Pervasive fitness trade-offs revealed by rapid adaptation in large experimental populations of *Drosophila melanogaster*

M.C. Bitter*[1], S. Greenblum*[1,2], S. Rajpurohit*[3,4], A.O. Bergland[1,5], J.A. Hemker[1], N.J. Betancourt[3], S. Tilk[1], S. Berardi[3], H. Oken, P. Schmidt+[3], and D.A. Petrov+[1,6]

1 – Department of Biology, Stanford University, Stanford, CA, USA
2 – DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
3 – Department of Biology, University of Pennsylvania, Philadelphia, PA, USA
4 – Division of Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Gujarat, India
5 – Department of Biology, University of Virginia, Charlottesville, VA, USA
6 – Chan Zuckerberg Biohub, San Francisco, CA, USA

* equal contribution
+ equal contribution

Corresponding authors: M.C.B (mcbitter@stanford.edu), P.S. (schmidtp@upenn.edu), and D.A.P (dpetrov@stanford.edu)

Abstract

Life-history trade-offs are an inherent feature of organismal biology that evolutionary theory posits play a key role in patterns of divergence within and between species. Efforts to quantify trade-offs are largely confined to phenotypic measurements and the identification of negative genetic-correlations among fitness-relevant traits. Here, we use time-series genomic data collected during experimental evolution in large, genetically diverse populations of *Drosophila melanogaster* to directly measure the manifestation of trade-offs in response to temporally fluctuating selection pressures on ecological timescales. Specifically, we quantify the genome-wide signal of antagonistic pleiotropy suggestive of trade-offs between reproduction and stress tolerance. We further identify a putative role of two cosmopolitan inversions in these trade-offs, and show that loci responding to selection during lab-based, reproduction selection exhibit signals of fluctuating selection in an outdoor mesocosm exposed to natural environmental conditions. Our results demonstrate the utility of time-series genomic data in revealing the presence and genomic architecture underlying fitness trade-offs, and add

credence to models positing a role of generic life history trade-offs in the maintenance of variation in natural populations.

Introduction

A fundamental tenet of evolutionary theory is that trait adaptation is restricted by trade-offs: the cost to individual fitness when an advantageous change in one trait occurs at the detriment to another (1,2). Trade-offs can be observed across closely related species, such as those between beak and body size among Darwin's finches (3,4). They can also arise as a consequence of microevolutionary processes among populations of the same species; for instance, the trade-offs to survival associated with distinct reproductive strategies among populations of salmon (5). Trade-offs oftentimes emerge, and are most readily studied, in the context of life-history traits, and form the basis of the theory of life history evolution (2). Ultimately, a key theoretical implication of trade-offs is that they maintain variation in fitness by constraining the simultaneous optimization of traits important for different environmental conditions (2,6–8).

At the molecular level, trade-offs likely reflect antagonistic pleiotropy, a phenomenon in which an allele that enhances one fitness-related trait may simultaneously impede another. Crucially however, since only a subset of traits may positively impact fitness in a given environment, selection on the allele is not net-neutral, but instead flips direction when the environment changes (7–9). Thus, a widely used method for inferring the presence of antagonistic pleiotropy and trade-offs is to compute genetic (co) variances using phenotypic data (10–12). Still, a more explicit way to quantify trade-offs, and also link them to the maintenance of variation, would be to directly measure population allele frequencies over time while selective pressures fluctuate. Indeed, several recent studies observed genome-wide fluctuations in allele frequencies through time, and across several distinct systems, suggesting underlying trade-offs (13–18).

A fundamental life-history trade-off is the balance between directing energy either towards survival or reproduction whereby periods of stressful environmental conditions favor somatic maintenance at the cost of reproductive investment (19). This tradeoff is clearly apparent among populations of *Drosophila melanogaster* inhabiting temperate environments, a well-studied system in which increased food accessibility and warmer climates during summer spur exponential population growth, followed in winter by harsher conditions and population collapse (20–22). These cyclical boom-bust population dynamics occur over approximately 10 generations and coincide with the evolution of several classic life history traits: traits conferring increased reproduction (e.g., fecundity, faster developmental rates) are favored during spring and summer as populations expand, followed by an increase in stress tolerance traits (e.g.,

desiccation and starvation resistance) throughout the harsh winter and subsequent population collapse (23–25). It is likely that trade-offs underpin these contrasting patterns of adaptation, as negative genetic correlations have been quantified between seasonally evolving reproduction and stress tolerance traits (10,26–28).

Recent genome-wide sequencing of wild populations, and outbred populations evolved in semi-natural mesocosms, has revealed evidence that hundreds of independent loci exhibit signatures of selection concurrent to these patterns of rapid phenotypic evolution, many of which switch in sign in concert with the fluctuating environment and evolving traits (13–16). Strikingly, despite differences in the particular environmental conditions in which patterns of evolutionary change were assayed (e.g., different observation years and geographic locations), there is high repeatability in the genomic regions subject to fluctuating selection across space and time (13–15). This suggests that the shared genomic signals of temporally varying selection could be underpinned by a generic response elicited by fundamental, life-history trade-offs. However, in the context of the natural environments in which these studies were conducted, it is exceptionally challenging to decompose which specific environmental pressures and fitness-relevant traits may interact to produce trade-offs and maintain variation in the system. For example, it is unknown whether the magnitude of phenotypic and genomic evolutionary change can be solely attributed to shifting ecological pressures, such as population expansion and collapse; or, whether the interaction of intraspecific ecological processes and changing abiotic conditions (e.g., temperature and precipitation) is necessary to solicit such dramatic adaptive responses and reversions in allele frequency trends. Decoupling these selective forces, and directly linking them to specific traits and patterns of allele frequency change, would be more robustly addressed in the context of simplified and controlled experimental settings.

Here, we used two lab-based experimental evolution studies to more explicitly link key components of life history evolution to patterns of fluctuating selection and maintenance of variation in *D. melanogaster*. In the first of these studies, we assayed the impact of ecological shifts, independent from any environmental selection. To do so, we monitored patterns of genomic variation in genetically diverse, replicate populations housed in a controlled, indoor environment throughout eleven discrete (non-overlapping) generations of population expansion, and a single bout of population truncation (Fig. 1; Fig S1). Our replicate populations were derived from a reconstituted, outbred population initiated from 145 inbred (DGRP) lines (Methods; Supplementary Data File 1), and evolution proceeded in four replicate cages (2 m^3). Food availability in the replicate cages was commensurate with increasing population sizes during the expansion phase, generating a selective regime that eliminated the impact of intraspecific competition (e.g. for food or egg laying space) and favored the reproduction-associated traits expected when resources are abundant, such as fecundity and developmental

rate (24). Contrarily, population truncation was induced by removing all food and water from our replicate cages and sampling the surviving individuals as the replicates collapsed. This mirrors dynamics of a population that has exceeded carrying capacity and, in effect, selects for the ability to withstand depleted resources via stress-tolerance mechanisms (29), which in our experiment is most likely to be desiccation resistance (30). We hypothesized that these contrasting selective regimes (hereafter, 'reproduction' and 'stress-tolerance' selection) would reveal trade-offs between life-history variation, whereby antagonistically pleiotropic alleles favored during sustained population expansion would become selected against during truncation.

The second experiment leveraged an independent inbred panel to construct a genetically diverse, outbred population, which was split into sets of indoor cage (N = 10) and outdoor mesocosm (N = 12) replicates. Through this, we used the same mapping population (thereby eliminating any confounding impacts of differences in linkage disequilibrium) to directly test whether loci identified under isolated selection for fast reproduction in the lab display predictable patterns of selection and antagonistic pleiotropy during adaptation in complex natural environments. In concert, our results provide evidence for: (1) strong, parallel selection induced under sustained population expansion (2) genome-wide tradeoffs associated with contrasting selective regimes of population expansion and truncation, (3) the role of cosmopolitan inversions underpinning fitness trade-offs during population expansion and collapse and (4) relevance of the alleles responding to selection for increased fecundity in a lab-based experiment to patterns of selection during adaptation to natural environmental fluctuations.

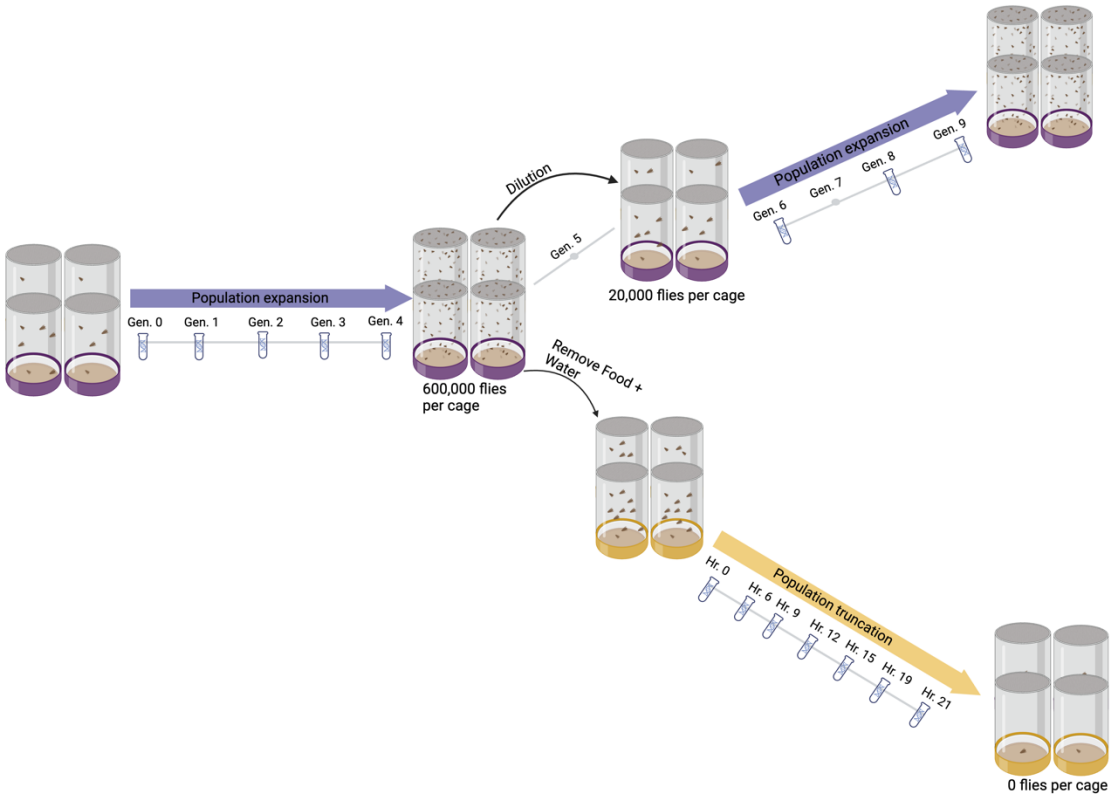

Figure 1. Schematic of lab-based, population expansion-truncation selection experiment | Four replicate, 2 m^3 cages were seeded with a genetically diverse, outbred population generated via four generations of recombination (N = 145 inbred lines). Evolution proceeded in discrete generations, whereby the amount of food was doubled every generation until replicate populations reached a census size of approximately 600,000 flies per cage (generation four). At this point, eggs from the generation 5 cohort were collected and all food and water was removed from the replicate cages. The adult cohort of the generation 4 flies was then sampled at seven time points (hours 0-21) as the absence of resources collapsed the populations. The generation 5 eggs were used to re-seed the replicate cages, which underwent continued population expansion for four additional generations. DNA in vials denotes generations of expansion, and hours of truncation, during which pooled samples were collected for pooled allele frequency calculation. Schematic generated with bioRender (https://www.biorender.com/).

Results

Selection during population expansion elicits genome-wide signals of strong, parallel selection

We observed parallel shifts in allele frequencies across the four replicate populations throughout the progression of population expansion, indicating adaptive responses to the shared selective pressures imposed in our experimental system. We quantified these systematic

allelic shifts using several approaches. We first computed genome-wide divergence as average F_{ST} across all segregating sites pairwise among all expansion samples and used these data to carry out several tests. F_{ST} divergence from the generation 0 populations increased steadily within each cage through time, readily exceeding differentiation between biological replicates ($F = 172.2$; p-value < 0.01; Fig. 2A & Fig. S2). We then used pairwise F_{ST} values as a distance metric to create multi-dimensional scaling (MDS) plots, in which divergence between samples is represented as distance among the points in a 2-D plane. Coloring samples based on the replicate cage identity indicated that genome-wide allele frequencies in one replicate were perturbed, shifting its points from the remaining three replicates throughout expansion (Fig. 2B). We hypothesize this was likely a result of potential bottlenecks during replicate cage founding. However, despite this offset, samples from all cages appeared to shift across the 2-D plane in the same direction through time, suggesting that the replicate populations experienced parallel shifts in genome-wide allele frequencies (Fig 2C and Fig S3; a trend that was also detected using principal component analysis; Fig. S4).

To more rigorously quantify the systematic movement of samples through time observed in the two-dimensional MDS space, we translated coordinates for each cage such that the centroid of all generation 0 samples was centered at the origin. We then used these translated points to fit a simple linear regression model to samples from expansion generations 0-4 (Methods; Fig. S5). The resulting axis represents the primary axis of variation in the 2D plane during early expansion for each replicate. We hypothesized that if sustained, directional selection imposed by population expansion was a primary driver of patterns of genomic variation, samples from the remaining expansion generations (generations 6-9) would continue to proceed along the established axis of variation in the same direction. Indeed, projection of late expansion samples onto this axis of variation indicated a significant correlation of sample collection generation and distance along the axis (Fig 2D.; significance of correlation determined via permutations). When segregating this analysis by chromosomal arm, however, this parallelism was only evident on chromosomal arms 2L, 3L, and 3R ; while correlations derived from SNPs on 2R and X were indistinguishable from permuted values (Fig. 2D). We re-sampled SNPs across the genome iteratively to match the number present on each chromosomal arm to confirm that the variation in parallelism was not a technical artifact. In each case, our sub-sampling yielded correlations that matched the genome-wide trend, indicating that differences observed among chromosomal arms (Fig. 2D) reflect systematic differences in the behavior of loci across the genome (Fig. S6, Table S2).

