

Parallel trait adaptation across opposing thermal environments in experimental *Drosophila melanogaster* populations

Ray Tobler,^{1,2} Joachim Hermisson,^{3,4} and Christian Schlötterer^{1,5}

¹Institut für Populationsgenetik, Vetmeduni Vienna, Veterinärplatz 1, Vienna, A-1210, Austria

²Vienna Graduate School of Population Genetics, Vetmeduni Vienna, Veterinärplatz 1, Vienna, A-1210, Austria

³Department of Mathematics, University of Vienna, Nordbergstrasse 15, 1090 Vienna, Austria

⁴Max F. Perutz Laboratories, Dr. Bohr-Gasse 9, 1030 Vienna, Austria

⁵E-mail: christian.schloetterer@vetmeduni.ac.at

Received October 13, 2014

Accepted June 2, 2015

Thermal stress is a pervasive selective agent in natural populations that impacts organismal growth, survival, and reproduction. *Drosophila melanogaster* exhibits a variety of putatively adaptive phenotypic responses to thermal stress in natural and experimental settings; however, accompanying assessments of fitness are typically lacking. Here, we quantify changes in fitness and known thermal tolerance traits in replicated experimental *D. melanogaster* populations following more than 40 generations of evolution to either cyclic cold or hot temperatures. By evaluating fitness for both evolved populations alongside a reconstituted starting population, we show that the evolved populations were the best adapted within their respective thermal environments. More strikingly, the evolved populations exhibited increased fitness in both environments and improved resistance to both acute heat and cold stress. This unexpected parallel response appeared to be an adaptation to the rapid temperature changes that drove the cycling thermal regimes, as parallel fitness changes were not observed when tested in a constant thermal environment. Our results add to a small, but growing group of studies that demonstrate the importance of fluctuating temperature changes for thermal adaptation and highlight the need for additional work in this area.

KEY WORDS: Experimental natural selection, fitness, temperature variation, thermal shock.

Determining which traits are involved in the adaptation of populations to new environments is a fundamental question in evolutionary biology. In moving to new environments, organisms are inevitably confronted by novel stressors, with thermal stress being particularly prominent (Hoffmann and Parsons 1991; Hoffmann and Hercus 2000; Bijlsma and Loeschcke 2005; Sørensen and Loeschcke 2007). Temperature affects diverse aspects of organismal physiology and metabolism and therefore represents a fundamental selective agent across all domains of life (Angilletta 2009). The ecological and evolutionary importance of thermal variation is readily observable in natural populations of ectothermic species whose geographical distribution is to a large extent determined by ambient temperature conditions (Andrewartha and Birch 1954).

The ectotherm *Drosophila melanogaster* has been the focus of numerous thermal evolution studies in both natural and experimental settings and has played a leading role in identifying which phenotypes are involved in thermal adaptation. Experimental evolution has proven to be invaluable to this end, allowing for specific thermal stressors to be investigated while simultaneously controlling demography and allowing replication. These studies can be broadly divided into two categories: those which use truncating selection on a stress-related phenotype, in which a specific trait is directly enriched in successive generations by the researcher, and those which use a novel thermal environment to impose the selection pressure. Although the former method guarantees a rapid change in the selected phenotype, the latter method, often called “experimental natural selection” (Garland and Rose 2009), allows



for the discovery of any trait that is critical for adaptation in new environments.

As small-bodied ectotherms, *D. melanogaster* and other insects are able to maintain optimal internal thermal conditions by altering their behavior, that is, through behavioral thermoregulation, or by adjusting their performance at different temperatures, that is, by altering the thermal sensitivity of certain traits (Angilletta et al. 2002). Although behavioral thermoregulation appears to be important in natural *D. melanogaster* populations (Feder et al. 2000), opportunities for this are expected to be limited in most experimental settings that typically employ constant temperatures and small spaces that allow for little thermal heterogeneity. Rather, experimental populations are expected to adapt by altering traits that optimize their sensitivity to their new thermal environments. Indeed, experimental *D. melanogaster* populations subjected to different constant temperatures show a range of putatively adaptive responses, with populations from hotter temperatures tending to have increased resistance to acute heat stress, decreased body size, and delayed development (Huey et al. 1991; Cavicchi et al. 1995; James and Partridge 1995; James et al. 1997). The adaptive nature of these responses is reinforced by geographical clines for each of these traits in natural *D. melanogaster* populations, including counter gradients for acute heat and cold stress resistance, which show genetic causation when assessed under laboratory conditions (Coyne and Beecham 1987; James et al. 1995, 1997; Hoffmann et al. 2002).

The most direct means of measuring adaption, however, is to quantify the fitness of evolving populations. Two important components of fitness in experimental *Drosophila* populations are larval competitiveness and viability (Lewontin 1955; Mather and Caligari 1981; De Miranda and Eggleston 1988). Such measurements can be obtained by directly competing different experimental populations; however, this is typically not possible in most studies due to the absence of population-specific markers. Nonetheless, robust measures of larval competitive ability and viability have been obtained previously for experimental *Drosophila* populations by assessing the relative performance of wild-type larvae competed against a mutant strain (e.g., Santos et al. 1992; Kraaijeveld and Godfray 1997; Mery and Kawecki 2003; Trotta et al. 2007; Vijendravarma et al. 2009). Moreover, previous work has demonstrated that good estimates of the average Darwinian fitness of whole populations can be obtained by combining larval competition with fecundity measures into a single assay (e.g., Knight and Robertson 1957; Hartl and Jungen 1979; Haymer and Hartl 1983; Yamazaki 1984; Nuzhdin et al. 1995). Notably, such methods have only rarely been used to quantify fitness of *Drosophila* populations evolving at different temperatures (e.g., James and Partridge 1998; Santos et al. 2006).

Here, we use such a multicomponent fitness assay to determine whether two experimental *D. melanogaster* populations

were adapting to opposing thermal environments that cycled daily between benign and stressful temperatures. Because the two environments differed primarily by the stressful temperature—either heat or cold stress—we predicted that populations would become better adapted to their own environments but not to the other. By extrapolating from results from other thermal natural selection studies (i.e., Huey et al. 1991; Cavicchi et al. 1995; James and Partridge 1995; James et al. 1997), we further hypothesized that selection would lead to improved stress resistance for the matching extreme temperature, and worse performance for the opposite stress, and slower (faster) developing, smaller (larger) individuals in the hot (cold) environment. However, while both populations were adapting to their respective environments, these traits tended to change in parallel rather than diverging through time. Furthermore, our results implicated the rapid daily temperature changes as the common stressor. Thus, our study demonstrates that temperature fluctuations represent an important and underappreciated selective agent that may contravene expectations based on constant thermal environments.

Methods

EXPERIMENTAL POPULATIONS

Female *D. melanogaster* were collected from Póvoa de Varzim, Portugal in July 2008 (Orozco-terWengel et al. 2012). Following the collection, females were maintained in the laboratory as 113 isofemale lines for five generations, at which point they were used to create 10 mass-bred replicate “base” populations. Each replicate base population comprised five females from each isofemale line (i.e., 565 females per replicate). The brief time between the collection and the establishment of the base population was used to limit the effects of drift within isofemale lines, thereby preserving much of the original segregating variation and decreasing the extent of linkage disequilibrium (LD) in the base population caused by amplifying inbred isofemale lines (also see Discussion). Five of the replicated base populations were thereafter maintained in either a “hot” or “cold” environment (see Experimental Environments). Replicate populations were maintained at approximately 1000 individuals, with approximately 50:50 sex ratios. Generations did not overlap; adults from the previous generation were collected following the peak of eclosion, randomly distributed into five new bottles and left to oviposit for two days, then transferred into another set of five new bottles for two more days of oviposition prior to being frozen for DNA storage. To prevent implicit selection for early fecundity, new generations were restricted to offspring from the second transfer when possible. Each replicate population was maintained in five 300 mL plastic bottles with 70 mL standard medium for the first 37 (23) generations in the hot (cold) environment, after which they were maintained in 170 mL plastic bottles with 40 mL of standard medium. Preliminary trials

revealed that bottle size had no impact on the eclosion probability for density-controlled populations (based on egg transfers). Furthermore, generation times remained consistent throughout the experiment (approximately two and four weeks in the hot and cold environments, respectively). Thus, consistent general population dynamics were maintained in each environment throughout the experiment, and this was not affected by the change in bottle size.

EXPERIMENTAL ENVIRONMENTS

Both the hot and cold environments maintained a bimodal temperature and light regime on a repeated 12:12 h cycle. Temperatures cycled between 18°C and 28°C, or 10°C and 20°C, in the hot and cold environment, respectively, with the light (dark) period coinciding with the hot (cold) phase. These temperatures were chosen because prolonged exposure to temperatures below 10°C or above 28°C are known to decrease fecundity and viability in European *D. melanogaster* (Pétavy et al. 2001; Chakir et al. 2002), whereas the two intermediate temperatures (i.e., 18°C and 20°C) are comparatively benign. The natural climate of the experimental founder populations is typified by mild temperatures and moderate diurnal temperature fluctuations throughout the year (Fig. S1).

By placing thermochronometers (Davidson et al. 2003) inside the bottles used in the experiment—with one encased in the food and another in the air near the top of the bottle—we were able to estimate the rate of cooling and heating that was experienced by the larvae and adults during the experiment. Heating was particularly fast, with 85% of temperature change taking approximately 30 min in the air, and approximately 60 min in the food, for both environments (Fig. S2). Cooling was slower, but still relatively rapid, with 85% of temperature change taking approximately 45 min in the air, and approximately 75 min in the food, for both environments. The last 15% (i.e., 1.5°C) of change typically took 1–2 h longer due to the smaller temperature difference inside and outside the bottle as the two environments approached thermal equilibrium.

