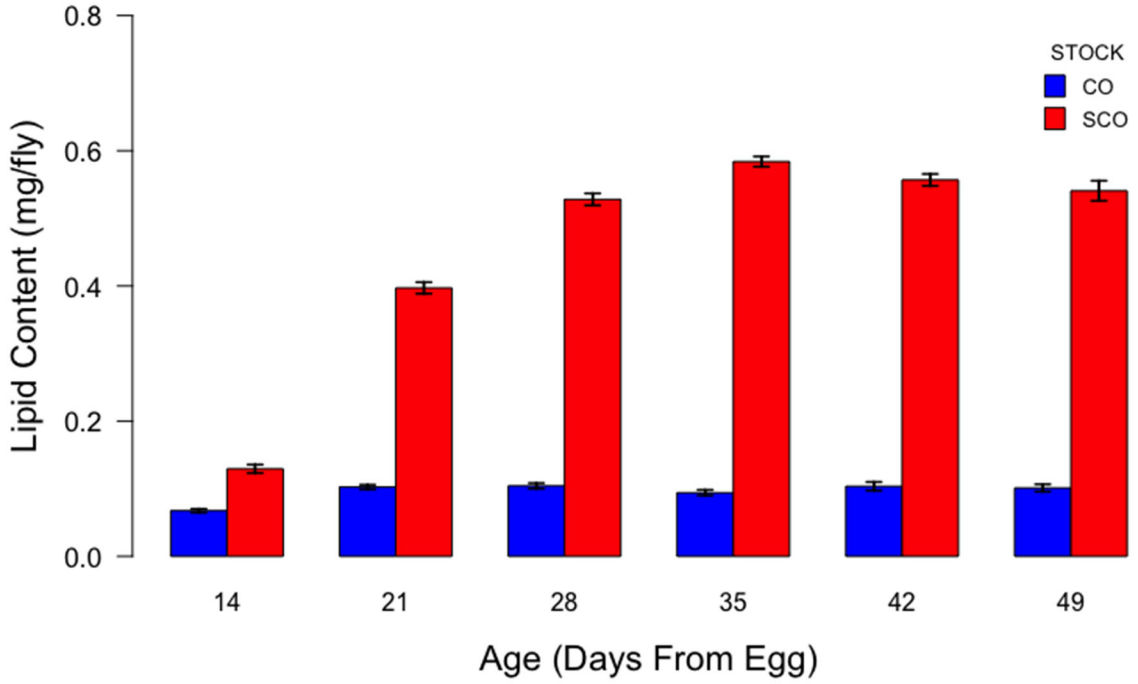
0.8
0.6
0.4
0.2
0.0
Lipid Content (mg/fly)
STOCK
CO
SCO
14
21
28
35
42
49
Age (Days From Egg)

Figure 2.

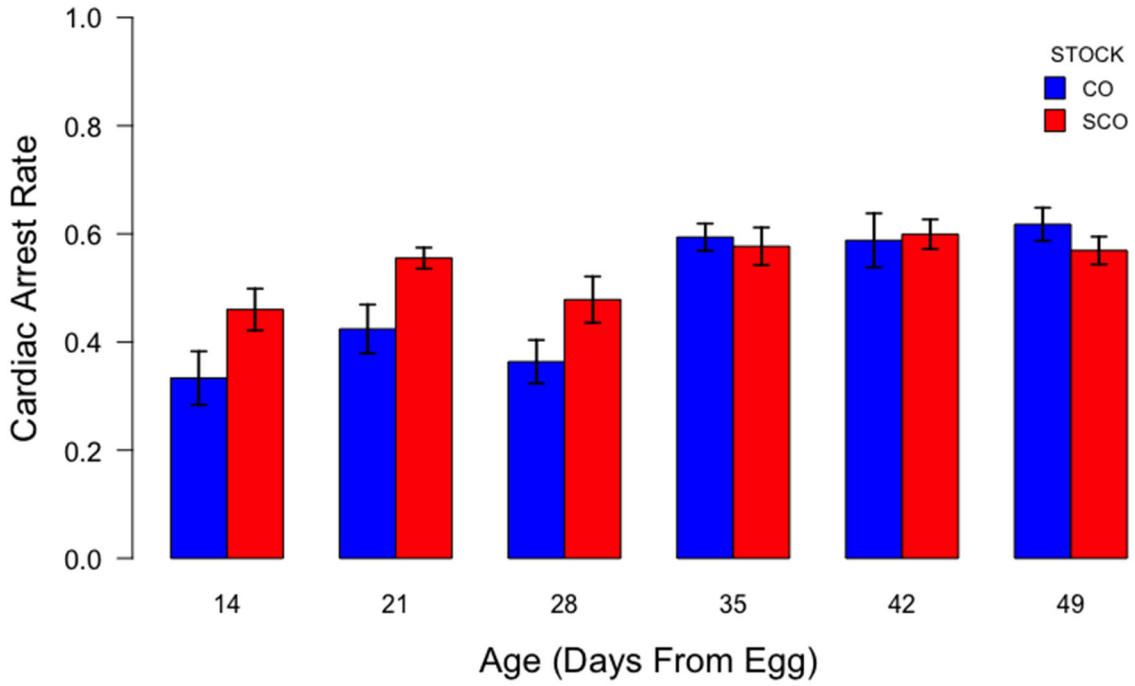
1.0
0.8
0.6
0.4
0.2
0.0
Cardiac Arrest Rate
STOCK
CO
SCO
14
21
28
35
42
49
Age (Days From Egg)

Figure 3.