To determine the distribution of SNPs (N = 1.7 M SNPs) contributing to the parallel patterns of divergence across cages, a generalized linear model was fit to allele frequencies from all expansion samples (generation 0-9). This model assessed the significance of the linear

relationship between allele frequency and generation of sampling across all cages, using a quasibinomial error model to reduce false positive associations (15,16,31). While an association between allele frequency and sampling timepoint within a single cage may represent either drift or selection, a significant parallel association across all four replicate cages indicates allele frequency trajectories that are both predictable over time and parallel across populations. In this case, selection (or linked selection) is the more parsimonious explanation. A total of 389,588 SNPs (22.9 % of all sites), across all chromosomal arms, showed significant parallelism after multiple testing correction (Benjamini-Hochberg false discovery rate <0.01 and allele frequency change > 2%). These SNPs showed systematic, directional selection across replicates throughout expansion (Fig. 2E; Fig. S7). Furthermore, the genomic distribution of these SNPs spanned all five chromosomal arms and were located both within and outside the breakpoints of five major cosmopolitan inversions that segregated at appreciable frequency (>4%) in our inbred reference panel (Fig. 2F).

We validated the parallelism inferred by GLM by implementing a leave-one-out cross validation. Specifically, we iteratively identified sets of parallel SNPs across expansion samples using a GLM fit to allele frequency data from three of the four replicate cages. We then quantified the frequency shifts of the rising allele at significant SNPs (FDR < 0.01, allele frequency shift > 2 %) in the left-out cage. In each iteration of this analysis, the left-out cage exhibited a magnitude of allele frequency change that exceeded the background allele frequency change (quantified using matched control SNPs; see Methods), and in a direction of change parallel to that observed in the other three cages (Fig 2G). The genome-wide median shifts of target SNPs across replicates ranged between 5 and 7.5%, indicating that these parallel patterns of adaptation were underpinned by strong selection, on the order of 10-20% per generation. While the per-chromosome Fst and MDS analyses described above, which focus on signal averaged across all segregating sites, suggest that non-parallel idiosyncratic movement may be the dominant force on certain chromosomes, our GLM analysis parallel SNPs can still be discovered on every chromosomal arm.

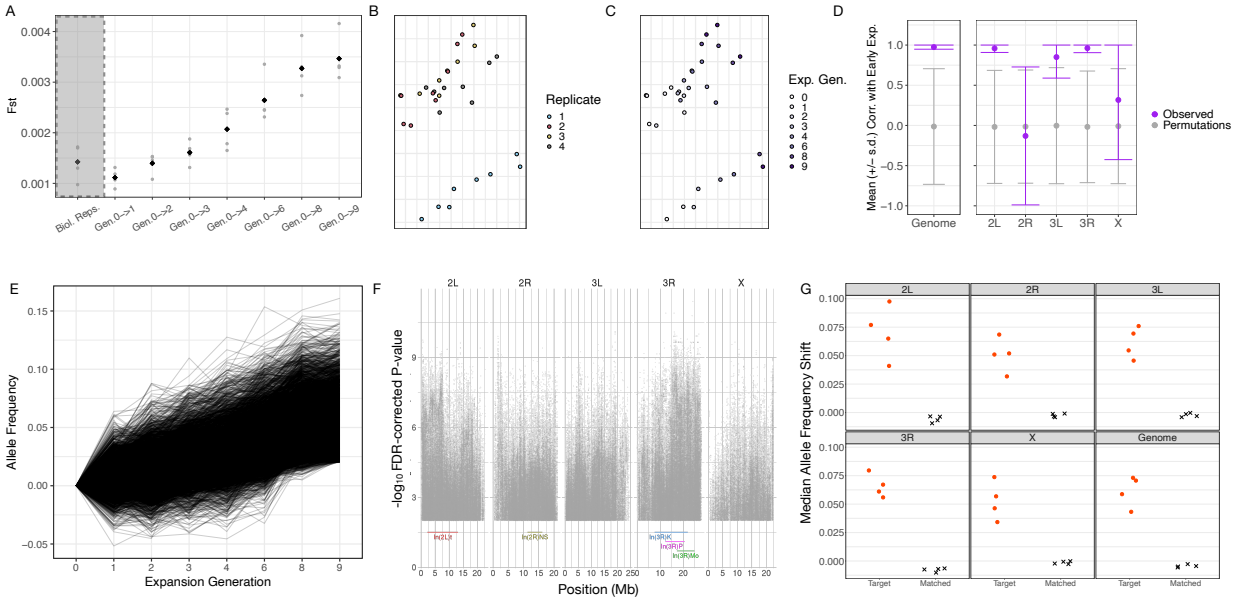

Figure 2. Adaptation under sustained population expansion elicits strong, genome-wide parallel responses | (A) Mean, genome-wide F_{ST} between biological replicates ('Biological Replicates'; same replicate/collection time point, different pooled sample/extraction of flies) and evolved replicate samples and their respective generation 0 expansion sample ('Gen. 0->n'). Grey points indicate individual replicate samples, and black diamonds the averaged value across replicates. F_{ST} differentiation from generation 0 samples increased monotonically as function of collection time ($F = 172.2$; p-value < 0.01). (B-C) MDS of F_{ST} values computed pairwise across all expansion samples, colored according to (B) replicate cage or (C) collection generation. (D) Average Pearson Correlation across replicates (+/- standard deviation) between sample expansion generation and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples (Fig. 2C). Purple points and error bars correspond to observed values (mean +/- standard deviation across cages), while grey points and error bars correspond to values derived from $N = 100$ permutations. (E) Trajectories of rising alleles at SNPs identified via GLM (FDR < 0.01 and effect size $> 2\%$) across all nine generations of expansion. (D) Genomic distribution of SNPs depicted in (E). The coordinates of the major cosmopolitan inversions segregating in our founding reference panel at greater than 4% frequency in our founding strains are depicted above the x-axis. (G) Leave-one-out cross-validation to infer replicate-specific parallelism of adaptation to sustained population expansion. Portrayed is the median shift of the rising allele for sets of target SNPs (points) identified via GLM (FDR < 0.01 and effect size $> 2\%$) in 3 of 4 cages, relative to medians derived from a matched control SNP set (X's), in the left-out replicate. Target SNP medians are colored red if the distribution of phased allele frequency shifts was significantly greater than that of matched control sites (two-tailed t-test, FDR < 0.05).

Genomic evidence of trade-offs induced by fluctuating selection across expansion and truncation

The parallel frequency shifts observed throughout expansion may be the product of three, not mutually exclusive, evolutionary dynamics: (1) directional selection in response to sustained fecundity selection, (2) adaptation to the lab environment, and/or (3) the purging of recessive deleterious alleles (i.e., negative selection) in outbred population. To disentangle these

various dynamics and, in turn, identify the presence of fitness-tradeoffs at putatively selected alleles, we leveraged our samples collected throughout truncation selection. Alleles identified during expansion that were a product of consistent lab selection and/or negative selection against unconditionally deleterious recessives should continue to show systematic, directional change throughout truncation (as their effect is not conditional on the specific treatment). However, alleles with treatment-specific behavior during expansion and truncation (i.e., moving in the opposite direction) represent those that likely underpin trade-offs between fecundity and stress tolerance selection.

We tested for evidence of context-specific behavior and trade-offs of genome-wide SNPs by re-conducting our MDS analysis of pairwise divergence values, this time including all samples collected throughout expansion and truncation. If the dominant direction of allele frequency change was sustained across both expansion and truncation (indicating sustained lab selection or purging of deleterious mutations), truncation samples would be ordered from early to late in a parallel manner as the expansion samples (Fig 2B). Instead, we observed that samples taken during truncation (initiated at expansion generation 5) shifted back towards earlier expansion samples, potentially suggesting a genome-wide reversion of allele frequencies (Fig. 3A; Fig. S8). We quantified these trends as above, translating the MDS coordinates of each truncation sample such that the centroid of all hour 0 samples was centered at the origin and then projecting them onto a the single axis linear regression model derived from early expansion points (Fig. 2D; Fig. S5). As hypothesized, correlations between collection time and distance along this axis were, genome-wide, significantly negative (Fig. 3B; Table S1). Segregating this analysis by chromosomal arm yielded more nuanced dynamics whereby 3L, 3R, and X yielded evidence of anti-parallel/reversions in allele frequencies, while 2L and 2R exhibited correlations that were insignificant relative to our permuted distributions (Fig. 3B). Again, as above, we validated that these differences among chromosomal arms were not simply a reflection of variation in number of SNPs available for measurement and reflected systematic differences in allele frequency dynamics across the genome (Fig. S9).

We garnered further evidence of trade-offs across fecundity and desiccation selection using allele frequency trajectories. Specifically, we selected those SNPs with evidence of systematic movement across replicates throughout expansion (GLM FDR < 0.01 and allele frequency shift $> 2\%$) and measured the average frequency shifts of the rising allele across cages throughout truncation (Figure 3D). If patterns of fluctuating selection and antagonistic pleiotropy dominated the data, we expected trajectories to be enriched for SNPs moving in the opposite (negative) direction during truncation. Indeed, across all expansion-identified SNPs, the ratio was significantly skewed towards those with exhibiting fluctuating selection, as

opposed to sustained directional selection, across expansion and truncation selection ($\chi^2 = 1123.6$; P-value < 0.001). These trajectories are depicted in Figure 3D, in which colored trajectories indicate those expansion-identified SNPs with additional, independent evidence of systematic movement across cages during truncation (GLM FDR < 0.05 and effect size $> 1\%$): purple for those SNPs with consistent directional selection across selection regimes, and orange for those with anti-parallel movement across regimes. As expected, this subset of SNPs (N = 24,814) exhibited even greater evidence of fluctuating selection and were 3.5 times more likely to fluctuate directions across treatments than exhibit sustained directional selection (19,158 relative to 5,656 SNPs, respectively). We assessed how consistent this dynamic was across the genome by quantifying the ratio of parallel vs. fluctuating selection genome-wide, and separately for each chromosomal arm. As expected based on the F_{ST} MDS analysis described above, there was strong evidence of enrichment for fluctuating selection on chromosomal arms 3L, 3R, and X. However, this analysis revealed an enrichment of 2L SNPs with sustained parallel selection across selection regimes, which was an undetected in our F_{ST} MDS analysis. We explored this dynamic further, using the underlying allele frequency shift distributions to compute empirical cumulative distribution functions, separately for alleles with sustained directional selection and for those with evidence of fluctuating selection (See Supp. Mat.). This verified an enrichment of alleles exhibiting fluctuating selection on arms 3L, 3R, and X, and sustained directional selection on 2L (Fig S10). In conclusion, we find evidence for the existence of pervasive, genome-wide trade-offs between fecundity and desiccation selection, as well as some evidence of sustained directional selection due to either adaptation to the lab environment or purging of the unconditionally deleterious alleles. It is important to note the patterns of fluctuating selection quantified here cannot be a spurious artifact driven by regression to the mean, as the data from which the trajectories were computed are entirely independent across expansion and truncation because they share no common datapoints.

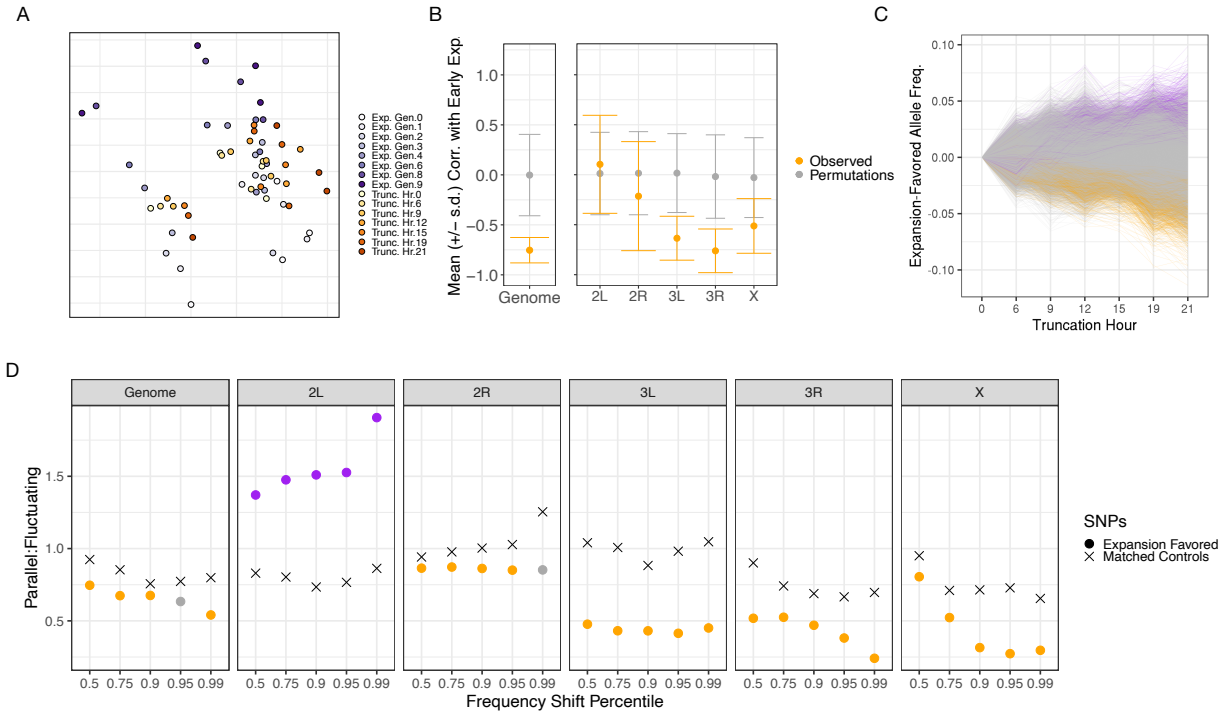

Figure 3. Genome-wide signal of fitness trade-offs between population expansion and truncation | (A) MDS of pairwise F_{ST} values across all samples collected throughout expansion (purple-hue points) and truncation (orange-hue points), shaded according to collection time point (darker hues indicate later expansion or truncation sampling generation/hour. (B) Average Pearson correlation across replicates (+/- standard deviation) between sample expansion generation and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples (Fig. 3A). Orange points and error bars correspond to observed values (mean +/- standard deviation across cages), while grey points and error bars correspond to values derived from N = 100 permutations. (C) Trajectories of SNPs identified via GLM regression across replicates throughout expansion, measured during truncation. Trajectories are colored for SNPs that exhibited evidence of systematic allele frequency changes across replicates during truncation (GLM FDR < 0.05 and allele frequency change > 1%) and displayed sustained directional (purple) or fluctuating selection (orange). (D) Ratio of parallel to fluctuating dynamics for expansion-favored alleles during truncation, segregated by chromosomal arm and percentile of truncation allele frequency shift. Points are further colored if the ratio of parallel to fluctuating behavior differed significantly from that expected based on a matched control set (χ^2 test P-value < 0.05).