RECONSTITUTED BASE POPULATION

To quantify and polarize the evolutionary changes during the experiment, the two evolving populations were compared to the original founder (base) population. Because we wanted to avoid any possible confounding environmental influences on the phenotypes due to assaying across different time points (see Controlling for Confounding Trait Variation), this required reconstituting five biological replicates of the base population from the isofemale lines in an identical fashion to the original base population as described above. The reconstituted populations should be good proxies of the original base population, providing that the effects of any genetic changes that have accrued within each isofemale line are random with respect to the two experimental environments. This seems to be a reasonable assumption for our exper-

iment, given that the lines were maintained at low to moderate densities (approximately 20–100 individuals/vial/generation) and a temperature (18°C) that should be nonstressful for cosmopolitan *D. melanogaster* (Chen et al. 2015). Under these conditions, selection should be relatively ineffective and drift should dominate within lines. Furthermore, only three isofemale lines (out of 113) became extinct during the study. This would have caused only minor deviations in allele frequencies and suggests that any lost variants were rare, whereby the average trait values should be largely unchanged. Finally, although the isofemale lines would have become increasingly inbred during the experiment—which has been associated with reduced thermal stress resistance in *D. melanogaster* (Ehiobu et al. 1989; Dahlgaard and Hoffmann 2000; Pedersen et al. 2005)—our experimental design meant that the reconstituted base population flies were essentially F1 crosses between different isofemale lines (see below). This should effectively mask the aggregate effects of inbreeding depression within lines. Further appraisal of the impact of this method on the interpretation of our findings is provided in the Discussion.

PHENOTYPIC ASSAYS

A series of phenotypic assays was performed on the evolving populations from the hot and cold environments—henceforth referred to as the hot population and cold population, respectively—along with a reconstituted base population. The following general trait categories were assayed: (1) acute thermal stress resistance, (2) body size, (3) developmental time, and (4) fitness. The first three have a long history of investigation in *Drosophila* thermal adaptation research and can be measured in a large-scale parallel fashion, whereas fitness has rarely been directly quantified. All assays were performed after two generations in a shared environment (i.e., common garden environment, CGE; see below). Serial transfers of first-generation adults were used to generate a large number of second-generation CGE adult flies for phenotyping. Unless stated otherwise, each assayed trait used a separate transfer of the second-generation CGE flies. All adult flies were aged between two and seven days when assayed, thereby closely matching the age range of flies used to seed new generations of the experimental environments (and thereby being most relevant for adult fitness). For all traits other than fitness, parallel assays were performed at two different time points during the experiment—(1) F34 cold and F59 hot and (2) F43 cold and F75 hot—with each utilizing an independently reconstituted base population. Fitness was only assayed for the later set of generations.

CONTROLLING FOR CONFOUNDING TRAIT VARIATION

To improve our estimation of any adaptive phenotypic response, we sought to minimize the contribution of potentially confounding variables by performing the following three steps:

1. Prior to the phenotypic assays, all populations were subjected to two generations in a CGE that consequently standardized both the environmental and trans-generational effects. Populations in the CGE were maintained under the standard culture conditions used in the experimental environments with 12-h light:12-h dark cycles that were synchronized across all environments. Four different common garden experiments were conducted, two at each of the cycling temperatures (18–28°C and 10–20°C) and two more at the mean temperature of each environment (i.e., 15°C and 23°C). Fitness traits were assayed in both cycling environments and the 23°C environment, while all other traits were limited to the two constant temperature environments. The choice of a single temperature for nonfitness traits was motivated by the desire to minimize any potential phenotypic variation that may have arisen due to development occurring within the bimodal experimental temperature regime.
2. To control for any trait-specific effects of differing larval densities, a fixed number of eggs (approximately 200) were used to seed the second generation of each CGE for all populations—that is, the hot, cold, and reconstituted base—and replicates. To do this, first-generation CGE adults were transferred to 170 mL bottles with new food and allowed to oviposit overnight. On the following morning, sections of food containing the required number of eggs were cut out and transferred into new 170 mL bottles (also containing food). The remaining first generation CGE adults were later discarded, and the bottles containing eggs returned to the relevant CGE. Two hundred eggs per bottle can be considered a moderate density relative to that in the experimental environments, where we were typically able to recover >1000 adults from a single transfer (i.e., five bottles) for each replicate.
3. For a given CGE, the hot, cold, and reconstituted base populations were assayed in parallel. In other words, all replicates of the two evolved populations, along with five newly reconstituted base population replicates, were simultaneously subjected to all steps of the CGE and subsequent phenotypic assays. This last step ensures that all environmental variables were shared among the hot, cold, and base populations for a given CGE.

THERMAL STRESS RESISTANCE

Two types of ecologically relevant thermal stress resistance were quantified: the time taken for adult flies to either become knocked down or to recover from an acute exposure to heat and cold stress, respectively (Berrigan and Hoffmann 1998; David et al. 1998). For knockdown time to heat stress, 12 adult flies were placed into a transparent 200 mL plastic tube (Greiner) that was then sealed with plastic wrap and placed into an incubator at 38.5°C. To alleviate possible desiccation, a damp 1 × 1 cm square of Whatman filter paper was added to each tube prior to the start of the assay.

Knockdown time was measured as the time taken for individuals to leave their feet or become inactive. For recovery from cold stress, 12 adult flies were moved into an empty vial that was placed in an ice-filled tray inside a 0°C freezer where they were kept for 4 h. Following this, they were transferred into an empty Petri dish in a room maintained at a constant 25°C ($\pm 1^\circ\text{C}$) where the time taken to return to a standing position was recorded. The number of individuals knocked down/recovered was recorded at the end of every minute, with each assay continuing until either 60 min had passed or all individuals were knocked down/recovered. For both stress phenotypes, eight sets of tubes/Petri dishes were assayed at a time. Flies from different populations, replicates, and sexes were assayed in separate tubes/Petri dishes, and were randomly distributed across assays. This was done in blocks for the heat stress assays to account for a weak (approximately 0.5°C), but persistent, thermal gradient across the incubator. Sexes were separated while anaesthetized with CO₂, with the assays taking place two days later to ensure sufficient recovery from the adverse effects of this procedure (Nilson et al. 2006; Colinet and Renault 2012). Overall 48 individuals were assayed for each combination of CGE, population, sex, and replicate (i.e., approximately 3000 flies per assay).

DEVELOPMENTAL TIME

As described above, the second-generation CGE populations were established by leaving first-generation CGE adults to oviposit in new food for 24 h after which approximately 200 eggs were collected, transferred to new food bottles, and returned to the relevant CGE. Developmental time was then measured as the time from the initial oviposition period to pupation time, that is, the larval development time (James and Partridge 1995). Pupae were counted everyday until no new pupae were observed. The pupae were not sexed. Four bottles were quantified in this way for each replicate in the first assay (F34 cold and F59 hot), and two in the second assay (F43 cold and F75 hot). Only the 23°C CGE was performed in the second assay.

BODY SIZE

Wing and leg measurements are correlated with general body size in *D. melanogaster* (Karan et al. 1998; Partridge et al. 1999; Turner et al. 2011). For each CGE, replicate and population, tibia length and wing area were measured for 30 males and females (i.e., 1800 flies per assay). A single rear leg and wing were dissected for each assayed individual and attached to a standard slide via a strip of adhesive tape. These were then photographed using a digital camera (DFC 300 FX) attached to a Leica M205-FA stereomicroscope and the resulting images analyzed using ImageJ software (ImageJ64, version 1.43; Schneider et al. 2012). Tibia length was measured as the distance between the adjoining tarsus and femur, and wing area was measured according to Gilchrist and Partridge (1999). All flies used for this assay were derived from

the eclosed adults of the developmental time assays. Flies were stored in a -20°C freezer for several months prior to dissection.

FITNESS

Previous work has shown that assays that combine key fitness components in experimental populations (i.e., female fecundity, larval competitiveness, and larval viability) are good proxies for the average Darwinian fitness of whole populations in *Drosophila* (Knight and Robertson 1957; Hartl and Jungen 1979; Yamazaki 1984). We followed a similar strategy to measure fitness for the hot (F75), cold (F43), and reconstituted base populations. For each of the hot, cold, and reconstituted base population, 40 gravid females were transferred to a 170 mL bottle along with 60 gravid white-eye *D. melanogaster* “reference” females (w1118). These were left to oviposit for two days, after which they were removed and their progeny expected to compete for resources, whereby larval competitiveness and viability, along with female fecundity, were captured by these assays. This was repeated twice for each replicate, population, and environment, such that there were 90 “subreplicates” in total. Following the emergence of the subsequent generation, fitness was determined in each subreplicate by quantifying the number of enclosing individuals that were wild-type (i.e., progeny of base, hot, or cold population flies). To ensure that the vast majority of surviving progeny were scored, eclosed adults were collected every five days for 15 days after the first individual eclosed; that is, three counts were performed in total. Statistical analyses (see below) were performed by combining consecutive counts of eclosed individuals for days ≤ 10 and ≤ 15 . Because the test results were qualitatively similar between the two categories, only the latter is reported. To expedite the counting, males and females were quantified together. Previous trial assays had revealed that there was no discernable fitness difference between the two sexes for any combination of population and environment (data not shown).

STATISTICAL ANALYSES

Because we were primarily interested in determining whether the populations were evolving to their new environments, we tested two specific adaptive hypotheses, having first fitted a generalized linear mixed model (GLMM) to each trait (see below). First, we tested whether each pair of hot, cold, and base populations were different conditional on sex and CGE. Second, we investigated whether phenotype values differed across the two assayed generations for each of the evolved populations, again conditional on sex and CGE. Thus, the first contrast tested whether the three populations differ after accounting for sex and CGE effects. The second contrast tested whether the populations were continuing to evolve through time, which required performing tests on transformed trait values that accounted for confounding variation across assays (see GLMM Fitting). All contrasts used the fitted values

from the full models; that is, model reduction was not performed to ascertain the best fitting set of independent variables. This was justified on the grounds that removing the nonsignificant factors from the GLMMs made no difference to the significance of the subsequent contrasts (given that the effects of the nonsignificant terms were small).