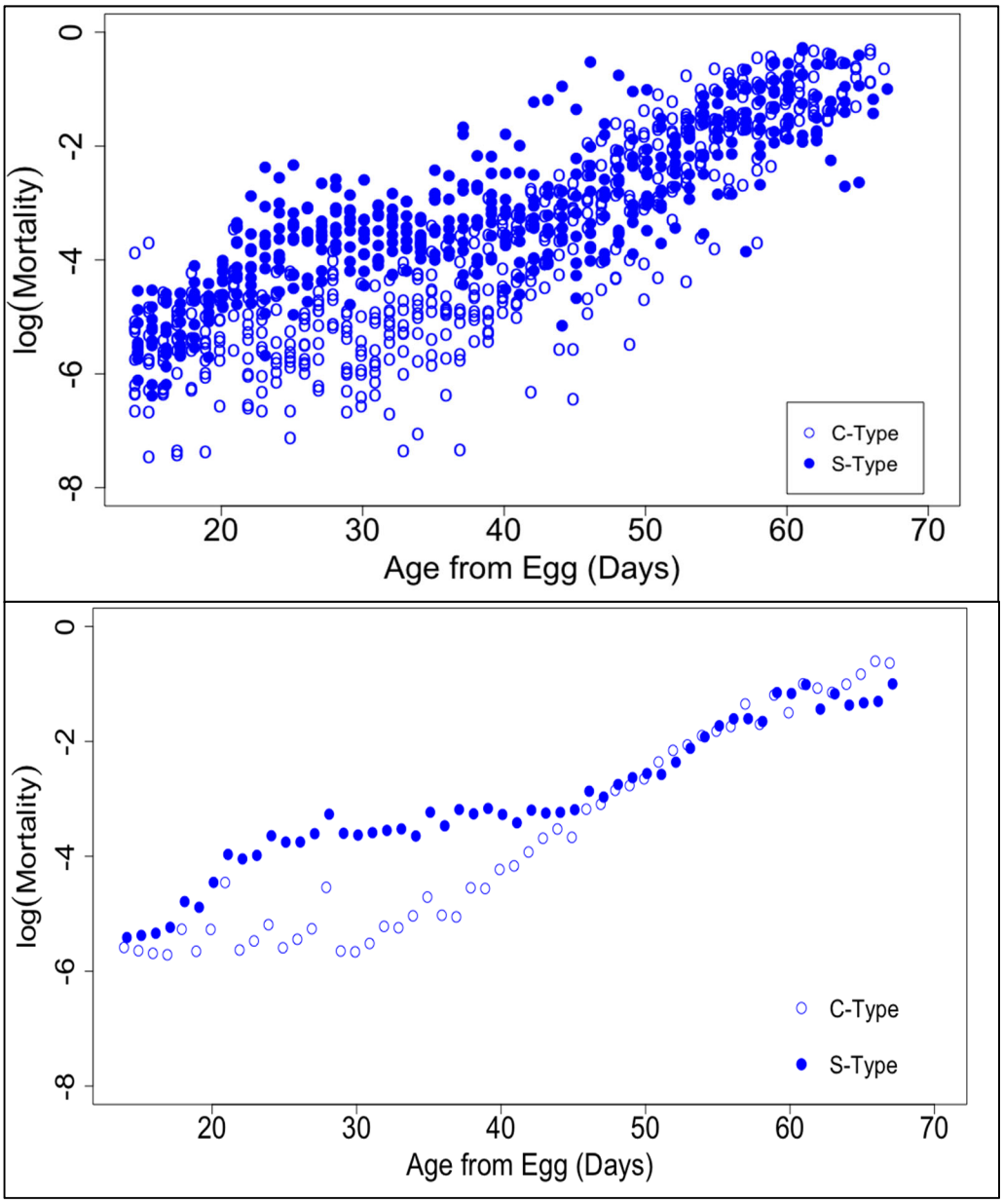
log(Mortality)
0
-2
-4
-6
-8
20
30
40
50
60
70
Age from Egg (Days)
C-Type
S-Type
log(Mortality)
0
-2
-4
-6
-8
20
30
40
50
60
70
Age from Egg (Days)
C-Type
S-Type

Figure 4.