The role of inversions in reproduction/stress-tolerance trade-offs

Theoretical, and emerging empirical, research suggests a dominant role of chromosomal inversions in adaptation in natural populations (32). Accordingly, we quantified whether five major cosmopolitan inversions in *D. melanogaster* (*In(2L)t*, *In(2R)NS*, *In(3R)K*, *In(3R)P*, *In(3R)Mo*), which all occurred at starting frequency > 4% in our inbred reference panel, exhibited dynamics consistent with adaptation and, if so, trade-offs between population expansion and truncation (Supplementary Data File 3). Generalized linear regression of inversion frequencies yielded evidence that all inversions shifted systematically across replicates during expansion, with magnitude of frequency shifts ranging between 1 and 7% (Fig. 4A-E; Table S2). *In(3R)K* and *In(3R)P* were the only two inversions that exhibited evidence of trade-offs across selection regimes, whereby each inversion was systematically favored during expansion and then became selected against during truncation (Fig. 4A-B).

Given the appreciable effect of inversions on patterns of adaptation throughout the experiment, we next explored the extent to which our inference of trade-offs was solely a function of variation within, or in tight linkage of, these structural variants. Specifically, we reconducted F_{ST}-based MDS analysis on a subset of SNPs, excluding all those within, or 100 kb away from, inversion breakpoints. Through this, we observed that the signal of reversions in allele frequencies were maintained on 3L and X, but eliminated on 3R. Still, the genome-wide signal of this analysis continued to indicate parallel movement across early and late-expansion samples, and that a signal of fluctuating selection dominates patterns of genomic variation between expansion and truncation (Fig. S11). We validated that this trend was not simply a function of the reduced set of SNPs available for analysis, relative to the genome-wide panel (Fig. S12).

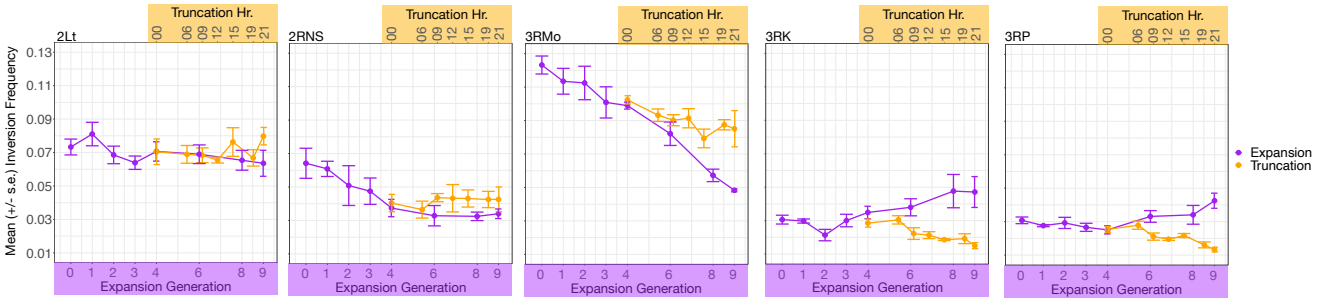

Figure 4. Role of inversions in reproduction and stress-tolerance trade-offs | Mean inversion frequency changes (+/- s.e.) across expansion generations (purple trajectories) and truncation hours (orange trajectories).

The emergence of tradeoffs in response to natural environmental fluctuations

We quantified whether alleles identified via selection under sustained population expansion in a laboratory setting can predict patterns of adaption and trade-offs in an outdoor environment, where populations adapt both to changes in population density as well as a suite of additional abiotic variables. Specifically, we leveraged data from an independent study year when we monitored patterns of genomic variation in a genetically diverse population that was split into a series of large, replicate cages maintained in both a controlled, indoor laboratory (N = 10 replicate populations), as well as outdoor mesocosms exposed to natural environmental fluctuations (N = 12 replicate populations; data previously reported in Bitter et al. 2024) (Fig. S13). As with our indoor expansion/truncation experiment described above, the population used in this experiment was derived via outbreeding an inbred reference panel, which in this case was originally collected from Linvilla Orchards, Media, PA (Supplementary Data File 2). The replicates in both environments evolved with overlapping generations, under constant food conditions, for a period of four months (see Methods). This induced rapid population expansion in each environment until a peak/stabilization of density was observed, suggesting the initiation of density control (Fig. S11; Table S7).

We first quantified the extent to which evolution in the outdoor environment is driven by the selective pressures solely associated with increasing population density (e.g., selection for increased fecundity, faster developmental rate). Specifically, we identified alleles systematically favored throughout concurrent sampling of the indoor and outdoor cages using a GLM (see Methods), and found far greater overlap than expected based on matched control SNPs (Fig. 5A). Next, we asked whether the dominant direction of selection was parallel across environments by quantifying frequency shifts in the outdoor replicates at the rising allele for each SNP identified within the indoor environment. We observed that the magnitude of allele frequency shifts was significantly greater than background allele frequency movement, and that the dominant direction of selection was conserved between environments (indicated via positive median shifts depicted in Fig. 5B; Table S6).

Finally, we tested whether the pervasive trade-offs we inferred from the previous experiment and results described above (Fig. 1-4) manifested in the outdoor mesocosms. Specifically, we leveraged an additional month of sampling in the outdoor mesocosms, during which time a population decline and ultimate collapse was observed as winter and the deterioration of abiotic conditions progressed (14). We hypothesized that, should those alleles identified in the indoor environment underpin antagonistic pleiotropy and trade-offs for fitness-relevant variation, allele frequencies should reverse in direction as the outdoor mesocosms collapsed. Indeed, we found a subtle, but significant genome-wide reversion in

allele frequencies in the outdoor cages throughout this period, which was ultimately driven by patterns of variation on 2L and 3R (Fig. 5C; Table S6).

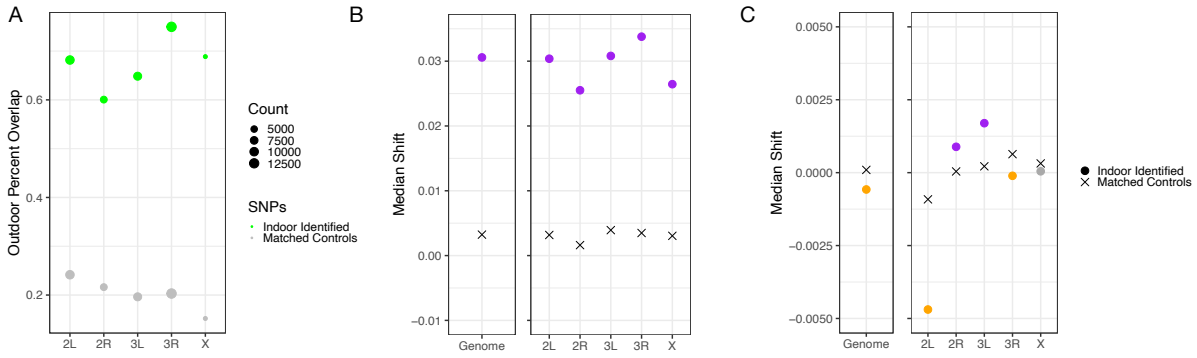

Figure 5. Selection on reproduction under controlled, lab-based conditions predicts patterns of adaptation and trade-offs in outdoor mesocosms | (A) Observed vs. expected overlap in SNPs with systematic allele frequency movement (GLM FDR < 0.01; allele frequency change > 2%) in a paired indoor and outdoor experimental evolution study. (B-C) Median shift of alleles identified during expansion in an indoor environment, quantified during population expansion (B) and collapse (C) in an outdoor mesocosm. Colored circles correspond to indoor-identified SNPs with allele frequency shift distributions that were either significantly greater (orange) or less (purple) than that observed for matched control SNPs (X's) (two-tailed t-test FDR < 0.05).

Discussion

We quantified genome-wide evidence of fitness trade-offs in response to selection on two key life-history traits in the context of ecologically realistic population boom and bust population dynamics. Our results hold relevance for the maintenance of fitness-relevant variation, the genomic basis of fitness-tradeoffs, and the interplay of ecological and evolutionary processes in natural populations.

We manipulated reproduction selection via nine generations of sustained population expansion in the absence of density regulation mechanisms. This process induced parallel, genome-wide shifts in allele frequencies across four, independent replicate populations. The strong selection coefficients quantified throughout this period (~10% per generation) are comparable to those quantified during sampling between summer and fall in wild populations and outdoor mesocosms (13–16). Thus, while it has been previously speculated that such patterns of evolution of *D. melanogaster* across seasons may be dominated by adaptation to shifting temperatures (e.g., Boulétreau-Merle, Fouillet, and Terrier 1987), our lab-based manipulation here provides direct evidence that the sole impact of shifting population densities can drive rapid adaptation over the course of several generations. In effect, this finding bolsters the notion that the interaction of ecological and evolutionary forces may be a fundamental force acting within natural populations (33,34).

By imposing a bout of truncation selection midway through population expansion, during which time individual survival and associated patterns of allele frequency change revealed differences in relative stress tolerance (e.g., desiccation resistance) among genotypes, we aimed to explore evidence of fitness-tradeoffs in the system. We quantified a pervasive, genome-wide signal of fluctuating selection, whereby patterns of allele frequency movement reversed direction across selection regimes. These patterns are indicative of antagonistic pleiotropy, whereby selected alleles conveying advantageous trait values for the phenotypes favored during reproduction selection convey disadvantageous trait values for the suite of phenotypes favored during truncation (7,35). While this dynamic was *a priori* expected based on the negative genetic correlations between traits likely under selection in our experiment (10,26–28), few data exist directly observing the manifestation of such trade-offs, particularly over the ecologically-relevant timescales assayed in our experiment. The key implication of this antagonistically pleiotropic behavior is that this process may, in effect, maintain variation in the population (7,8). Specifically, mutations with unconditionally advantageous effects will ultimately be driven to fixation, while those with context-specific behavior can in principle be maintained for longer periods of time in the presence of fluctuating selective pressures (36,37). The key challenge for future work will be to discern whether these dynamics indeed balance alleles over long time-periods, which will ultimately be aided by refining the causal loci driving the patterns observed here and observation across multiple bouts of fluctuating selection.

The genomic architecture of trade-offs was not dominated by a single, large-effect locus, nor were antagonistically pleiotropic alleles distributed evenly across chromosomal arms. Rather, we quantified a dominant signal of fluctuating selection across reproduction and stress-tolerance selection on chromosomal arms 3L, 3R, and X, and extensively showed that this signal was not spurious due to variation in the number of SNPs available for analysis across arms. We further showed that the trade-offs quantified on 3R were likely underpinned by inversions segregating in our experimental population. Inversions are a common form of structural variation both across species and within populations and have long been recognized as important for adaptive evolution. For example, suppressed recombination within the breakpoints of relatively young inversions can give rise to the accumulation of putatively adaptive alleles (38–41). Indeed, in natural populations of *D. melanogaster*, inversions are implicated in adaptation to spatially and temporally varying selection pressures, which may, in part, be underpinned by the selective pressures manipulated in our lab-based study (15,42–44). We add to this growing body of research, here reporting evidence for a role of each of five major chromosomal inversions during reproduction selection, with two of these inversions, *In(3R)P*, and *In(3R)K*, showing a behavior consistent with trade-offs between reproduction and stress-tolerance selection. In congruence with this result, previous experimental work has

465 demonstrated that inverted karyotypes of *In(3R)P* are shorter-lived and less stress resistant than
466 non-inverted karyotypes, as would be predicted based on the frequency trajectories of this
467 inversion in our experiment (45). Still, resolving the functional relevance of each of these major
468 inversions is nascent and provides a fruitful avenue of future research.
469 It is noteworthy that the trade-offs quantified here manifested during a single bout of
470 truncation selection, which proceeded over the course of just 24 hours. This illustrates how
471 finely adaptation via standing variation can track shifting environmental conditions, and that
472 selection can be detected on temporal scales shorter than a single generation. Indeed, mounting
473 research has shown how both anomalous (e.g., heat wave, hurricane, or poaching) and non-
474 anomalous (e.g., weekly to seasonal shifts in abiotic and biotic conditions) environmental
475 perturbations can elicit patterns of adaptation on similar timescales in natural populations (13–
476 16,44,46–49). This growing body of research increasingly suggests that the selection coefficients
477 underpinning the standing, functional variation in natural populations may be substantially
478 larger than previously recognized, and ultimately augment adaptation to the accelerated shifts
479 in mean environments associated with global climate change.
480 Our paired indoor cage and outdoor mesocosm experiment demonstrated how selection
481 induced by population expansion under controlled, lab-based conditions can identify alleles
482 that are relevant to adaptation and trade-offs in response to natural environmental fluctuations
483 (Fig. 5). Such repeatability in the targets of selection suggests a set of shared selective pressures
484 acted across environments, which in this case are most likely those associated with changes in
485 population density (as abiotic conditions in the indoor environment were controlled). Thus, the
486 repeated signals of fluctuating selection quantified in *D. melanogaster*, both across populations
487 and through time (13–16), may be underpinned by fundamental life-history trade-offs that
488 emerge as a result of shifts in the selective environment induced by population boom-bust
489 demographic dynamics. These boom-bust demographic dynamics are a generic feature of
490 populations across taxa, and this phenomenon may thus be more widespread than previously
491 recognized and provide a key instance of the interplay of ecological and evolutionary forces in
492 natural populations (33,34,50,51). Furthermore, given the dramatic differences in indoor and
493 outdoor environments, the shared allele frequency patterns observed here suggest that core
494 alleles may underpin adaptation during generic boom and bust cycles, regardless of specific
495 abiotic condition. Such a 'coarse-graining' of the architecture of the adaptive response would
496 position fluctuating selection as a key force in maintaining variation in natural populations.
497 While we note the possibility that a unique set of loci become fitness-relevant under specific
498 abiotic contexts, our goal here was to simply query whether a common set of loci underpinned
499 trade-offs and responded generically to fluctuating selection across abiotic contexts, for which
500 we found substantial evidence.