GLMM FITTING

For all traits except fitness, the explanatory variables comprised two fixed factors, population (i.e., base, hot, or cold) and CGE (i.e., 15°C or 23°C), and replicate as a nested random factor. Sex was included as an additional fixed factor for both the thermal shock and morphological traits. A block factor was also included for the heat stress phenotype to account for a thermal gradient across the incubator. For the fitness traits, the CGE factor was replaced with a factor for the environment of the assay (i.e., $10-20^{\circ}\text{C}$, $18-28^{\circ}\text{C}$, or constant 23°C). As for the predictor variables, the thermal stress (knockdown time or recovery time) and the developmental time (day of pupation) data were all fitted with a gamma response distribution (i.e., inverse link), because this resulted in a better fit of the data than using either normal or log-normal distributions (results not shown). The morphological data (tibia length and wing area) were fitted with a standard Gaussian error function, and the fitness data with binomial response distribution (i.e., logit link), as the predicted variable was the proportion of the wild-type adults. The fitted models contained all possible interactions between fixed factors, which therefore included three and four way interactions in some cases.

Initial screening of the raw data suggested that the measurements were affected by uncontrolled environmental variation for some traits (see Results). This meant that using the raw observed values to make comparisons between datasets that were collected at different time-points—namely, different CGEs from the same population, or comparisons across the two assayed generations—could lead to spurious results. Hence, for each trait we transformed the predicted variable by dividing all observed values of the evolved populations by the mean base population value conditional on sex, CGE, and assayed generation (i.e., either the first or second set of assays). Because the base population is expected to be genetically homogeneous for both assays (see Reconstituted Base Population), between-assay phenotypic variation for this population should be attributable to uncontrolled experimental variables. Consequently, dividing all data by the relevant base means should control for such variation and facilitate comparisons across the different assayed CGEs and generations. After transforming the data for each trait, GLMMs were refitted to the combined data for the two assayed generations, with an additional fixed explanatory variable included to capture any assay effects. Visual inspection of QQ plots and residual homoscedasticity indicated that the adjusted data fulfilled the assumptions of their

respective response distributions (data not shown). Tables S1 and S2 provide the ANOVA results from the GLMMs fitted to the raw and adjusted data, respectively.

SOFTWARE

All analyses were performed using R (version 3.0.1; R Core Team 2013). The GLMMs used the lme4 package (version 1.0-4; Bates et al. 2013) with subsequent significance testing performed via the ANOVA function from the car package (version 2.0-19; Fox and Weisberg 2011). Parameters were estimated using restricted maximum likelihood (i.e., the “REML” option). Contrasts used the phia package (version 0.1-5; De Rosario-Martinez 2013), with the “Holm” method (Holm 1979) being used to correct for multiple testing.

Results

FITNESS

As predicted, both evolved populations were significantly better adapted to their native environment than the base population. Unexpectedly, however, both evolved populations were also significantly fitter than the base population in their non-native environment (Fig. 1, Table S3). For the hot population, the mean proportion of wild-type flies increased by approximately 50 and 15% in the hot and cold environment, respectively, whereas the cold population increased by approximately 20% in both hot and cold CGEs. Thus, the evolved populations had the highest fitness in their native environments—although the difference between the evolved populations was only significant in the hot environment—implying that selection was driven by habitat-specific and common factors.

Although many stressors common to both environments are conceivable, we suspected that the rapid temperature fluctuations in each environment were a major driver of the parallel adaptive response. The flies must tolerate the metabolic and physiological effects induced by these thermal fluctuations and are unlikely to have encountered such rapid changes in natural settings (see Discussion). To test whether the rapid temperature fluctuations were imposing a common stress in both environments, we exposed flies to a single fixed temperature, 23°C, while holding all other experimental parameters constant.

Indeed, the differences between the base and evolved populations were substantially reduced in this constant environment (Fig. 2, Table S3). The hot populations now exhibited a small, but significant, increase in fitness relative to the cold and base populations, whereas the difference between the cold and base populations was negligible.

Finally, because the number of eclosed adult flies was similar for the three populations, most notably for the two evolved

populations (Figs. 1 and 2), the fitness differences between the three populations were not simply the result of differential female fecundity. Rather, fitness appeared to be largely driven by differences in larval competitiveness and viability between the three populations.

THERMAL STRESS RESISTANCE

Our expectation of habitat-specific improvements in thermal stress resistance was not realized. Rather, both acute heat and cold stress resistance had a tendency to increase in both evolved populations relative to the base population, and this difference was significant for many of the contrasts (Fig. 3, Table S3). Cold stress resistance at 15°C CGE was a notable exception, with all three populations showing a similar level of resistance. This is indicative of a hardening effect at this temperature, which outweighed any selection on this trait. Notably, the two evolved populations tended to be similar to one another in all cases, regardless of sex or developmental temperature.

After correcting for the assay-specific variation (see Methods), several sex–CGE combinations showed significant evidence for ongoing intergenerational change in both hot and cold populations (Fig. S3, Table S4). Of the significant combinations, the majority (6 of 9) had significantly increased stress resistance across generations, such that stress resistance was improving in most, but not all, cases.

DEVELOPMENTAL TIME

In keeping with results from a previous *D. melanogaster* experimental evolution study (James et al. 1997), the hot population evolved a significantly longer development time than the cold population in both CGEs and assayed generations. However, development time was significantly increased in both of the evolved populations when compared to the base, which contradicts our expectation of a divergent response for this trait (Fig. 4, Table S3). The scale of differences between treatments was sufficiently large—for example, nearly two days separated the base and cold populations in the 15°C CGE—to suggest that these patterns were unlikely to be a simple consequence of differential oviposition time or egg retention across the treatments during the 24 h egg laying period prior to assay. The difference in development time between the base and evolved populations began to decrease again between the two assayed generations (after correcting for assay-specific variation) in the 23°C CGE, significantly so in the case of the hot population (Fig. S4, Table S4).

BODY SIZE

The influence of temperature on plasticity in body size is well characterized—body size increases with decreasing developmental temperature (Kingsolver and Huey 2008)—and this pattern was replicated across all of our populations. However, the

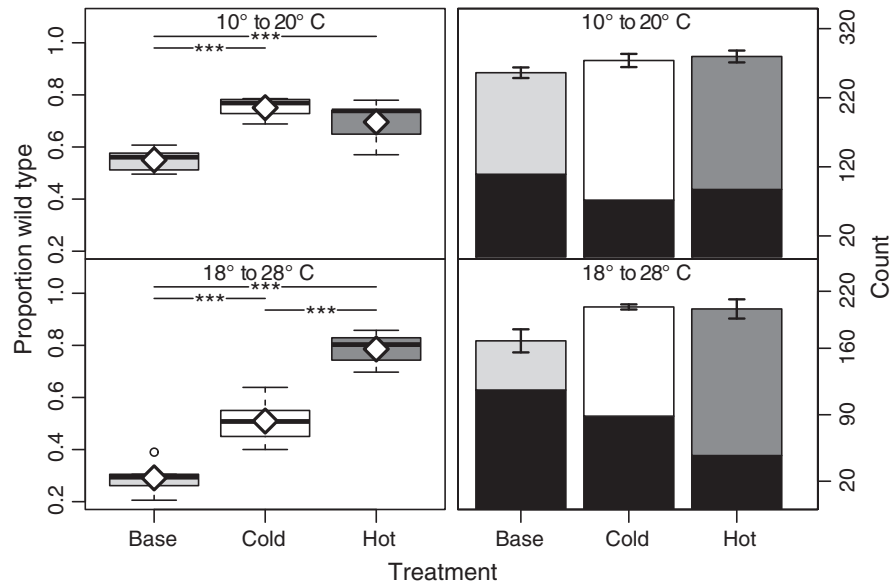


Figure 1. Boxplots showing the fitness of the hot (dark gray), cold (white), and base (light gray) populations quantified as the proportion of adult progeny (wild-type) that eclosed after larval competition with an inbred white-eye strain (left-hand panels). White diamonds specify the mean values. The total number of eclosing wild-type (coded as above) and white-eye (black boxes) adults is shown in the adjacent barplots (right-hand panels). Competition took place in each of the two experimental environments (cold, top panels; hot, bottom panels). Analyses were based on all adult flies observed to eclose for 15 days after the first observed eclosion. Significance for each pairwise comparison between populations is shown at the top of each panel ("." ≤ 0.1 , "*" ≤ 0.05 , "***" ≤ 0.01 , "****" ≤ 0.001).