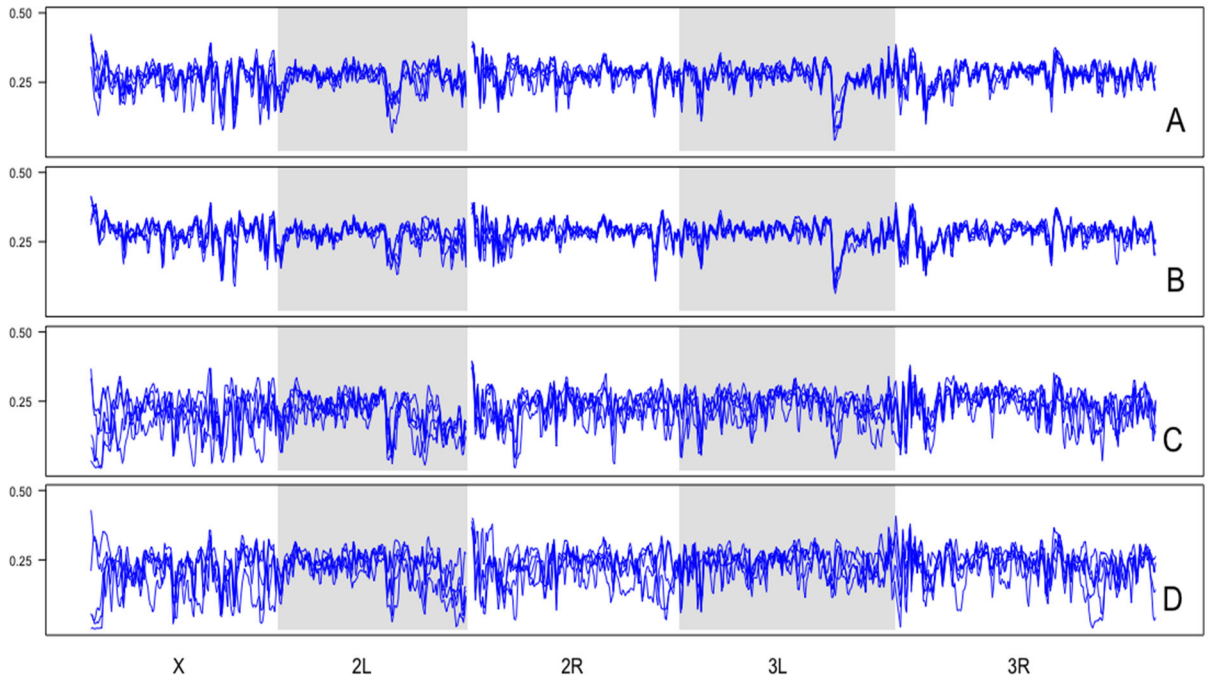

Figure 5.

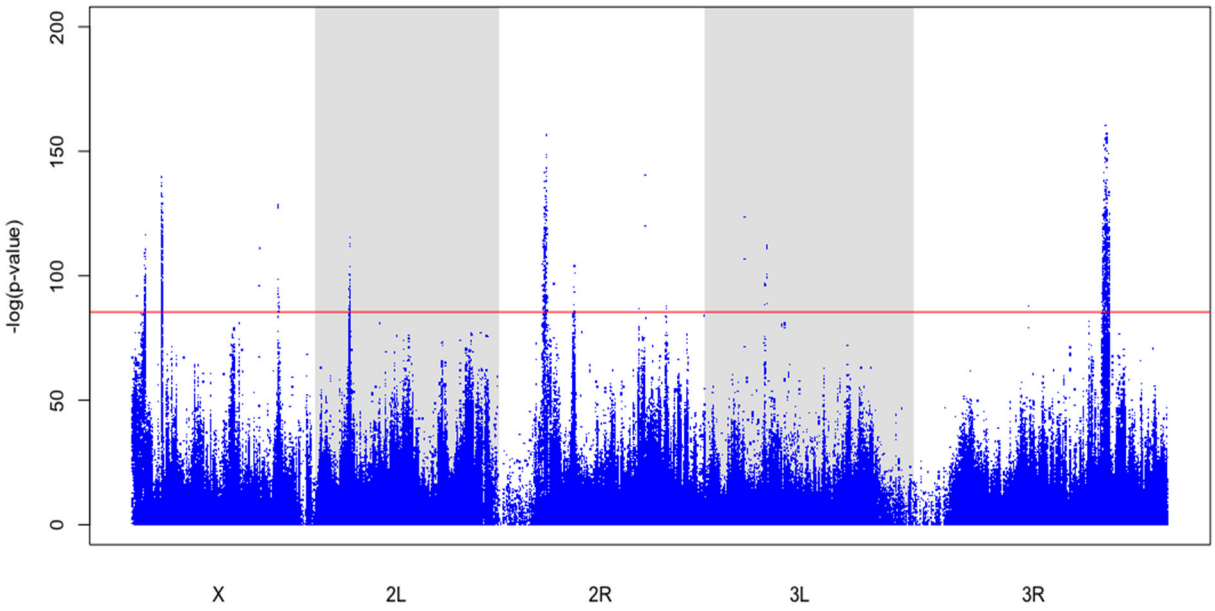

Figure 6.