Finally, we observed differences in the architecture of trade-offs between our indoor population expansion/truncation selection experiment and the paired indoor-outdoor mesocosm study. For example, chromosomal arm 2L displayed evidence of sustained selection across reproduction and stress tolerance selection in the first experiment, but exhibited strong evidence of antagonistic pleiotropy between population expansion and collapse during our outdoor mesocosm experiment. Furthermore, we did not detect any shared enrichment of SNPs with evidence of linked selection across these experiments. While there were distinct methodological and environmental differences that may have played a role in this discordance, a more salient possibility is the use of a different set of inbred lines to generate the outbred mapping population for each experiment. As our analyses ultimately identify sets of SNPs in tight linkage to an underlying causal locus, differences in patterns of linkage across mapping populations could, in effect, lead to dramatic differences in the relative effect size of marker alleles across studies. This process is analogous to the oftentimes poor portability of association studies across human populations (52,53), and is an important consideration for future research using experimental evolution approaches to characterize the architecture of complex trait adaptation.

In conclusion, the data presented here demonstrate how well-resolved time-series genomic data can reveal the presence of, and genomic architecture underlying, fitness trade-offs. Such trade-offs can emerge in response to generic shifts in intraspecific ecological conditions and may in turn serve as a key force maintaining variation in natural populations.

Methods

Population construction and replicate cage seeding for expansion/truncation selection experiment

We constructed a genetically diverse founder population via outbreeding 145 lines of the Drosophila melanogaster Genetic Reference Panel (DGRP; Mackay et al. 2012). Ten mated females per DGRP inbred line, each of the same age cohort and from density-controlled line cultures, were pooled into a single 0.3m x 0.6m x 0.3m cage (P0 generation). Over the course of 4d, eggs were collected in 64 culture bottles (P1 generation) using cornmeal molasses medium (16 bottles added per day, then capped, removed, and 16 new bottles added); the 64 bottles were then randomly assigned to 1 of the 4 replicate cages; from this point forward, all replicates were cultured independently. Once flies eclosed, they were again released into 4 replicate 0.3m x 0.7m x 0.3m cages, and eggs were collected over 24h on two culture trays (0.5m x 0.3m x 0.1m; 1L of Drosophila cornmeal molasses medium per tray) per replicate. Once these flies eclosed (the first true F1 generation), the flies were released into 4 replicate, medium sized cages (0.6m x 0.6m x 1.2m) and allowed to oviposit on 4 trays of culture medium over 24h (embryos were the F2 generation). The F2 embryos from these trays were sealed and collected, and the adults

discarded. Once the F2 flies eclosed, they were then released into large, experimental cages (3 m^3) and given 8 trays of media for oviposition over 24h.

Population expansion/truncation selection in large, indoor cages

Culturing within the large, experimental cages proceeded via discrete generations, throughout which the amount of food was doubled every generation (starting at 8 trays for the F2). By doubling food every generation we generated sustained population expansion, thereby continually selecting for the fecundity-associated traits that are advantageous in wild populations when resources are abundant, most likely reproductive output and increased developmental rate. Hereafter, we refer to collection timepoints in accordance with the number of generations since release into large cages (the first 8 tray stage), which was two generations from the actual founding of the experimental populations (i.e., the samples labeled "generation 0" are F2's, "generation 1" are F3's, and so forth). Once flies had laid eggs for a period of 24h, the adults were removed from each cage and their volume measured to estimate census size. The remaining embryos (i.e., subsequent generation of flies) were left to develop and eclose within the same replicate cage. Across replicates, the population sizes rapidly expanded across generations, a dynamic selecting for individuals exhibiting increased reproduction (e.g., fecundity and developmental rate) (Fig. 1).

We maintained food doubling until generation 4 at which point, due to logistical constraints, it was not possible to continue the doubling of population size per cage beyond the 64 tray per cage generation. In order to continue the selection regime of population expansion (and selection for early fecundity, fast developmental rate), we instituted a random dilution followed by re-expansion. Specifically, the fifth generation was founded with embryos collected over 24h on two food trays (containing eggs of approximately 20,000 flies), from the from the fourth generation. These generation 5 flies were allowed to develop and eclose in the large indoor cages, at which point 8 trays of medium were added to each replicate cage, and oviposition was carried out for 24h and after which adults were discarded. Once the sixth generation of flies eclosed, they were allowed to oviposit for 24h on 2 food trays. These generation 7 embryos were then allowed to develop, eclose, and expanded out to 8 trays for the eighth generation. The generation 8 flies were allowed to oviposit over 24h on 8 trays, consistent with the previous generation (thus each generation of flies resulted from egg laying over a period of 24h from the preceding generation). We collected a random sample of 100 male and 100 female flies in generations 0, 1, 2, 3, 4, 6, 8, and 9, and pooled sexes separately for later whole genome shotgun sequencing (see below). We took additional samples of each replicate at generations 6 and 8 to use as biological replicates (i.e., same replicate/generation, different set of 100 flies) to quantify noise in our allele frequency estimates (see below). Census estimates (based on the total volume of dead flies) of adult flies within each cage were conducted on

generations 1-4. Calibration of census estimates was conducted by counting the number of flies (desiccated and dried) in $1cm^3$ and then measuring the volume in a given sample.

Concurrent to the generation 4 dilution and re-expansion described above, we induced a bout of truncation selection that segregated flies based on stress-tolerance, most likely desiccation resistance as water availability is expected to drive mortality at a much faster rate than starvation (30). Specifically, after collecting generation 5 embryos for continued expansion, we retained the adult flies within their respective cages, removed all food and water, and sampled the surviving adults at various timepoints by direct aspiration using vacuums until all flies had died. The live samples were immediately sorted by sex and timepoint and preserved in ethanol (as with expansion samples, 100 male and 100 female flies were isolated at each collection time point for whole genome shotgun sequencing). We continued this process until there were an insufficient number of flies (i.e., < 100 individuals) for sampling, resulting in samples for DNA analysis collected at hours 0, 6, 9, 12, 15, 19, and 21. Samples collected throughout this process were expected to reveal changes in the relative frequency of genotypes of differing degrees of stress-tolerance, traits expected to trade-off with the reproductive traits favored during the nine generations of expansion selection (10,26).

Pooled genomic sequencing and allele frequency estimation of expansion/truncation selection experiment

Genomic DNA from pools of 100 male and 100 female flies were extracted and sequenced in two rounds. First, multiplexed libraries were created for 56 expansion samples using Illumina Nextera DNA Prep with e-gel size selection and i7 indexing. Barcoded fragments were mixed and loaded evenly into 4 lanes, then sequenced with 100 bp paired-end reads on an Illumina HiSeq2000 sequencer, with target coverage of 10x per sample. After QC, resulting per/sample coverage was quite variable, and 26 samples from this round with coverage <2x were later re-sequenced with 150-bp dual-indexed reads on an Illumina HiSeq4000 sequencer. Reads from the same sample were merged after all QC and mapping steps. A separate round of sequencing was conducted for the remaining 96 samples, including remaining expansion and truncation samples. These samples were sequenced with 150-bp paired-end dual-index reads on a NextSeq550 high output machine, with a target coverage 7x per sample. All samples were demultiplexed, adapter sequences were trimmed, and reads with any 3' bases with quality score < 20 or trim length <18 were discarded. Overlapping forward and reverse reads were subsequently assembled and reads were mapped separately to the *D.mel* v5.39 reference genome using bwa and default parameters (54). Aligned reads were deduplicated using Picard tools (http://broadinstitute.github.io/picard/), and all reads were re-aligned around indels using GATK v4 IndelRealigner (https://gatk.broadinstitute.org/).

We used the founder line genome sequencing data to compute haplotype-informed allele frequency estimates at 2.7 M, previously identified, segregating sites using a local

609 inference method and pipeline developed and described by (55,56). We conducted haplotype
610 inference in window sizes that varied proportionally to the length of un-recombined haplotype
611 blocks expected as a function of the estimated number of generations since the construction of
612 the outbred population. We have previously provided extensive validation that our haplotype-
613 informed allele frequencies are replicable across different sets of subsampled reads from the
614 same sample, and produce an accuracy of allele frequency estimates that is comparable to deep
615 sequencing and standard methods for pooled, allele frequency estimation (14,16,56). We filtered
616 our allele frequencies to only include those sites with an average minor allele > 0.02 across all
617 samples, resulting in 1.7M SNPs for analysis. Finally, we averaged technical replicate allele
618 frequencies (i.e. the male and female pool) for each replicate and collection time point, resulting
619 in 4 cages per time point, across 8 expansion generations and 7 truncation time points (60 total
620 samples).
621 *Identifying parallel patterns of allele frequency change and genome-wide evidence of trade-offs*
622 Statistical analysis of allele frequency data was conducted using R v. 3.5.6. We first
623 explored patterns of genomic divergence between samples as average F_{ST} across all segregating
624 sites. We compared F_{ST} between biological replicates (same replicate cage/collection time point,
625 different pool/extraction) to those values obtained for each replicate cage between its
626 generation 0 sample and all subsequent generation expansion samples. We quantified whether
627 evolutionary divergence increased throughout the course of expansion, and at what point it
628 exceeded the sources of biological noise impacting our allele frequency estimates, using a linear
629 regression.
630 We next explored whether increasing differentiation through time was underpinned by
631 parallel allele frequency shifts across replicates. Specifically, we used our pairwise divergence
632 values (i.e., F_{ST} between samples) to create multi-dimensional scaling (MDS) plots, in which
633 divergence between samples is represented as distance between points in a 2-D plane (*cmdscale*
634 function in *stats* package R). If the observed evolutionary change across replicates was
635 dominated by parallel responses to shifting population densities, then the segregation of
636 samples across the 2-D plane should, in part, correspond to the generation from which samples
637 were derived. This analysis was conducted on mean, genome-wide F_{st} values, as well as F_{st}
638 values computed separately for SNPs on each chromosomal arm. To quantify trends visualized
639 using MDS, we translated the coordinates for each cage such that the centroid of all generation 0
640 samples was centered at the origin and then projected all points onto a single axis which was
641 constructed by fitting a simple linear regression model to samples from generations 0-4
642 (regressions were constructed separately for each replicate cage). The resulting axis represents
643 the primary axis of variation in the 2D plane during early expansion for each replicate (Fig. S5).
644 We hypothesized that if sustained, directional selection imposed by population expansion was a

primary driver of patterns in genomic variation, the samples from the remaining expansion generations would continue to proceed along the established axis of variation in the same direction. To evaluate the degree of concordance of early and late expansion generation samples we computed Pearson correlations between collection time and position along this axis. The significance of these correlations was determined via permutations of sample collection time point ($N = 100$ permutations). We conducted this analysis across all genome-wide SNPs, as well as separately for SNPs on each chromosomal arm. Finally, to disentangle whether differences in dynamics observed among chromosomal arms reflected true systematic differences in allele frequency behavior, as opposed technical artifacts driven by differences in the number of SNPs available for analysis across arms, we re-conducted this analysis using randomly sampled subsets of SNPs matching that present on each arm.

We quantified the extent to which individual SNPs exhibited parallel movement across cages and throughout expansion by fitting a generalized linear model to allele frequencies (formula: allele frequency ~ expansion generation; *glm* function in base R) (31). Allele frequencies were weighted by the total number of chromosomes sequenced (N = 200) and depth per sample (56). This model assessed the significance of the linear relationship between allele frequency and generation of sampling across all cages, using a quasibinomial error model to reduce false positive associations (15,16,31). While an association between allele frequency and sampling timepoint within a single cage may represent either drift or selection, a significant association across all four replicate cages indicates allele frequency trajectories that are both predictable over time and parallel across populations. In this case, selection (or linked selection) is the more parsimonious explanation. P-values were adjusted using the Benjamini-Hochberg false discovery rate (FDR) correction (*p.adjust* package in base R). We considered a SNP significant if it exhibited an FDR < 0.01 and effect size > 2%. We further validated the parallelism inferred by the GLM by implementing a leave-one-out cross validation. Specifically, we iteratively identified sets of parallel SNPs (FDR < 0.01 and effect size > 2%) across expansion samples using a GLM and allele frequency data from three of the four replicate cages. We then quantified the frequency shifts of the favored allele at significant sites in the left-out cage between generation 0 and 9 of expansion. For sets of parallel SNPs identified via each iteration of the leave-one-out validation, we matched each parallel SNP to a control SNP based on chromosomal arm, starting frequency (within 5% of generation 0 frequency), inversion status, and recombination rate. We compared the distribution of allele frequency shifts at parallel vs. matched control sites (using a two-tailed, paired t-test) for each left out cage to infer if the magnitude of allele frequency change exceeded background allele frequency movement, and whether the dominant direction of allele frequency change was in a direction concordant with

the three training cages. We conducted this analysis for SNPs identified genome-wide, as well as separately for each chromosomal arm.