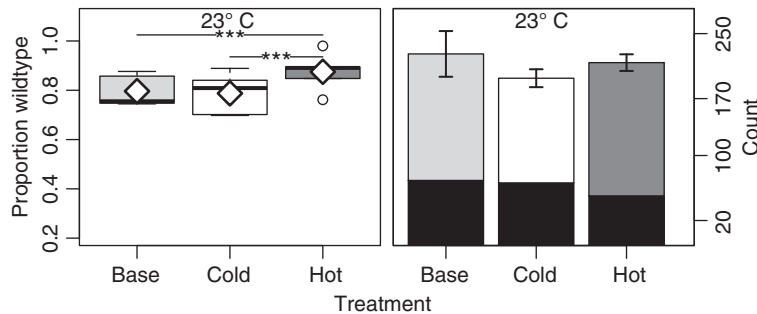


Figure 2. As for Figure 1 but with fitness assays carried out in a constant 23°C thermal environment. Analyses were based on all adult flies observed to eclose for 15 days after the first observed eclosion. Significance for each pairwise comparison between populations is shown at the top of each panel ("." ≤ 0.1 , "*" ≤ 0.05 , "***" ≤ 0.01 , "****" ≤ 0.001).

negative correlation between body size and temperature reported in previous *Drosophila* experimental evolution studies (Anderson 1972; Cavicchi et al. 1989; Partridge et al. 1994) did not evolve in our populations. Rather, no consistent trend emerged for either evolved population across the different time points (Fig. 5, Fig. S5, Table S3). Importantly, the variation in body size between the two different time points was considerably smaller for the base population than for the evolved populations. Furthermore, a strong positive correlation was observed between wing and leg dimensions ($r = 0.688$ and 0.687 , based on all 1500 individuals for the first and second assays, respectively). This confirms the reliability of our measurements because these traits are known to be correlated (e.g., Robertson 1962; Turner et al. 2011)

and suggests that these morphological traits were be reasonably robust to uncontrolled environmental variation introduced by our assaying procedure.

Discussion

Our results clearly demonstrate that both of our experimental populations were adapting to their new environments, showing significant fitness gains relative to the base population. However, the expected responses were not observed for any trait; rather each phenotype either changed in parallel (e.g., thermal stress resistance and development), or fluctuated without any obvious pattern (e.g., morphology). Importantly, the consistency of the

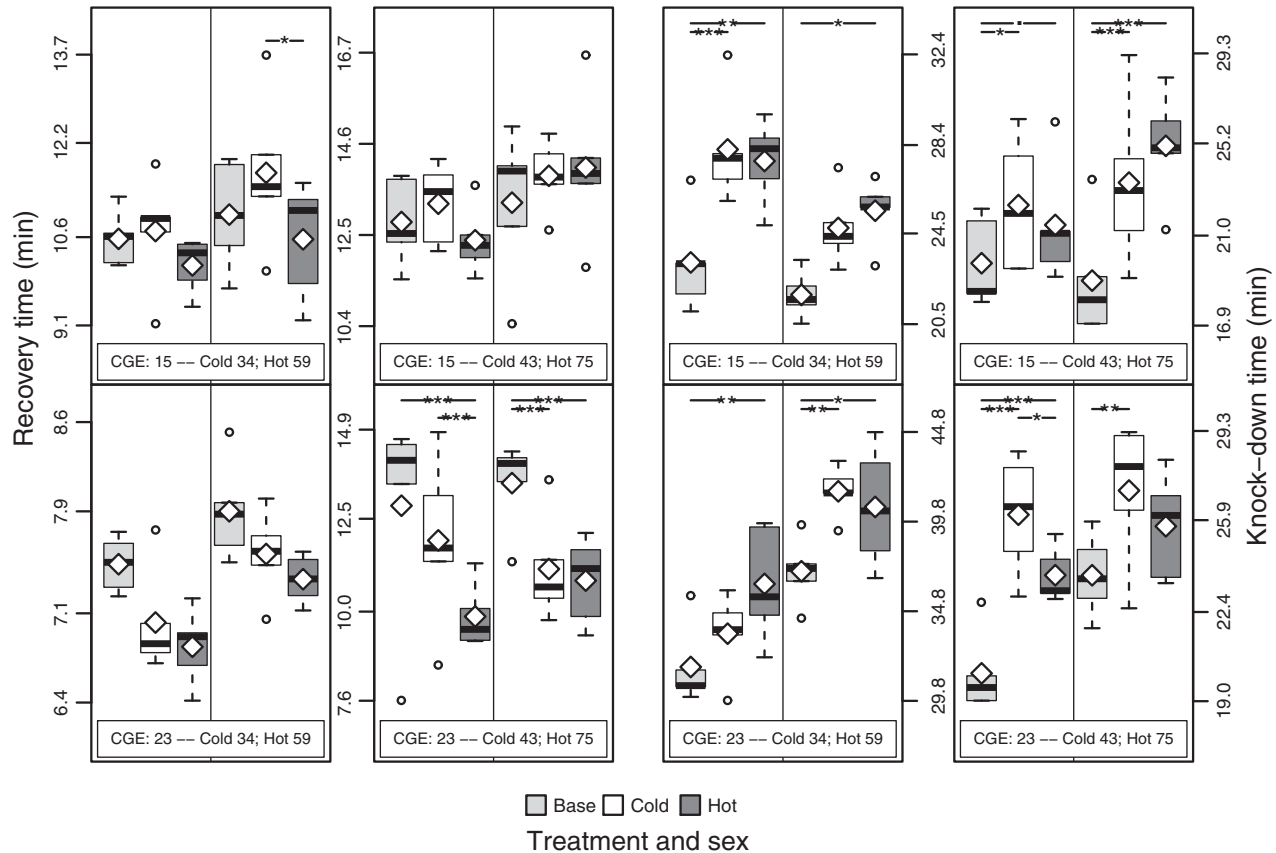


Figure 3. Box plot of recovery time (two left-most columns) and knock down time (two right-most columns) from acute cold and heat stress, respectively. Information on CGE and generation is inset in each panel, with females (males) on the left (right) of the central vertical line in each panel and population status shown in the key beneath the plot. The white diamonds specify the mean values. Significance for each pairwise comparison between populations is shown at the top of each panel ("." ≤ 0.1 , "*" ≤ 0.05 , "***" ≤ 0.01 , "****" ≤ 0.001).

phenotypic response across replicates, in combination with the use of common environments for trait assays, strongly imply that selection—and not plasticity, trans-generational effects, or drift—drove these changes. These findings imply that a common selective agent was a major driver of adaptation in both evolved populations in conjunction with environment-specific thermal stress. Possible explanations for these patterns are discussed in the following four sections.

POTENTIAL BIAS INTRODUCED BY THE EXPERIMENTAL DESIGN

The observed parallel responses may have arisen if our phenotype measurements for the reconstituted base populations were systematically biased relative to those of the original base population. One possible source of bias comes from increasing inbreeding depression within the isofemale lines across the experiment. Nonetheless, our experimental design is expected to mask such effects by generating hybrids from inter-isofemale line crosses (see Methods). Indeed, the similarity in fitness between base and evolved populations in the constant 23°C environment

demonstrates that any putative within-isofemale line inbreeding depression did not influence the fitness of the reconstituted base populations. Within-isofemale line selection during the experiment represents another possible source of bias, as suggested by a previous study where thermal stress resistance responded to selection in *D. melanogaster* isofemale lines (Morrison and Milkman 1976). However, it is unclear why selection in the constant 18°C environment occupied by the isofemale lines would lead to decreased thermal resistance (or other trait changes) in both of the experimental environments. Furthermore, the adaptive responses in the Morrison and Milkman study were driven by strong truncating selection. We anticipate that similarly strong selection would be necessary for alleles to consistently escape stochastic loss from drift in our isofemale lines given their small population sizes. In the unlikely scenario that such a selected locus did consistently escape drift and had a measurable effect on the average phenotypes in our experimental conditions, laboratory selection stands out as a plausible candidate driver. Notably, such loci would ultimately lead to fitter reconstituted base populations in relation to the initial base population, making our estimates

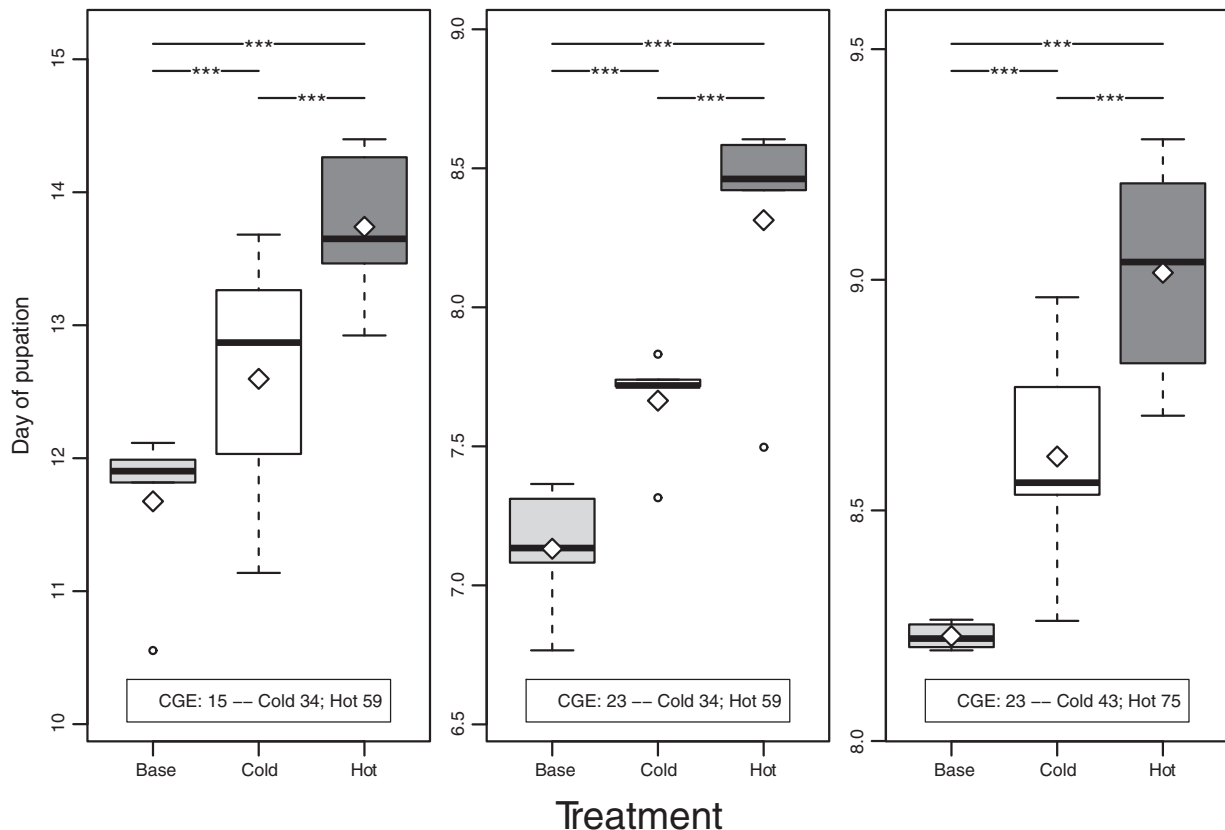


Figure 4. Box plot of developmental time. Information on CGE and the generations assayed are displayed within each panel. The x-axis labels indicate the population. Note that there was no assay for the 15°C CGE in the second set of assayed generations. The white diamonds specify the mean values. Significance for each pairwise comparison between populations is shown at the top of each panel ("." ≤ 0.1 , "*" ≤ 0.05 , "**" ≤ 0.01 , "***" ≤ 0.001).

of fitness change more conservative overall. In summary, we consider the reconstituted base populations to be a robust basis for quantifying the trait changes during the experiment.

Another potential source of bias comes from the 15-day interval of adult collection used in the fitness assays, which may have allowed for larval competition between slow developing larvae and the descendents of more rapidly developing individuals. Thus, because developmental rates had genetically diverged across the base, cold, and hot populations, larval competition may have been most intense for the faster developing populations, particularly in the hotter assay environments. Such differential larval competition may explain the parallel fitness increases observed in the hot (18–28°C) and cold (10–20°C) environments, as well as the reduced number of adults in the former, but is incompatible with the results from the constant 23°C environment. Nonetheless, it is possible that development was generally slower in the 23°C environment than in the fluctuating hot environment, as developmental rates under constant temperatures can be retarded compared to cycling temperatures with the same mean in *D. melanogaster* (Siddiqui and Barlow 1972). Anecdotal observations indicate that such differences did not arise in our populations, however. Indeed,

cycling thermal schemes similar to our own are known to result in comparable, or even delayed, development relative to constant environments (Pétavy et al. 2001). In any case, the number of adults enclosing in both the constant 23°C and hot environments were very similar, which does not conform to the larger number expected had competition been relaxed in the constant environment. Consequently, our fitness measurements are unlikely to suffer from this type of competition-induced bias and should reflect the true fitness patterns.

LABORATORY ADAPTATION

When populations are maintained for multiple generations within a laboratory environment, they are inevitably exposed to novel conditions that may give rise selection on a variety of traits (Harshman and Hoffmann 2000; Matos et al. 2002; Huey and Rosenzweig 2008; Santos et al. 2010). This phenomenon, referred to as laboratory adaptation, has been thoroughly studied in *D. subobscura*, where it is known to drive convergent adaptive responses in life-history traits, including increased early fecundity, larval competitiveness, and development time (Matos et al. 2000; Simões et al. 2008; Fragata et al. 2014). Results from *D.*

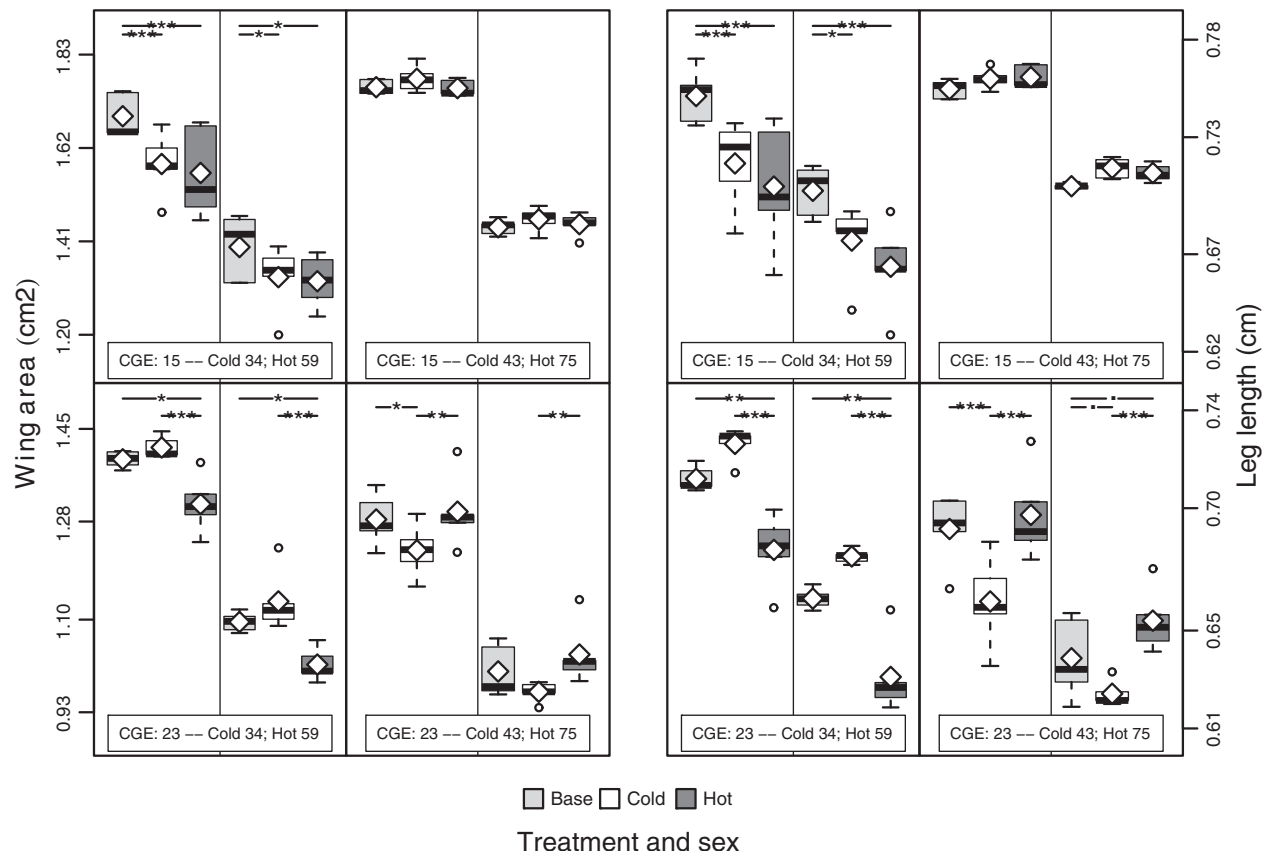


Figure 5. Box plot of wing area (two left-most columns) and rear tibia length (right right-most columns). Information on CGE and generation is inset in each panel, with females (males) on the left (right) of the central vertical line in each panel and population status shown in the key beneath the plot. The white diamonds specify the mean values. Significance for each pairwise comparison between populations is shown at the top of each panel ("." ≤ 0.1 , "*" ≤ 0.05 , "**" ≤ 0.01 , "***" ≤ 0.001).

melanogaster generally support these findings, with increased larval competitiveness, slower development, and larger body size all being reported for populations maintained in the laboratory over several generations (Frankham and Loebel 1992; Latter and Mulley 1995; Sgrò and Partridge 2000). Whether thermal tolerance traits are also subject to laboratory adaptation in *Drosophila* remains unclear, however. On the one hand, no change in either survivability or knockdown time under heat shock was found for *D. melanogaster* (Krebs et al. 2001) or sibling species *D. birchii* (Griffiths et al. 2005), respectively, following several generations of laboratory culturing. However, the latter population did evolve faster recovery to cold shock over the same culturing period (Griffiths et al. 2005). Although ecological drivers behind laboratory selection remain the subject of debate, increased larval densities and the restriction of oviposition to young adults—which are typical of *Drosophila* culturing methods, including our own—are likely suspects (Mueller 1991; Sgrò and Partridge 2000).

These results, in combination with the brief period that our flies spent in the laboratory prior to the start of the experiment,

may imply that the parallel responses reported here were largely determined by laboratory adaptation. Nonetheless, different lines of evidence suggest that this was not true for all traits. First, the cold population often displayed significantly larger changes for body size and thermal stress traits than the hot population. Second, while developmental time initially increased for both evolved populations, it had started to decrease again when measured in the later generations. Thus, trait changes tended not to be proportionate to time spent in the laboratory, or were not strictly monotonic, contravening the expectations of laboratory selection. Finally, laboratory selection should have led to the evolved populations being fitter than the base population across all three environments in which fitness was assessed in this study, given that these environments shared all aspects of the culturing regime other than temperature. That this was not the case suggests that fitness was independent of the culturing conditions in our experiment (i.e., food conditions, transfer regime, adult age, larval densities, etc.). Hence, our results indicate that laboratory selection was not the sole driver of the parallel responses observed for all traits, with fitness being the most notable exception.

ADAPTATION TO THERMAL FLUCTUATIONS

Two previous *D. melanogaster* experimental evolution studies showed that flies that had evolved in diurnally fluctuating thermal environments were fitter in both constant and fluctuating temperatures than those that had evolved under constant temperatures (Beardmore and Levine 1963; Long 1970). The authors attributed these fitness differences to adaptive plastic changes accruing in the populations that evolved to fluctuating temperatures, which improved their capacity to tolerate novel environments in general (i.e., not just those differentiated by temperature). Such plastic changes seem not to have evolved in our experimental populations, however, as the fitness differences between the evolved and base population were relatively minor in the constant environment. Rather, the environment-specific fitness gains observed in our populations strongly imply that direct exposure to the thermal fluctuations was necessary to elicit the adaptive response. The combination of the swift rate and large amplitude of temperature change occurring twice daily in our experiment (i.e., 10°C change at approximately 0.15° to $0.25^{\circ}\text{C min}^{-1}$) is probably rarely experienced by natural *D. melanogaster* populations, including the progenitor population used here (see Fig. S1 for details). Thus, because our experimental populations comprise ectotherms that are unable to relocate away from the stressful temperatures, the rapid temperature fluctuations are a plausible candidate driver for the parallel fitness responses.