Quantifying evidence of fecundity and desiccation tolerance trade-offs

We next used samples collected throughout truncation, during which patterns of survival and selection were most likely underpinned by differences in desiccation resistance among genotypes, to quantify the presence of fitness-tradeoffs at those alleles systematically favored during expansion. First, we re-conducted our MDS analysis of pairwise divergence values, this time including all samples collected across both phases of the experiment. While general qualitative inference may be obtained from the visualization of this analysis, we explicitly quantified the relative movement of expansion and truncation samples across this 2D plane to provide a more rigorous investigation into the relative direction of allele frequency movement across phases. Specifically, as above, we translated MDS coordinates for each cage such that the centroid of all hour 0 truncation samples was centered at the origin, and then projected onto a single axis which, as above, was constructed by fitting a simple linear regression model to samples from expansion generations 0-4. We hypothesized that if our contrasting reproduction/stress-tolerance selection environments were inducing genome-wide reversions in allele frequency movement, then the projection of samples collected throughout truncation would regress in the opposite direction as that observed for the expansion samples. The degree of concordance/discordance for truncation samples were quantified as the Pearson Correlation between collection time and position along this axis, and the significance of these correlations was determined via permutations of sample collection time point ($N = 100$ permutations). This analysis was also conducted both genome-wide and separately for each chromosomal arm and, as above, we used genome-wide subsampling to infer whether variation in dynamics among chromosomal arms was a technical artifact.

Finally, we aimed to corroborate patterns observed in the dimensionality reduction of genome and chromosome-wide F_{ST} values described above by accruing additional evidence of genome-wide trade-offs using allele frequency trajectories. We first isolated those SNPs with strong evidence of linkage to a selected locus throughout expansion (GLM FDR < 0.01 and allele frequency shift > 2%). We then quantified the behavior of the rising allele at each SNP throughout the progression of truncation, and evaluated whether these alleles were more likely to exhibit a reversion in trajectory direction, or continue in the same direction, using a Chi-squared test and set of control SNPs (matched on chromosome, starting frequency, recombination rate, and inversion status). We further assessed if and how the results of this analysis varied when only evaluating trajectories of expansion-identified SNPs that also exhibited independent evidence of linkage to a selected locus (GLM FDR < 0.05 and allele frequency shift > 1%) during truncation. Finally, we further quantified the distribution of

716 expansion-identified SNPs, genome-wide and per chromosomal arm, using empirical
717 cumulative distribution functions (eCDF). The eCDFs for SNPs with sustained directional
718 movement across treatments and evidence of fluctuating selection were evaluated
719 independently via comparison to eCDFs generated via a set of matched control SNPs, and using
720 a Kolmorgorov-Smirnov test (*stats* package in R).

721 *Assessing the role of inversions in fitness trade-offs*

722 From the known cosmopolitan inversions found in the DGRP lines (57), we analyzed
723 those that occurred in at least 10 of our founding strains (i.e., $> 4\%$). We used chromosome-level
724 haplotype frequencies to compute the frequencies of each of these inversions independently
725 during expansion and truncation. Next, we regressed inversion frequencies through time using
726 a generalized linear model (logistic link function, and quasibinomial error variance) to classify
727 inversions moving systematically across replicates throughout each phase (Benjamini-Hochberg
728 corrected P-value < 0.1). For those inversions displaying significant frequency change across
729 both expansion and truncation, we characterized those that exhibited thus fluctuating selection
730 and trade-offs (i.e. switched directions) across phases.

731 We next aimed to quantify the extent to which SNPs within, or in tight linkage of, of the
732 assayed inversions underpinned genome-wide signals of trade-offs quantified in our
733 experiment. We thus first generated a winnowed set of genome-wide SNPs, eliminating those
734 either within, or up to 100 Kb away from, inversion breakpoints. We then explored whether we
735 retained signals of parallel adaptation throughout expansion and trade-offs during truncation
736 using this inversion-free SNP panel and re-conducting our F_{ST}-based MDS analysis of
737 expansion and truncation samples (described in *Quantifying evidence of fecundity and desiccation*
738 *tolerance trade-offs*, above).

739 *Quantifying the emergence of trade-offs during adaptation to natural environmental fluctuations*

740 We next quantified whether alleles identified via selection under sustained population
741 expansion in a laboratory setting can predict patterns of adaption and trade-offs in an outdoor
742 environment. Specifically, we leveraged data from an independent study year when we
743 monitored patterns of genomic variation in a genetically diverse, outbred population that was
744 split into a series of large, replicate cages maintained in a controlled, indoor laboratory, as well
745 as outdoor mesocosms exposed to natural environmental fluctuations. Allele frequency data for
746 the outdoor mesocosms were previously described, analyzed, and reported in Bitter *et al.* (2024).
747 Briefly, the outdoor mesocosms were located in Philadelphia, Pennsylvania and consisted of
748 twelve, replicate 2 m^3 cages, each consisting of a single dwarf peach tree and exposed to natural
749 environmental fluctuations. The paired, indoor cage study (not reported or analyzed by Bitter *et*
750 *al.* (2024)) consisted of ten, replicate ~0.5 m^3 cages housed in a temperature-controlled laboratory
751 at the University of Pennsylvania.

The replicate cages for both the outdoor mesocosms and indoor cage study were seeded with a genetically diverse, outbred population, derived from a panel of 76 inbred strains originally collected wild from Linvilla Orchards, Media, PA (Supplementary Data File 2). Constat food was supplied to the replicates within each environment, whereby four hundred ml of Drosophila media ('Spradling cornmeal recipe') was provided in 900 cm^3 aluminum loaf pans within each cage three times per week. These pans provided the only source of food and egg laying substrate, and egg laying upon each loaf pan was carried out for two days, after which a new pan was added, and the original pan was covered with a mesh lid to prevent any further laying. The lids were removed after eclosure was first observed, causing each replicate to experience a near continual input of new flies that evolved with overlapping generations. The census size of the adult population in the outdoor mesocosms were estimated five times throughout the progression of the experiment (19 July, 5 August, 20 August, 17 September, and 20 October) (14), while the final census in the indoor cages were estimated at the end of the monitoring period via collection and volumetric quantification of all flies remaining in each replicate. The outdoor mesocosms were sampled to quantify patterns of genomic variation weekly for 9 weeks (13 July – 7 September), after which samples were collected on 21 September, 20 October, and following the first freeze and population crash on 20 December (12 total time points). The indoor cages were sampled at the first, second, eighth, and eleventh time points. The overlapping monitoring period of the outdoor and indoor cages encompassed the rapid population expansion and ultimate stabilization of cage densities as the replicates in each environment presumably reached carrying capacity (Fig S11) (14). The selective landscape experienced during this shared monitoring period probably mirrors that experienced under the sustained population expansion imposed by population expansion portion of the first experiment of this study, described above. The notable differences between these experiments, however, is that the first experiment increased food availability as a function of population density throughout expansion, while our latter experiment maintained a constant food substrate throughout the study period. Thus, this paired, indoor-outdoor study likely also imposed selection on traits impacted by density regulation (e.g., larval competition). Analogous to the truncation portion of the first experiment, the final collection interval of the outdoor mesocosm study (time point 11->12) encompassed the onset of winter, during which time a total crash of the adult populations in the cages was observed. Selection during this phase of the experiment was thus hypothesized to favor increased in stress tolerance traits associated with the deteriorating abiotic conditions.

Methods for sampling flies for sequencing from the outdoor mesocosms and indoor cages were similar to those described above: eggs were collected overnight directly from each replicate cage, and larval development was carried out in the indoor environment where eggs

788 developed and eclosed to F1 adults in 30 cm^3 cages. A random set of 100 females were sampled
789 from each cage 3-5 days post-eclosure and preserved in 99% ethanol at -20° C (14). Genomic
790 DNA was extracted from pools of flies for each replicate and time point using the Monarch
791 Genomic DNA Purification Kit (New England Biolabs). Libraries were nonstructured using the
792 Illumina DNA Prep Tagmentation Kit and all samples were sequenced on Illumina Novaseq
793 6000 flow cells using 150 bp, paired end reads. Raw sequencing reads from each sample were
794 trimmed of adapter sequences and bases with quality score < 20, and aligned to the *Drosophila*
795 *melanogaster* v5.39 reference genome using bwa and default parameters(54). Aligned reads were
796 deduplicated using Picard tools (http://broadinstitute.github.io/picard/) and the final set of
797 reads for each sample was down-sampled to obtain an equivalent, genome-wide coverage of 8x
798 across all samples (>100x effective coverage) (14,56). Haplotype informed allele frequencies
799 from reads from each replicate and time point were then generated using the local inference
800 pipeline described above (see *Allele Frequency Calculation*) (55,56). The final set of allele
801 frequencies only included those sites with an average minor allele frequency > 0.02 in the
802 baseline population, and present in at least one evolved sample at a MAF > 0.01, ultimately
803 yielding 1.9 M SNPs (14).
804 We quantified the extent to which evolution in the outdoor environment is driven by the
805 selective pressures solely associated with increasing population density (e.g., selection for
806 increased fecundity, faster developmental rate, increased competitive ability). Specifically, we
807 identified alleles with systematic behavior independently in the indoor and outdoor cages using
808 GLM (FDR < 0.01; allele frequency change > 2%), and compared the observed overlap between
809 these lists using a Chi squared test and with null expectations generated via a set of matched
810 control SNP set. Next, we queried whether the dominant direction of selection throughout the
811 shared monitoring period of the indoor and outdoor environment was consistent. Specifically,
812 we quantified the mean frequency shift of the rising allele at SNPs identified in the indoor
813 environment via GLM, in the outdoor cages during the same time interval. We determined
814 whether the magnitude of frequency shifts exceeded background allele frequency movement
815 expected via matched control SNPs, and whether the dominant direction was parallel or anti-
816 parallel given the sign of observed median shift (analysis conducted both genome-wide and
817 separately for each chromosomal arm using a paired t-test and sign test). Finally, to test
818 whether the total population crash in the outdoor mesocosms elicited evidence of the trade-offs,
819 we quantified the behavior of the indoor identified alleles throughout this period. Through this,
820 we used a common set of alleles, determined as linked to an advantageous locus during
821 increasing population density in a controlled, lab-based setting, to infer the presence of trade-
822 offs during a bout of population expansion and collapse in a semi-natural setting.
823

Acknowledgements: We are grateful to members of the Petrov and Schmidt labs for discussion during experimental design, lab work, and data analysis. We thank our funding organizations: the National Science Foundation (NSF PRFB 2109407 to M.C.B.) and the National Institutes of Health (NIH 5R35GM118165-07 to D.A.P and NIH R01GM100366 and R01GM137430 to P.S.), as well as funding support from the Chan Zukerberg Biohub.

Author Contributions: The experiment was conceived by D.A.P, P.S, A.O.B., S.R., and M.C.B. Data curation for the population expansion/truncation experimental evolution study was conducted by S.R., N.B., S.T., A.O.B., S.G., and P.S. Data curation paired indoor and outdoor mesocosm study was conducted by M.C.B., S.B., H.O., and P.S. Formal analysis was carried out by S.G., J.H., and M.C.B. The original manuscript was prepared by M.C.B., S.G., P.S., and D.A.P. All authors reviewed and edited the final version of the manuscript.

Data availability: Sequencing data from the DGRP lines used in the population expansion/truncation experiment are publicly available at http://dgrp2.gnets.ncsu.edu/data.html. Haplotype-informed allele frequency estimates from evolved, outbred samples and used in statistical analysis will be available at upon article publication. Founder line sequences for the paired indoor/outdoor mesocosm study are available at NCBI Accession PRJNA722305. Outdoor mesocosm sequences generated via pooled sequencing for this experiment are available at NCBI accession PRJNA1031645 and alle frequency data is available at the following Dryad repository: https://doi.org/10.5061/dryad.xd2547dpv. Indoor sequences and allele frequency data will be available upon article publication.

Code availability: Code associated with all analyses conducted in this manuscript are publicly available at the following repository: https://github.com/MarkCBitter/Drosophila-fitness-trade-offs .

References

1. Roff DA, Fairbairn DJ. The evolution of trade-offs: where are we? Journal of Evolutionary Biology. 2007 Mar 1;20(2):433–47.

2. Stearns SC. Trade-Offs in Life-History Evolution. Functional Ecology. 1989;3(3):259–68.

3. Grant PR, Abbott I, Schluter D, Curry RL, Abbott LK. Variation in the size and shape of Darwin's finches. Biological Journal of the Linnean Society. 1985 May 1;25(1):1–39.

4. Shoval O, Sheftel H, Shinar G, Hart Y, Ramote O, Mayo A, et al. Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space. Science. 2012 Jun;336(6085):1157–60.

5. Christie MR, McNickle GG, French RA, Blouin MS. Life history variation is maintained by fitness trade-offs and negative frequency-dependent selection. Proceedings of the National Academy of Sciences. 2018 Apr 24;115(17):4441–6.

6. Charlesworth B. Evolution in Age-Structured Populations. 2nd ed. Cambridge: Cambridge University Press; 1994.

7. Curtsinger JW, Service PM, Prout T. Antagonistic Pleiotropy, Reversal of Dominance, and Genetic Polymorphism. The American Naturalist. 1994 Aug;144(2):210–28.

8. Rose MR. Antagonistic pleiotropy, dominance, and genetic variation. Heredity. 1982 Feb;48(1):63–78.

9. Mitchell-Olds T, Willis JH, Goldstein DB. Which evolutionary processes influence natural genetic variation for phenotypic traits? Nat Rev Genet. 2007 Nov;8(11):845–56.

10. Hoffmann AA, Parsons PA. Selection for increased desiccation resistance in Drosophila melanogaster: additive genetic control and correlated responses for other stresses. Genetics. 1989 Aug 1;122(4):837–45.

11. Schmidt PS, Paaby AB, Heschel MS. Genetic variance for diapause expression and associated life histories in Drosophila melanogaster. Evolution. 2005 Dec 1;59(12):2616–25.

12. Schou MF, Engelbrecht A, Brand Z, Svensson EI, Cloete S, Cornwallis CK. Evolutionary trade-offs between heat and cold tolerance limit responses to fluctuating climates. Science Advances. 2022 May 27;8(21):eabn9580.

13. Bergland AO, Behrman EL, O'Brien KR, Schmidt PS, Petrov DA. Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in Drosophila. PLOS Genetics. 2014 Nov 6;10(11):e1004775.

14. Bitter MC, Berardi S, Oken H, Huynh A, Lappo E, Schmidt P, et al. Continuously fluctuating selection reveals fine granularity of adaptation. Nature. 2024 Aug 14;1–8.

15. Machado HE, Bergland AO, Taylor R, Tilk S, Behrman E, Dyer K, et al. Broad geographic sampling reveals the shared basis and environmental correlates of seasonal adaptation in Drosophila. eLife. 2021 Jun 22;10:e67577.

16. Rudman SM, Greenblum SI, Rajpurohit S, Betancourt NJ, Hanna J, Tilk S, et al. Direct observation of adaptive tracking on ecological time scales in Drosophila. Science. 2022 Mar 18;375(6586):eabj7484.

17. Tataru D, Leon MD, Dutton S, Perez FM, Rendahl A, Ferris KG. Fluctuating selection in a Monkeyflower hybrid zone. bioRxiv; 2024. Available from: https://www.biorxiv.org/content/10.1101/2024.06.14.599085v1

18. Troth A, Puzey JR, Kim RS, Willis JH, Kelly JK. Selective trade-offs maintain alleles underpinning complex trait variation in plants. Science. 2018 Aug 3;361(6401):475–8.

19. Flatt T, Amdam GV, Kirkwood TBL, Omholt SW. Life-History Evolution and the Polyphenic Regulation of Somatic Maintenance and Survival. The Quarterly Review of Biology. 2013 Sep;88(3):185–218.

20. Ives PT. The genetic structure of American populations of Drosophila melanogaster. Genetics. 1945 Mar 15;30(2):167–96.

21. Band HT, Ives PT. Correlated changes in environment and lethal frequency in a natural population of drosophila melanogaster. Proceedings of the National Academy of Sciences. 1961 Feb;47(2):180–5.

22. Gleason JM, Roy PR, Everman ER, Gleason TC, Morgan TJ. Phenology of Drosophila species across a temperate growing season and implications for behavior. PLOS ONE. 2019 May 16;14(5):e0216601.

23. Boulétreau-Merle J, Fouillet P, Terrier O. Seasonal variations and balanced polymorphisms in the reproductive potential of temperate D. Melanogaster populations. Entomologia Experimentalis et Applicata. 1987;43(1):39–48.

24. Behrman EL, Watson SS, O'Brien KR, Heschel MS, Schmidt PS. Seasonal variation in life history traits in two Drosophila species. Journal of Evolutionary Biology. 2015;28(9):1691–704.

25. Behrman EL, Schmidt P. How predictable is rapid evolution? bioRxiv; 2022. Available from: https://www.biorxiv.org/content/10.1101/2022.10.27.514123v1

26. Hoffmann AA, Parsons PA. An integrated approach to environmental stress tolerance and life-history variation: desiccation tolerance in Drosophila. Biological Journal of the Linnean Society. 1989;37(1–2):117–36.

27. Service PM. Physiological Mechanisms of Increased Stress Resistance in Drosophila melanogaster Selected for Postponed Senescence. Physiological Zoology. 1987 May;60(3):321–6.

923 28. Zera AJ, Harshman LG. The Physiology of Life History Trade-Offs in Animals. Annual
924 Review of Ecology, Evolution, and Systematics. 2001 Nov 1;32(Volume 32, 2001):95–126.

925 29. Del Monte-Luna P, Brook BW, Zetina-Rejón MJ, Cruz-Escalona VH. The carrying capacity
926 of ecosystems. Global Ecology and Biogeography. 2004;13(6):485–95.

927 30. Da Lage JL, Capy P, David JR. Starvation and desiccation tolerance in *Drosophila*
928 *melanogaster* adults: Effects of environmental temperature. Journal of Insect Physiology.
929 1989 Jan 1;35(6):453–7.

930 31. Wiberg RAW, Gaggiotti OE, Morrissey MB, Ritchie MG. Identifying consistent allele
931 frequency differences in studies of stratified populations. Methods in Ecology and Evolution.
932 2017;8(12):1899–909.

933 32. Berdan EL, Barton NH, Butlin R, Charlesworth B, Faria R, Fragata I, et al. How
934 chromosomal inversions reorient the evolutionary process. J Evol Biol. 2023
935 Dec;36(12):1761–82.

936 33. Carroll SP, Hendry AP, Reznick DN, Fox CW. Evolution on ecological time-scales.
937 Functional Ecology. 2007;21(3):387–93.

938 34. Hendry AP. Eco-evolutionary Dynamics. Princeton University Press; 2016.

939 35. Connallon T, Chenoweth SF. Dominance reversals and the maintenance of genetic variation
940 for fitness. PLOS Biology. 2019 Jan 29;17(1):e3000118.

941 36. Turelli M, Barton NH. Polygenic Variation Maintained by Balancing Selection: Pleiotropy,
942 Sex-Dependent Allelic Effects and G × E Interactions. Genetics. 2004 Feb 1;166(2):1053–
943 79.

944 37. Wittmann MJ, Bergland AO, Feldman MW, Schmidt PS, Petrov DA. Seasonally fluctuating
945 selection can maintain polymorphism at many loci via segregation lift. Proceedings of the
946 National Academy of Sciences. 2017 Nov 14;114(46):E9932–41.

947 38. Sturtevant AH. A Case of Rearrangement of Genes in Drosophila1. Proceedings of the
948 National Academy of Sciences. 1921 Aug;7(8):235–7.

949 39. Sturtevant AH, Beadle GW. The relations of inversions in the X chromosome of Drosophila
950 melanogaster to crossing over and disjunction. Genetics. 1936 Sep 1;21(5):554–604.

951 40. Kirkpatrick M, Barton N. Chromosome Inversions, Local Adaptation and Speciation.
952 Genetics. 2006 May 1;173(1):419–34.

953 41. Kirkpatrick M. How and Why Chromosome Inversions Evolve. PLOS Biology. 2010 Sep
954 28;8(9):e1000501.

42. Kapun M, Mitchell ED, Kawecki TJ, Schmidt P, Flatt T. An Ancestral Balanced Inversion Polymorphism Confers Global Adaptation. Molecular Biology and Evolution. 2023 Jun 1;40(6):msad118.

43. Kapun M, Fabian DK, Goudet J, Flatt T. Genomic Evidence for Adaptive Inversion Clines in Drosophila melanogaster. Molecular Biology and Evolution. 2016 May 1;33(5):1317–36.

44. Nunez JCB, Lenhart BA, Bangerter A, Murray CS, Mazzeo GR, Yu Y, et al. A cosmopolitan inversion facilitates seasonal adaptation in overwintering Drosophila. Genetics. 2024 Feb 1;226(2):iyad207.

45. Durmaz E, Benson C, Kapun M, Schmidt P, Flatt T. An inversion supergene in Drosophila underpins latitudinal clines in survival traits. Journal of Evolutionary Biology. 2018;31(9):1354–64.

46. Campbell-Staton SC, Arnold BJ, Gonçalves D, Granli P, Poole J, Long RA, et al. Ivory poaching and the rapid evolution of tusklessness in African elephants. Science. 2021 Oct 22;374(6566):483–7.

47. Campbell-Staton SC, Cheviron ZA, Rochette N, Catchen J, Losos JB, Edwards SV. Winter storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard. Science. 2017 Aug 4;357(6350):495–8.

48. Donihue CM, Herrel A, Fabre AC, Kamath A, Geneva AJ, Schoener TW, et al. Hurricane-induced selection on the morphology of an island lizard. Nature. 2018 Aug;560(7716):88–91.

49. Rodríguez-Trelles F, Tarrío R, Santos M. Genome-wide evolutionary response to a heat wave in Drosophila. Biol Lett. 2013 Aug 23;9(4):20130228.

50. Boor GKH, Schultz CB, Crone EE, Morris WF. Mechanism matters: the cause of fluctuations in boom—bust populations governs optimal habitat restoration strategy. Ecological Applications. 2018;28(2):356–72.

51. Kendall BE, Briggs CJ, Murdoch WW, Turchin P, Ellner SP, McCauley E, et al. Why Do Populations Cycle? A Synthesis of Statistical and Mechanistic Modeling Approaches. Ecology. 1999;80(6):1789–805.

52. de Bakker PIW, Burtt NP, Graham RR, Guiducci C, Yelensky R, Drake JA, et al. Transferability of tag SNPs in genetic association studies in multiple populations. Nat Genet. 2006 Nov;38(11):1298–303.

53. Martinez KL, Klein A, Martin JR, Sampson CU, Giles JB, Beck ML, et al. Disparities in ABO blood type determination across diverse ancestries: a systematic review and validation in the All of Us Research Program. Journal of the American Medical Informatics Association. 2024 Jun 25;ocae161.

54. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009 Jul 15;25(14):1754–60.

55. Kessner D, Turner TL, Novembre J. Maximum Likelihood Estimation of Frequencies of Known Haplotypes from Pooled Sequence Data. Mol Biol Evol. 2013 May 1;30(5):1145–58.

56. Tilk S, Bergland A, Goodman A, Schmidt P, Petrov D, Greenblum S. Accurate Allele Frequencies from Ultra-low Coverage Pool-Seq Samples in Evolve-and-Resequence Experiments. G3 (Bethesda). 2019 Oct 21;9(12):4159–68.

57. Huang W, Massouras A, Inoue Y, Peiffer J, Ràmia M, Tarone AM, et al. Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. Genome Res. 2014 Jul 1;24(7):1193–208.

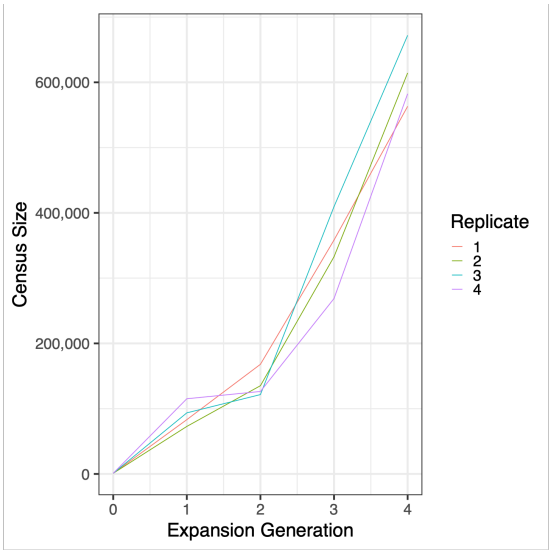

Figure S1. Census estimates of indoor replicate cages during the first four generations of population expansion.

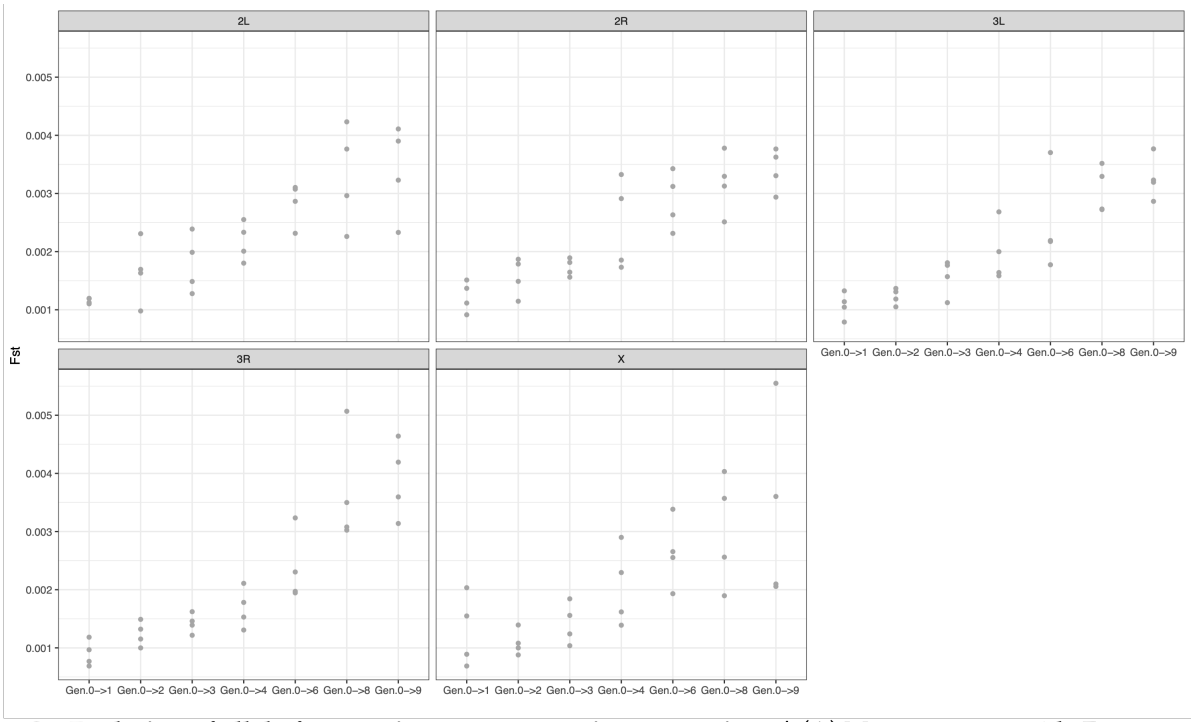

Figure S2. Evolution of allele frequencies across expansion generations | (A) Mean, genome-wide Fst between each evolved replicate and its generation 0 sample (Gen.0->n), segregated by chromosomal arm. Each grey dot corresponds to an individual replicate.