Whether the rapid thermal fluctuations were also responsible for the parallel changes observed in other traits remains unclear. One possibility is that the parallel increases in thermal stress resistance were a by-product of selection for improved tolerance of the metabolic changes associated with both rapid heating and cooling. Thus, because the thermal stress assays will also induce brisk changes in internal body temperature in the captive flies, the evolved populations may have subsequently become more robust in these environments. Alternatively, selection for increased resistance to one type of stress can lead to correlated changes in other stress and life-history related traits in *D. melanogaster* (e.g., Hoffmann and Parsons 1989; Bubliy and Loeschcke 2005). In particular, developmental time was observed to decrease in *D. melanogaster* populations selected for increased cold (Mori and Kimura 2008) and heat (Bubliy and Loeschcke 2005; Mori and Kimura 2008) resistance, whereas body size changes tend not to be correlated with thermal stress resistance (Anderson et al. 1991; Bubliy and Loeschcke 2005; also see Williams et al. 2012), echoing the results in our experiment. Selection for improved cold resistance is also known to produce more heat-resistant flies, but not vice versa, in *D. melanogaster* (Bubliy and Loeschcke 2005; Mori and Kimura 2008). These results suggest that a correlated response to selection for improved thermal stress resistance may have been sufficient to drive some of the parallel phenotypic changes observed in our experimental populations.

The mechanistic basis underlying these putative intertrait correlations—for example, pleiotropy, LD between causal genes, or performance trade-offs (Arnold 1992)—is not clear. A separate study revealed that moderate increases in LD had accrued in the hot population relative to the ancestral natural population after 67 generations of evolution, being almost uniform for single nucleotide polymorphisms (SNPs) separated by up to 50 Kb (Franssen et al. 2014). Because these LD differences represent the joint effects of selection and drift during the experiment, as well as the initial LD generated by amplifying inbred isofemale lines, the latter mechanism did not appear to have caused extensive LD in our base populations. However, selected haplotype blocks, some spanning several megabases, were inferred to rise from low to intermediate frequencies over 67 generations (Franssen et al. 2014). Consequently, the putative correlated phenotypic responses could have arisen in our study if multiple different functional alleles occupied the same rare selected haplotype blocks.

ADAPTATION TO HABITAT-SPECIFIC THERMAL STRESS

In addition to the parallel adaptive response, habitat-specific fitness gains were also observed for both populations in their native environments, though only the hot population was significantly fitter. Reductions in viability (Chakir et al. 2002) and fertility (Pétavy et al. 2001) were found to be much stronger at 10°C than 28°C in other European *D. melanogaster* populations, however, suggesting that habitat-specific selection for thermal tolerance should have been stronger in the cold environment in our study. Although much of this apparent discrepancy could be due to the cold population having experienced many fewer generations of evolution, another potential explanation comes from a previous genomic analysis of these populations (Tobler et al. 2014). This study revealed that the majority of alleles that were selected over the first 15 generations of the experiment started at low frequencies in the hot environment, but were more likely to arise from intermediate to high frequencies in the cold environment. Based on the time of collection of the wild progenitor population, the authors concluded that this population had probably recently undergone selection for increased cold tolerance. Under this scenario, the base population would have preadapted the cold temperature, but relatively unfit under the hot temperature, used in the experiment. Thus, the observed habitat-specific fitness patterns could have resulted from a combination of the reduced number of generations and partial preadaptation within the cold environment.

COMPARISON WITH PREVIOUS GENOMIC STUDY

In contrast to the parallel phenotypic responses described in the current study, Tobler et al. (2014) found that the genomic response was largely environment-specific over the first 15 generations of the experiment. In particular, the top candidates in the hot and cold

populations were enriched for genes specifically associated with either heat- or cold-stress resistance, respectively, but not *vice versa*, while genes conferring resistance to both stresses (which were mostly heat shock proteins; Hsps) were not enriched in either environment (Fig. S4). This discrepancy between the genomic and phenotypic results implies that the hot and cold populations may have followed different genetic pathways in adapting to the rapid temperature fluctuations. Alternatively, temperature-specific selection may have dominated the early part of the experiment, whereas selection for tolerance to rapid temperature fluctuations may have been initially weaker, but eventually dominated fitness in the cold population in later generations. Either way, it seems unlikely that putative stress resistance genes, including Hsps, contributed to the shared fitness response observed for both evolving populations. Indeed, while Hsps facilitate general stress responses in *D. melanogaster* and other organisms (Feder and Hofmann 1999; Sørensen et al. 2003), long-term or repeated expression of Hsp genes is known to be detrimental in *D. melanogaster* (Sørensen et al. 2003). Thus, it appears that alternative pathways may be favored to deal with rapid thermal oscillations in our experiment. Elucidation of these pathways will be a focus of future work.

TEMPORAL THERMAL VARIATION AND ADAPTATION

A small but growing set of experimental studies have recently appeared that specifically address the role of temporal temperature variation on adaptation in *D. melanogaster*, either in environments where the temperature cycled across generations (Cooper et al. 2010, 2012; Yeaman et al. 2010) or within generations, but where the mean temperature increased linearly through time (Schou et al. 2014). In the latter study, resistance to desiccation and heat showed a negative correlation in fluctuating environments (Schou et al. 2014), opposing the positive association reported in studies conducted in constant environments (Hoffmann and Parsons 1989; Bubliy and Loeschcke 2005; Bubliy et al. 2012). Another recent study showed that *D. melanogaster* have an afternoon peak in locomotor activity when tested in cycling light and temperature environments, which was absent in environments where temperature (but not light) was uniform (Vanin et al. 2012). This revealed a previously unknown link between temperature variation and circadian rhythms in this species. Furthermore, simply imposing thermal variation during the development of *D. melanogaster* is also sufficient to produce a range of phenotypic changes (Economos and Lints 1986), including smaller body sizes (Pétavy et al. 2004; Kjærsgaard et al. 2012; Czarnoleski et al. 2013) and altered development (Pétavy et al. 2001), relative to populations that developed in constant thermal environments with the same mean temperature. Indeed, such plasticity may even be sensitive to the predictability of the thermal fluctuations (Manenti et al. 2014). Overall, these results

demonstrate that expectations based on experimental populations that evolved in constant environments, and even for natural populations that are reared at constant temperatures, need not be readily generalizable to environments where temperature varies.

Conclusions

The basic expectation of experimental natural selection—that populations should adapt to their new experimental environment—was fulfilled in our study. Our expectations of opposing trajectories of phenotypic change were not observed for any of the measured traits, however. Although it remains unclear to what extent laboratory selection drove these patterns, and whether the different traits were part of a correlated response, the evidence strongly implies that tolerance to rapid temperature fluctuations was a target of selection in each environment. This and a handful of other studies clearly demonstrate that temporal variation in thermal stress, at a variety of scales, may lead to contrasting phenotypic outcomes from studies where temperature is kept constant. Because of the default use of constant temperatures in most previous work, it remains largely unknown how important temporal variation is for shaping the adaptive response to thermal stress in *D. melanogaster* and other species. Considering that environmental heterogeneity is the norm in natural populations, it is clear that further work is needed to address these important issues.

ACKNOWLEDGMENTS

The authors are especially grateful to the enormous efforts of the technicians—D. Martynow, G. Aleti, S. Widhalm, S. Lilja, N. O. Mompel, R. Weiss, A. Macon, S. Lehrner, and A. Gaidziza—who collected the data for the phenotypic assays. They give special thanks to P. Prasser for the morphological measurements, A. Futschik for statistical advice, P. Orozco-terWengel and V. Nolte for assistance in the design of the phenotypic assays, and D. Fabian for critical discussions regarding data interpretation. This work was supported by a Ph.D. fellowship from the Vetmeduni Vienna, the Austrian Science Fund (FWF, W1225-B20), and the European Research Council (ERC) grant “ArchAdapt” awarded to CS. The authors declare that they have no competing interests.

DATA ARCHIVING

The doi for our data is <http://dx.doi.org/10.5061/dryad.r8mb0/1>.

LITERATURE CITED

Anderson, A. R., A. A. Hoffmann, and S. W. McKechnie. 1991. Response to selection for rapid chill-coma recovery in *Drosophila melanogaster*: physiology and life-history traits. *Genet. Res.* 85:15–22.