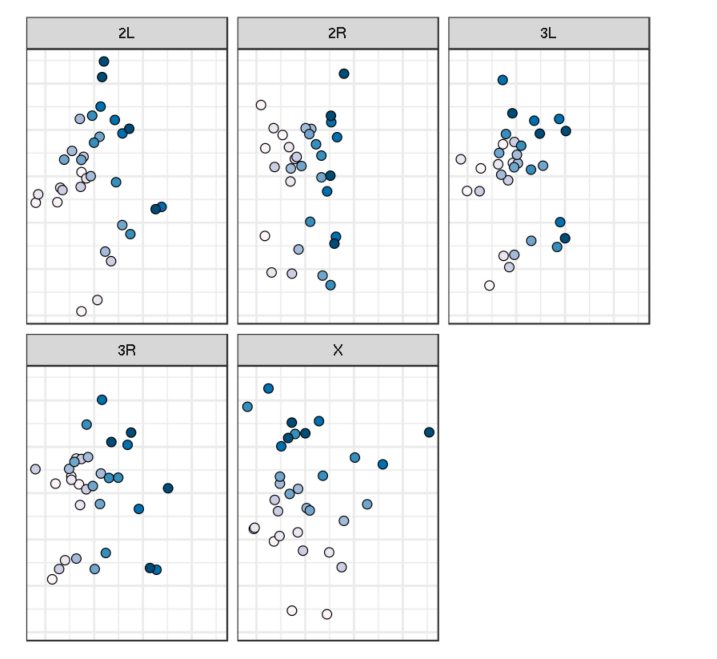

Figure S3. MDS of pairwise Fst values computed per chromosomal arm and and across all expansion samples. Point Color corresponds to sample expansion collection generation.

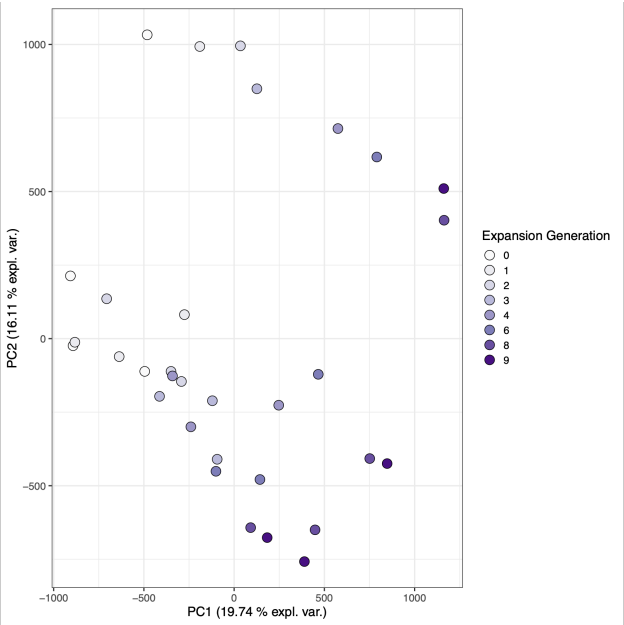

Figure S4. Principal component analysis of samples collected throughout expansion using allele frequency data across all 1.7 M SNPs. Samples are projected onto the first two principal components and colored in accordance with expansion generation collection time point.

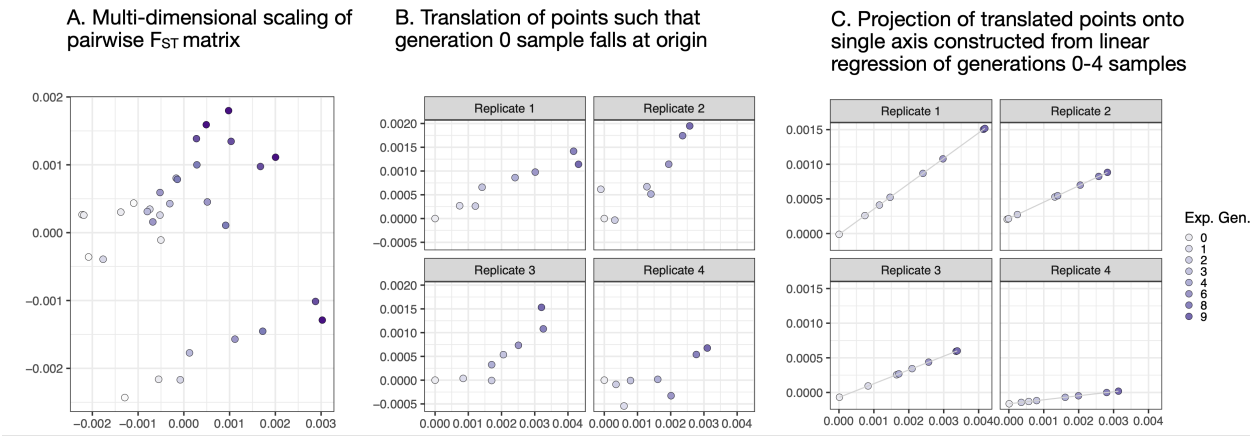

Figure S5. F_{ST}-based MDS analysis workflow. (A) Multi-dimensional scaling of expansion sample pairwise F_{ST} matrix as in main figure 2. (B) Translation of expansion points such that, for each replicate, the generation 0 sample falls at the plot origin. (C) Projection of translated points (B) onto a single axis constructed via a simple linear regression of generation 0-4 translated points.

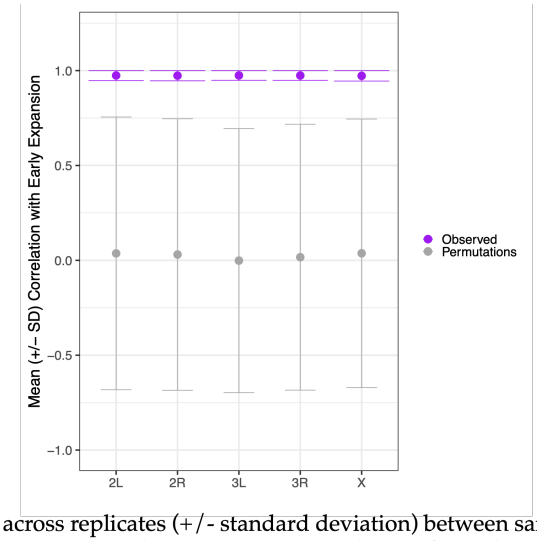

Figure S6. Average Pearson Correlation across replicates (+/- standard deviation) between sample expansion generation and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples. Purple points and error bars correspond to empirical values, while grey points and error bars correspond to values derived from empirical permutations (N = 100). Here, Fst values and MDS analysis was run on SNPs randomly sampled throughout the genome, matching the number of SNPs present on each chromosomal arm.

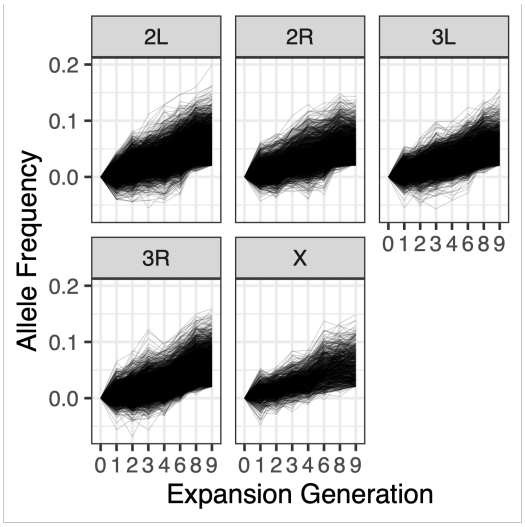

Figure S7. Trajectories of SNPs identified via GLM as exhibiting parallel frequency shifts across replicates throughout expansion (GLM FDR < 0.01 and effect size > 2%), segregated by chromosomal arm.

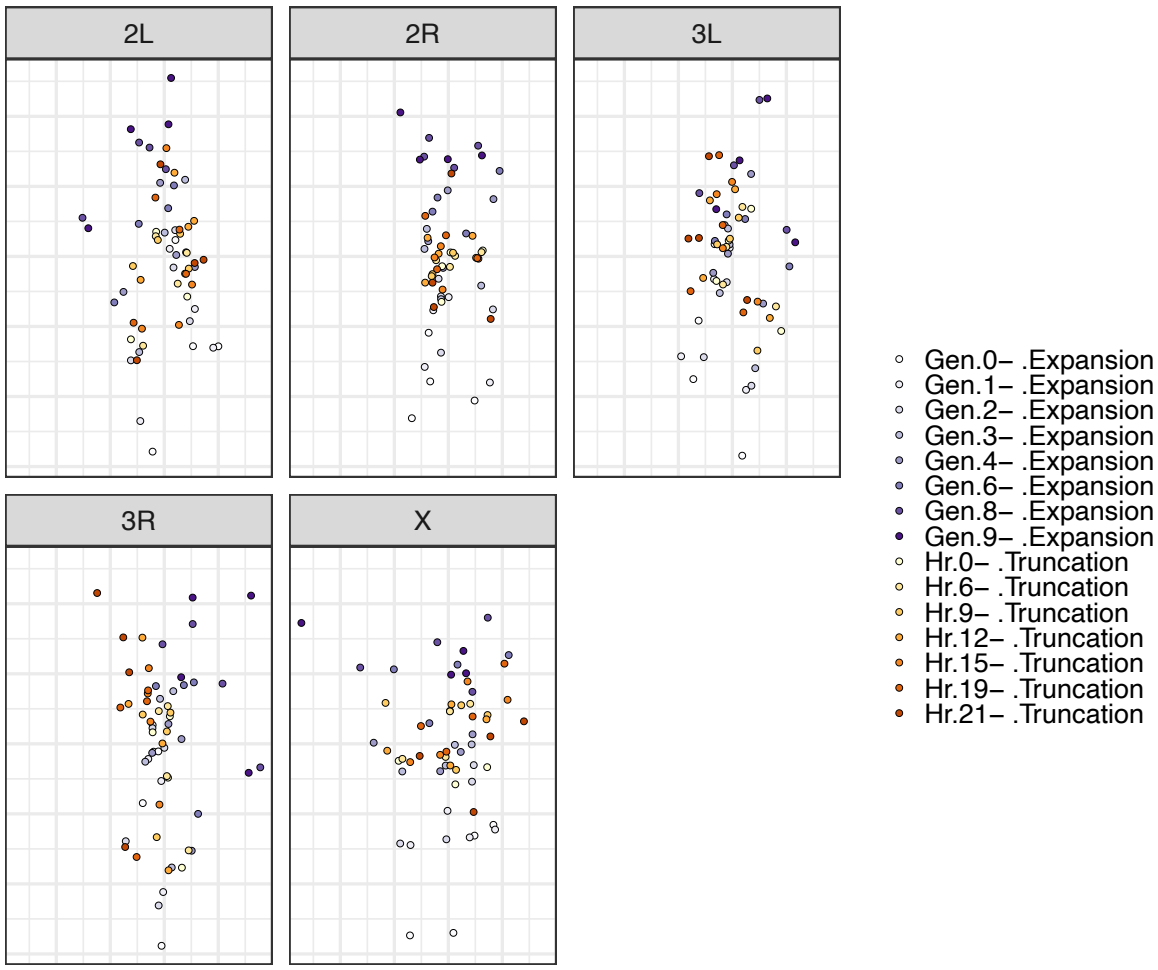

Figure S8. MDS of pairwise Fst values for each chromosomal arm and across all samples collected throughout expansion (purple-hue points) and truncation (orange-hue points). Points are shaded according to collection time point (darker hues indicate later expansion or truncation sampling generation/hour).

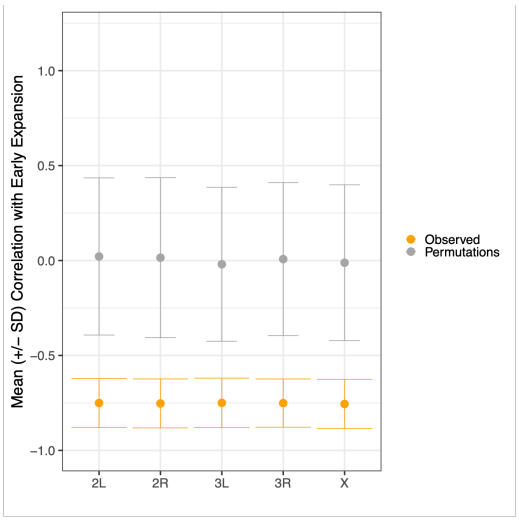

Figure S9. Average Pearson Correlation across replicates (+/- standard deviation) between sample truncation hour and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples. Orange points and error bars correspond to empirical values, while grey points and error bars correspond to values derived from empirical permutations (N = 100). Here, Fst values were derived from randomly sampled SNPs throughout the genome, matching the number of SNPs present on each chromosomal arm.

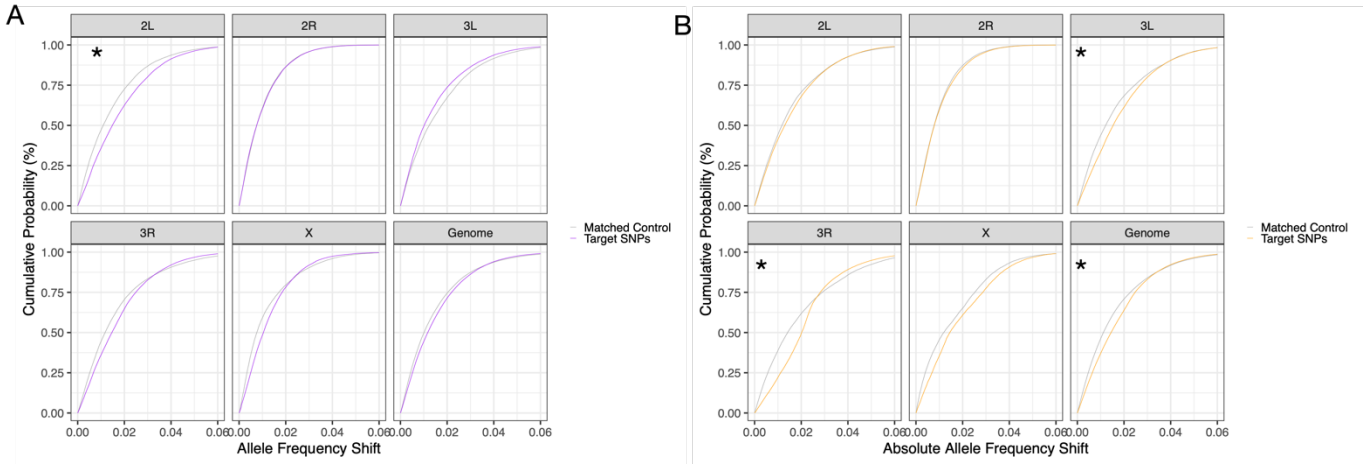

Figure S10. eCDFs of expansion-favored SNPs during truncation, segregated by those with sustained direction of movement (A) or reversals in direction (B). Asterisks correspond to instances in which the observed eCDF differed significantly from that generated via matched control SNPs (grey lines) (one-tailed Kolmogorov-Smirnov test, p-value < 0.05)

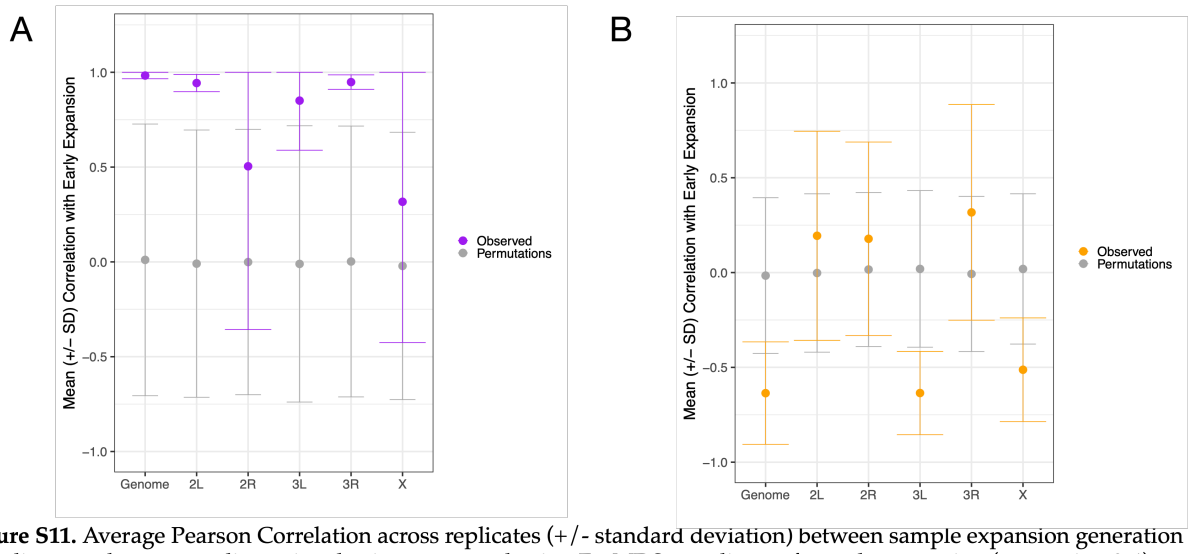

Figure S11. Average Pearson Correlation across replicates (+/- standard deviation) between sample expansion generation and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples. Purple points and error bars correspond to empirical values, while grey points and error bars correspond to values derived from empirical permutations (N = 100). Here, Fst values were derived from SNPs excluding those within, or up to 100 Kb away from, inversion breakpoints.

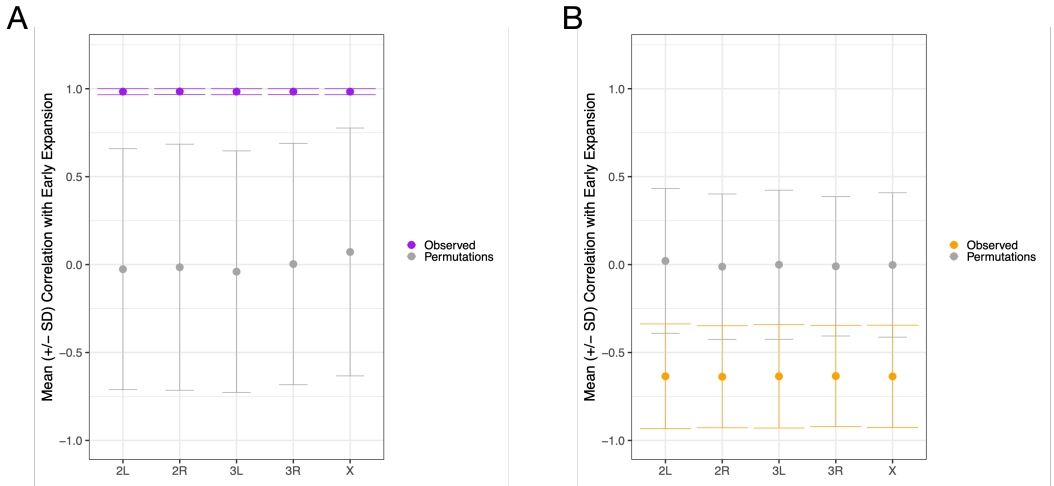

Figure S12. Average Pearson Correlation across replicates (+/- standard deviation) between sample expansion generation and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples. Purple points and error bars correspond to empirical values, while grey points and error bars correspond to values derived from empirical permutations (N = 100). Here, Fst values were derived from randomly sampled SNPs throughout the genome matched to the total number of SNPs present on each chromosomal arm and excluding those within, or up to 100 Kb away from, inversion breakpoints.

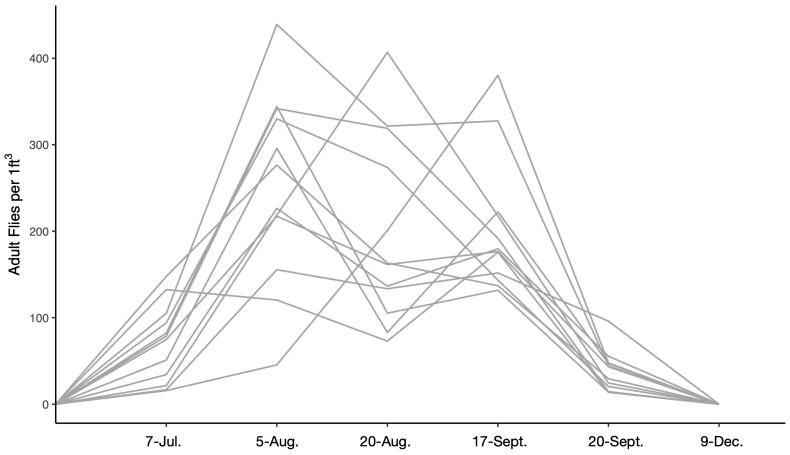

Figure S13. Census estimates of flies in the outdoor mesocosm experiment from 2021. Figure modified from Bitter *et al* . (2024).

Table S1 Statistics for projection of late expansion (generations 6-9) and truncation samples onto a one-dimensional axis constructed using F_{ST} MDS coordinates of early expansion (generation 0-4) samples. Correlation coefficients were computed between sample collection time point and distance along this axis for each cage, and reported are the median values across cage for each chromosomal arm, as well as when the analysis was run genome-wide. P-values were generated via comparison of each median correlation to a distribution generated via permutation of time point labels (N = 100 total permutations) .

Test Interval	Chromosome	Median Correlation	Permutation p-value
Late Expansion	2L	0.97213201683991	0.0337499999999999
Late Expansion	2R	-0.244238340356563	0.375
Late Expansion	3L	0.972604523272359	0.08875
Late Expansion	3R	0.988165487304083	0.05375
Late Expansion	X	0.315657111616651	0.425
Late Expansion	Genome	0.975609936558498	0.06875
Truncation	2L	-0.0443704606050802	0.43875
Truncation	2R	-0.317142428585079	0.23625
Truncation	3L	-0.578233415528078	0.06375
Truncation	3R	-0.862222504266232	0.005
Truncation	X	-0.504215848697299	0.13
Truncation	Genome	-0.753283710728129	0.01625

Table S3 Inversion Frequency Dynamics | GLM regression summary statistics of the expansion and truncation dynamics of four major cosmopolitan inversions segregating in our outbred population of DGRP lines. The FDR corrected p-value for each inversion during both phases of the experiment (expansion and truncation) are presented, alongside the mean inversion frequency shift in cases in which the GLM regression was significant.

Inversion	Selection Phase	FDR P-value	Frequency Shift
In(2L)t	Expansion	0.094*	-0.97%
In(2L)t	Truncation	0.426	Not significant
In(2R)NS	Expansion	<0.001**	-3.02%
In(2R)NS	Truncation	0.439	Not significant
In(3R)K	Expansion	<0.001**	1.65%
In(3R)K	Truncation	<0.001**	-1.35%
In(3R)Mo	Expansion	0.003**	-7.49%
In(3R)Mo	Truncation	0.017**	-1.73%
In(3R)P	Expansion	0.001**	1.17%
In(3R)P	Truncation	<0.001**	-1.21%

Table S4 Statistics for projection of late expansion (generations 6-9) and truncation samples onto a one-dimensional axis constructed using F_{ST} MDS coordinates of early expansion (generation 0-4) samples. Correlation coefficients were computed between sample collection time point and distance along this axis for each cage, and reported are the median values across cage for each chromosomal arm, as well as when the analysis was run genome-wide. P-values were generated via comparison of each median correlation to a distribution generated via permutation of time point labels (N = 100 total permutations). Here, Fst values and MDS analysis was run on a set of SNPs excluding those within, or up to 100 Kb away from, inversion breakpoints.

Test Interval	Chromosome	Median Correlation	Permutation p-value
Late Expansion	2L	0.941877272974043	0.15875
Late Expansion	2R	0.906664673672926	0.12875
Late Expansion	3L	0.972604523272357	0.0925
Late Expansion	3R	0.949409974764516	0.10375
Late Expansion	X	0.315657111616649	0.39
Late Expansion	Genome	0.988690601491595	0.0855
Truncation	2L	0.193910187619202	0.345
Truncation	2R	0.1106453838666	0.41375
Truncation	3L	-0.578233415528078	0.0775
Truncation	3R	0.516639815656395	0.10625
Truncation	X	-0.504215848697302	0.10125
Truncation	Genome	-0.660159220395907	0.047

Table S5 Statistics for projection of late expansion (generations 6-9) and truncation samples onto a one-dimensional axis constructed using F_{ST} MDS coordinates of early expansion (generation 0-4) samples. Correlation coefficients were computed between sample collection time point and distance along this axis for each cage, and reported are the median values across cage for each chromosomal arm, as well as when the analysis was run genome-wide. P-values were generated via comparison of each median correlation to a distribution generated via permutation of time point labels (N = 100 total permutations). Here, Fst values and MDS analysis was run on SNPs randomly sampled throughout the genome, matched to the number of SNPs present on each chromosomal arm, and excluding those within, or up to 100 Kb away from, inversion breakpoints.

Test Interval	Chromosome	Median Correlation	Permutation p-value
Late Expansion	2L	0.988413197497469	0.0525
Late Expansion	2R	0.989942495115707	0.05125
Late Expansion	3L	0.989390312932461	0.05
Late Expansion	3R	0.990068206319797	0.06375
Late Expansion	X	0.989855205403952	0.08125
Truncation	2L	-0.665094200789921	0.04375
Truncation	2R	-0.661142450550023	0.055
Truncation	3L	-0.660888966235905	0.0525
Truncation	3R	-0.653623292619365	0.03875
Truncation	X	-0.660601423436917	0.045

Table S6 Statistics for comparison of rising allele frequency shifts of indoor-identified SNPs in an outdoor mesocosm during population expansion and collapse. Target values (Mean.Target, Median.Target) indicate those derived from indoor-identified SNPs, while matched values indicate those derived from control SNPs. The distributions of target and matched sites were compared using both a two-tailed t-test and sign test. The resulting test statistics are provided.

Chromosome	Mean.Target	Median.Target	Mean.Matched	Median.Matched	t	FDR.t.test	FDR.sign.test	Test.Interval
2L	0.0319985981378007	3.038013e-02	0.00316831802589668	3.178635e-03	90.4650925090862	0.000000e+00	8.326673e-16	Expansion
2R	0.028969065444205	2.551098e-02	0.00115295987609255	1.615365e-03	66.0420368142833	0.000000e+00	8.326673e-16	Expansion
3L	0.0335003699797639	3.080976e-02	0.00286597441953921	3.928358e-03	90.8841932663739	0.000000e+00	8.326673e-16	Expansion
3R	0.0369291946950844	3.378144e-02	0.00264119051901511	3.484494e-03	129.54052005962	0.000000e+00	8.326673e-16	Expansion
X	0.0288667950506608	2.644817e-02	0.0025943771973813	3.045992e-03	60.462689719572	0.000000e+00	8.326673e-16	Expansion
Genome	0.0333054320018882	3.057965e-02	0.00260823894732216	3.234185e-03	199.627700493635	0.000000e+00	8.326673e-16	Expansion
2L	-0.00487214737932314	-4.692767e-03	-0.00104825093438945	-9.152650e-04	-19.5690789120401	2.356927e-84	7.180716e-52	Collapse
2R	0.000478905306333812	8.868033e-04	-0.000171925992292266	4.171333e-05	2.8948882077324	3.958655e-03	9.993639e-05	Collapse
3L	0.00187994456347232	1.696414e-03	-7.9971808333333e-05	2.173192e-04	8.39801534836484	5.830216e-17	1.057355e-15	Collapse
3R	-0.000895085879399349	-1.074300e-04	0.000631816786767623	6.350982e-04	-8.12116003302136	5.475990e-16	3.364728e-06	Collapse
X	-0.000821591180249633	4.711083e-05	-0.000622166078071464	3.165075e-04	-0.573004050626995	5.666659e-01	3.673839e-02	Collapse
Genome	-0.00109679281216485	-5.765000e-04	-0.000135973486759733	9.451667e-05	-9.42612164714719	4.364745e-21	2.861040e-08	Collapse

Table S7 Final fly census size in indoor cages from paired indoor-outdoor mesocosm study. Census sizes were estimated volumetrically using the total remaining flies at the final collection time point (each replicate was originally seeded with 1,000 total flies).

Replicate	Census estimate
1	67200
2	120000
3	83400
4	67800
5	90000
6	90600
7	97800
8	70200
9	69000
10	51600