- Anderson, W. W. 1972. Genetic divergence in body size among experimental populations of *Drosophila pseudoobscura* kept at different temperatures. *Evolution* 27:278–284.
- Andrewartha, H. G., and L. C. Birch. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago, IL.
- Angilletta, M. J. 2009. Thermal adaptation: a theoretical and empirical synthesis. Oxford University Press, Oxford, UK.
- Angilletta, M. J., B. S. Cooper, M. S. Schuler, and J. G. Boyles. 2002. The evolution of thermal physiology in endotherms. *J. Therm. Biol.* 27:249–268.
- Arnold, S. 1992. Constraints on phenotypic evolution. *Am. Nat.* 140:85–107.
- Bates, D., M. Maechler, B. Bolker, and S. Walker. 2013. lme4: linear mixed-effects models using Eigen and S4. R package version 1.0-5, <http://CRAN.R-project.org/package=lme4>
- Beardmore, J., and L. Levine. 1963. Fitness and environmental variation. I. A study of some polymorphic populations of *Drosophila pseudoobscura*. *Evolution* 17:121–129.
- Berrigan, D., and A. Hoffmann. 1998. Correlations between measures of heat resistance and acclimation in two species of *Drosophila* and their hybrids. *Biol. J. Linn. Soc.* 64:449–462.
- Bijlsma, R., and V. Loeschcke. 2005. Environmental stress, adaptation and evolution: an overview. *J. Evol. Biol.* 18:744–749.
- Bubli, O. A., and V. Loeschcke. 2005. Correlated responses to selection for stress resistance and longevity in a laboratory population of *Drosophila melanogaster*. *J. Evol. Biol.* 18:789–803.
- Bubli, O. A., T. N. Kristensen, V. Kellermann, and V. Loeschcke. 2012. Plastic responses to four environmental stresses and cross-resistance in a laboratory population of *Drosophila melanogaster*. *Funct. Ecol.* 26:245–253.
- Cavicchi, S., D. Guerra, V. Natali, C. Pezzoli, and G. Giorgi. 1989. Temperature-related divergence in experimental populations of *Drosophila melanogaster*. II. Correlation between fitness and body dimensions. *J. Evol. Biol.* 2:235–251.
- Cavicchi, S., D. Guerra, V. LaTorre, and R. B. Huey. 1995. Chromosomal analysis of heat-shock tolerance in *Drosophila melanogaster* evolving at different temperatures in the laboratory. *Evolution* 49:676–684.
- Chakir, M., A. Chafik, B. Moreteau, P. Gibert, and J. R. David. 2002. Male sterility thermal thresholds in *Drosophila*: *D. simulans* appears more cold-adapted than its sibling *D. melanogaster*. *Genetica* 114:195–205.
- Chen, J., V. Nolte, and C. Schlötterer. 2015. Temperature stress mediates decanalization and dominance of gene expression in *Drosophila melanogaster*. *PLoS Genet.* 11(6): e1005315.
- Colinet, H., and D. Renault. 2012. Metabolic effects of CO₂ anaesthesia in *Drosophila melanogaster*. *Biol. Lett.* 8:1050–1054.
- Cooper, B. S., M. Czarnoleski, and M. J. Angilletta. 2010. Acclimation of thermal physiology in natural populations of *Drosophila melanogaster*: a test of an optimality model. *J. Evol. Biol.* 23:2346–2355.
- Cooper, B. S., J. M. Tharp, I. I. Jernberg, and M. J. Angilletta. 2012. Developmental plasticity of thermal tolerances in temperate and subtropical populations of *Drosophila melanogaster*. *J. Therm. Biol.* 37:211–216.
- Coyne, J. A., and E. Beecham. 1987. Heritability of two morphological characters within and among natural populations of *Drosophila melanogaster*. *Genetics* 117:727–737.
- Czarnoleski, M., B. S. Cooper, J. Kierat, and M. J. Angilletta. 2013. Flies developed small bodies and small cells in warm and in thermally fluctuating environments. *J. Exp. Biol.* 216:2896–2901.
- Dahlgard, J., and A. a. Hoffmann. 2000. Stress resistance and environmental dependency of inbreeding depression in *Drosophila melanogaster*. *Conserv. Biol.* 14:1187–1192.
- David, J. R., P. Gibert, E. Pla, G. Petavy, D. Karan, and B. Moreteau. 1998. Cold stress tolerance in *Drosophila*: analysis of chill coma recovery in *D. Melanogaster*. *J. Therm. Biol.* 23:291–299.
- Davidson, A. J., F. Aujard, B. London, M. Menaker, and G. D. Block. 2003. Thermochron iButtons: An inexpensive method for long-term recording of core body temperature in untethered animals. *J. Biol. Rhythms* 18:430–432.
- De Miranda, J. R., and P. Eggleston. 1988. Larval competition in *Drosophila melanogaster*. II. Comparing biological and competitive parameters. *Heredity* 60:213–219.
- De Rosario-Martinez, H. 2013. phia: post-hoc interaction analysis. R package version 0.1-5, <http://cran.r-project.org/web/packages/phia/index.html>
- Economos, A. C., and F. A. Lints. 1986. Developmental temperature and life span in *Drosophila melanogaster*. II. Oscillating temperature. *Gerontology* 32:28–36.
- Ehiobu, N. G., M. E. Goddard, and J. F. Taylor. 1989. Effect of rate of inbreeding on inbreeding depression in *Drosophila melanogaster*. *Theor. Appl. Genet.* 77:123–127.
- Feder, M. E., and G. E. Hofmann. 1999. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu. Rev. Physiol.* 61:243–282.
- Feder, M. E., S. P. Roberts, and A. C. Bordelon. 2000. Molecular thermal telemetry of free-ranging adult *Drosophila melanogaster*. *Oecologia* 123:460–465.
- Fox, J., and S. Weisberg. 2011. An {R} companion to applied regression. 2nd ed. Sage, Thousand Oaks, CA.
- Fragata, I., P. Simões, M. Lopes-Cunha, M. Lima, B. Kellen, M. Bárbaro, J. Santos, M. R. Rose, M. Santos, and M. Matos. 2014. Laboratory selection quickly erases historical differentiation. *PLoS One* 9: e96227.
- Frankham, R., and D. A. Loebel. 1992. Modeling problems in conservation genetics using captive *Drosophila* populations: rapid genetic adaptation to captivity. *Zoo Biol.* 342:333–342.
- Franssen, S. U., V. Nolte, R. Tobler, and C. Schlötterer. 2014. Patterns of linkage disequilibrium and long range hitchhiking in evolving experimental *Drosophila melanogaster* populations. *Mol. Biol. Evol.* 32:495–509.
- Garland, T., and M. R. Rose, eds. 2009. Experimental evolution: concepts, methods, and applications of selection experiments. University of California Press, Berkeley and Los Angeles, CA.
- Gilchrist, A. S., and L. Partridge. 1999. A comparison of the genetic basis of wing size divergence in three parallel body size clines of *Drosophila melanogaster*. *Genetics* 153:1775–1787.
- Griffiths, J. A., M. Schiffer, and A. A. Hoffmann. 2005. Clinal variation and laboratory adaptation in the rainforest species *Drosophila birchii* for stress resistance, wing size, wing shape and development time. *J. Evol. Biol.* 18:213–222.
- Harshman, L., and A. Hoffmann. 2000. Laboratory selection experiments using *Drosophila*: what do they really tell us? *Trends Ecol. Evol.* 15:32–36.
- Hartl, D. L., and H. Jungen. 1979. Estimation of average fitness of populations of *Drosophila melanogaster* and the evolution of fitness in experimental populations. *Evolution* 33:371–380.
- Haymer, D., and D. Hartl. 1983. The experimental assessment of fitness in *Drosophila*. II. A comparison of competitive and noncompetitive measures. *Genetics* 104:343–352.
- Hoffmann, A. A., and M. J. Hercus. 2000. Environmental stress as an evolutionary force. *Bioscience* 50:217–226.
- Hoffmann, A. A., and P. A. Parsons. 1989. Selection for increased desiccation resistance in *Drosophila melanogaster*: additive genetic control and correlated responses for other stresses. *Genetics* 122:837–845.

- . 1991. Evolutionary genetics and environmental stress. Oxford University Press, Oxford, UK.
- Hoffmann, A. A., A. Anderson, and R. Hallas. 2002. Opposing clines for high and low temperature resistance in *Drosophila melanogaster*. *Ecol. Lett.* 5:614–618.
- Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6:65–70.
- Huey, R. B., and F. Rosenzweig. 2008. Laboratory evolution meets catch-22. Pp. 671–702 in T. Garland and M. R. Rose, eds. *Experimental evolution: concepts, methods, and applications of selection experiments*. University of California Press, Berkeley and Los Angeles, CA.
- Huey, R. B., L. Partridge, and K. Fowler. 1991. Thermal sensitivity of *Drosophila melanogaster* responds rapidly to laboratory natural selection. *Evolution* 45:751–756.
- James, A. C., and L. Partridge. 1995. Thermal evolution of rate of larval development *Drosophila melanogaster* in laboratory and field populations. *Genetics* 130:315–330.
- . 1998. Geographic variation in competitive ability in *Drosophila melanogaster*. *Am. Nat.* 151:530–537.
- James, A. C., R. Azevedo, and L. Partridge. 1995. Cellular basis and developmental timing in a size cline of *Drosophila melanogaster*. *Genetics* 140:659–666.
- . 1997. Genetic and environmental responses to temperature of *Drosophila melanogaster* from a latitudinal cline. *Genetics* 146:881–890.
- Karan, D., J. P. Morin, B. Moreteau, and J. R. David. 1998. Body size and developmental temperature in *Drosophila melanogaster*: analysis of body weight reaction norm. *J. Therm. Biol.* 23:301–309.
- Kingsolver, J. G., and R. B. Huey. 2008. Size, temperature, and fitness: three rules. *Evol. Ecol. Res.* 10:251–268.
- Kjærsgaard, A., N. Le, D. Demontis, Z. K. Novicic, V. Loeschcke, and C. Pertoldi. 2012. The effect of developmental temperature fluctuation on wing traits and stressed locomotor performance in *Drosophila melanogaster*, and its dependency on heterozygosity. *Evol. Ecol. Res.* 14:803–819.
- Knight, G. R., and A. Robertson. 1957. Fitness as a measurable character in *Drosophila*. *Genetics* 42:524–530.
- Kraaijeveld, A. R., and H. C. J. Godfray. 1997. Trade-off between parasitoid resistance and larval competitive ability in *Drosophila melanogaster*. *Nature* 389:278–280.
- Krebs, R. A., S. P. Roberts, B. R. Bettencourt, and M. E. Feder. 2001. Changes in thermotolerance and Hsp70 expression with domestication in *Drosophila melanogaster*. *J. Evol. Biol.* 14:75–82.
- Latter, B., and J. Mulley. 1995. Genetic adaptation to captivity and inbreeding depression in small laboratory populations of *Drosophila melanogaster*. *Genetics* 139:255–266.
- Lewontin, R. 1955. The effects of population density and composition on viability in *Drosophila melanogaster*. *Evolution* 9:27–41.
- Long, T. 1970. Genetic effects of fluctuating temperature in populations of *Drosophila melanogaster*. *Genetics* 66:401–416.
- Manenti, T., J. G. Sørensen, N. N. Moghadam, and V. Loeschcke. 2014. Predictability rather than amplitude of temperature fluctuations determines stress resistance in a natural population of *Drosophila simulans*. *J. Evol. Biol.* 27:2113–2122.
- Mather, K., and P. Caligari. 1981. Competitive interaction in *Drosophila melanogaster*. II. Measurement of competition. *Heredity* 46:239–254.
- Matos, M., M. R. Rose, M. Pité, C. Rego, and T. Avelar. 2000. Adaptation to the laboratory environment in *Drosophila subobscura*. *J. Evol. Biol.* 13:9–19.
- Matos, M., T. Avelar, and M. Rose. 2002. Variation in the rate of convergent evolution: adaptation to a laboratory environment in *Drosophila subobscura*. *J. Evol. Biol.* 15:673–682.
- Mery, F., and T. J. Kawecki. 2003. A fitness cost of learning ability in *Drosophila melanogaster*. *Proc. Biol. Sci.* 270:2465–2469.
- Mori, N., and M. T. Kimura. 2008. Selection for rapid and slow recovery from chill- and heat-coma in *Drosophila melanogaster*. *Biol. J. Linn. Soc.* 95:72–80.
- Morrison, W. W., and R. Milkman. 1976. Modification of heat resistance in *Drosophila* by selection. *Nature* 273:49–50.
- Mueller, L. D. 1991. Ecological determinants of life-history evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 332:25–30.
- Nilson, T. L., B. J. Sinclair, and S. P. Roberts. 2006. The effects of carbon dioxide anesthesia and anoxia on rapid cold-hardening and chill coma recovery in *Drosophila melanogaster*. *J. Insect Physiol.* 52:1027–1033.
- Nuzhdin, S. V., J. D. Fry, and T. F. C. Mackay. 1995. Polygenic mutation in *Drosophila melanogaster*: the causal relationship of bristle number to fitness. *Genetics* 139:861–872.
- Orozco-terWengel, P., M. Kapun, V. Nolte, R. Kofler, T. Flatt, and C. Schlötterer. 2012. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. *Mol. Ecol.* 21:4931–4941.
- Partridge, L., B. Barrie, K. Fowler, and V. French. 1994. Evolution and development of body size and cell size in *Drosophila melanogaster* in response to temperature. *Evolution* 48:1269–1276.
- Partridge, L., R. Langelan, K. Fowler, B. Zwaan, and V. French. 1999. Correlated responses to selection on body size in *Drosophila melanogaster*. *Genet. Res.* 74:43–54.
- Pedersen, K. S., T. N. Kristensen, and V. Loeschcke. 2005. Effects of inbreeding and rate of inbreeding in *Drosophila melanogaster*—Hsp70 expression and fitness. *J. Evol. Biol.* 18:756–762.
- Pétavy, G., J. R. David, P. Gibert, and B. Moreteau. 2001. Viability and rate of development at different temperatures in *Drosophila*: a comparison of constant and alternating thermal regimes. *J. Therm. Biol.* 26:29–39.
- Pétavy, G., J. R. David, V. Debat, P. Gibert, and B. Moreteau. 2004. Specific effects of cycling stressful temperatures upon phenotypic and genetic variability of size traits in *Drosophila melanogaster*. *Evol. Ecol. Res.* 6:1–18.
- R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Robertson, F. W. 1962. Changing the relative size of the body parts of *Drosophila* by selection. *Genet. Res.* 3:169–180.
- Santos, M., K. Fowler, and L. Partridge. 1992. On the use of tester stocks to predict the competitive ability of genotypes. *Heredity* 3:489–495.
- Santos, M., D. Brites, and H. Laayouni. 2006. Thermal evolution of pre-adult life history traits, geometric size and shape, and developmental stability in *Drosophila subobscura*. *J. Evol. Biol.* 19:2006–2021.
- Santos, M., I. Fragata, J. Santos, P. Simões, A. Marques, M. Lima, and M. Matos. 2010. Playing Darwin. Part B. 20 years of domestication in *Drosophila subobscura*. *Theory Biosci.* 129:97–102.
- Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9:671–675.
- Schou, M. F., T. N. Kristensen, V. Kellermann, C. Schlötterer, and V. Loeschcke. 2014. A *Drosophila* laboratory evolution experiment points to low evolutionary potential under increased temperatures likely to be experienced in the future. *J. Evol. Biol.* 27:1859–1968.
- Sgrò, C. M., and L. Partridge. 2000. Evolutionary responses of the life history of wild-caught *Drosophila melanogaster* to two standard methods of laboratory culture. *Am. Nat.* 156:341–353.
- Siddiqui, W. H., and C. A. Barlow. 1972. Population growth of *Drosophila melanogaster* (Diptera: Drosophilidae) at constant and alternating temperatures. *Ann. Entomol. Soc. Am.* 65:993–1001.

- Simões, P., J. Santos, I. Fragata, L. D. Mueller, M. R. Rose, and M. Matos. 2008. How repeatable is adaptive evolution? The role of geographical origin and founder effects in laboratory adaptation. *Evolution* 62:1817–1829.
- Sørensen, J. G., and V. Loeschcke. 2007. Studying stress responses in the post-genomic era: its ecological and evolutionary role. *J. Biosci.* 32:447–456.
- Sørensen, J. G., T. N. Kristensen, and V. Loeschcke. 2003. The evolutionary and ecological role of heat shock proteins. *Ecol. Lett.* 6:1025–1037.
- Tobler, R., S. U. Franssen, R. Kofler, P. Orozco-terWengel, V. Nolte, J. Hermisson, and C. Schlötterer. 2014. Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. *Mol. Biol. Evol.* 31:364–375.
- Trotta, V., F. C. F. Calboli, M. Ziosi, and S. Cavicchi. 2007. Fitness variation in response to artificial selection for reduced cell area, cell number and wing area in natural populations of *Drosophila melanogaster*. *BMC Evol. Biol.* 7(Suppl. 2):S10.
- Turner, T. L., A. D. Stewart, A. T. Fields, W. R. Rice, and A. M. Tarone. 2011. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation *Drosophila melanogaster*. *PLoS Genet.* 7:e1001336.
- Vanin, S., S. Bhutani, S. Montelli, P. Menegazzi, E. W. Green, M. Pegoraro, F. Sandrelli, R. Costa, and C. P. Kyriacou. 2012. Unexpected features of *Drosophila* behavioural rhythms under natural conditions. *Nature* 484:371–375.
- Vijendravarma, R. K., A. R. Kraaijeveld, and H. C. J. Godfray. 2009. Experimental evolution shows *Drosophila melanogaster* resistance to a microsporidian pathogen has fitness costs. *Evolution* 63:104–114.
- Williams, B. R., B. van Heerwaarden, D. K. Dowling, and C. M. Sgrò. 2012. A multivariate test of evolutionary constraints for thermal tolerance in *Drosophila melanogaster*. *J. Evol. Biol.* 25:1415–1426.
- Yamazaki, T. 1984. Measurement of fitness and its components in six laboratory strains of *Drosophila melanogaster*. *Genetics* 108:201–211.
- Yeaman, S., Y. Chen, and M. C. Whitlock. 2010. No effect of environmental heterogeneity on the maintenance of genetic variation in wing shape in *Drosophila melanogaster*. *Evolution* 64:3398–3408.

Associate Editor: I. Dworkin
Handling Editor: R. Shaw

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Supplemental Table S1. Results of ANOVA (car R package, see methods) testing for all traits in both assayed generations (Cold 34 & Hot 59 or Cold 44 & Hot 75).

Supplemental Table S2. Results of ANOVA (car R package, see methods) testing for all traits except fitness after adjusting all values and adding the assayed generations as an additional fixed factor (Assay).

Supplemental Table S3. Results of contrasts (phia R package, see methods) testing all pairwise population differences given sex and CGE for all traits in both assayed generations (Cold 34 & Hot 59 or Cold 44 & Hot 75).

Supplemental Table S4. Results of contrasts (phia R package, see methods) testing for differences across the two assayed generations (Cold 34 & Hot 59 or Cold 44 & Hot 75) for each trait given population, sex and CGE. CGE = common garden experiment (constant 15% C, constant 23% C), Assay I = Cold 34 & Hot 59, Assay II = Cold 44 & Hot 75.

Supplemental Figure 1. Boxplot showing the maximum (red boxes), minimum (blue boxes), mean (green boxes) and range (gray boxes) of daily temperatures by month for the native habitat of the progenitor population Póvoa de Varzim, Portugal.

Supplemental Figure 2. A 24hr time-series showing the temperature cycles maintained in the hot (red) and cold (blue) experimental environments.

Supplemental Figure 3. Box plot showing the adjusted of recovery time (2 left-most columns) and knock down time (2 right-most columns) from acute cold and heat stress, respectively.

Supplemental Figure 4. Box plot of adjusted developmental time for hot (red) and cold (blue populations). The x-labels denote the assay number (I = F34 cold & F59 hot, II = F43 cold & F75 hot).

Supplemental Figure 5. Box plot showing the adjusted of wing area (2 left-most columns) and rear tibia length (4 right-most columns) from acute cold and heat stress, respectively. Sex and population are listed in the panels, CGE and assay (I = F34 cold & F59 hot, II = F43 cold & F75 hot) in the & labels.

Supplemental Figure 6. Enrichment of putatively selected SNPs from the cold (left panel) and hot (right panel) experimental populations in resistance genes unique to heat stress (red line), cold stress (blue line), or common to both stresses (black line).