Full title: Pervasive fitness trade-offs revealed by rapid adaptation to shifting population densities in large experimental populations of *Drosophila melanogaster*

Short title: Fitness trade-offs in *Drosophila melanogaster*

M.C. Bitter*[1] (mcbitter@stanford.edu), S. Greenblum*[1,2] (greensi@lbl.gov), S. Rajpurohit*[3,4] (subhash.rajpurohit@ahduni.edu.in), A.O. Bergland[1,5] (aob2x@virginia.edu), J.A. Hemker[1] (jahemker@stanford.edu), Egor Lappo[1] (elappo@stanford.edu), N.J. Betancourt[3] (nicolas.betancourt.0011@gmail.com), S. Tilk[1] (susannetilk@gmail.com), S. Berardi[3] (berardis@sas.upenn.edu), H. Oken (hayesoken@gmail.com), P. Schmidt[+3] (schmidtp@sas.upenn.edu), and D.A. Petrov[+1,6] (dpetrov@stanford.edu)

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

Statement of Authorship: 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., E.L., 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 and code 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 allele 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 associated with all analyses conducted in this manuscript are publicly available at the following repository: https://github.com/MarkCBitter/Drosophila-fitness-trade-offs .

Type of Article: Letter
Keywords: eco-evolutionary feedbacks; fluctuating selection; adaptation; trade-offs; population density fluctuations; genomics; experimental evolution; Drosophila melanogaster
Abstract word count: 204
Main text word count: 10,387
Number of references: 61
Number of figures: 6
Number of tables: 0
Number of text boxes: 0
Correspondance: Mark C. Bitter (mcbitter@stanford.edu)

Abstract

Trade-offs are an inherent feature of organismal biology that are expected play a fundamental role in the evolution of natural populations. 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 fluctuating selection on ecological timescales. Specifically, we first conducted a lab-based selection experiment to quantify a genome-wide signal of antagonistic pleiotropy elicited in response to shifting population densities and associated with reproduction and stress tolerance selection. In doing so, we identified a putative role of two cosmopolitan inversions in these trade-offs. We then conducted an independent experiment to show that a simple manipulation of increasing population density under controlled lab-based conditions identified loci that are relevant to selection during population expansion and collapse in a complex, semi-natural setting. In concert, our results reveal how adaptation in complex, natural environments can be coarse-grained in such a manner to drive repeatable and predictable patterns of genomic variation, and further add credence to models positing a role of generic fitness 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 by constraining the simultaneous optimization of traits associated with fitness under different environmental conditions (2,6–8).

At the molecular level, trade-offs likely imply antagonistic pleiotropy, a phenomenon in which an allele that enhances one fitness-related trait may simultaneously impede another in a particular environment. Then, when the environment changes and selection favors the previously impeded trait, the previously favored allele becomes disadvantageous (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 another, putatively more direct, signature of trade-offs are fluctuations in population allele frequencies that occur in concert over time with selective pressures. 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). Evidence of such a trade-off is widespread across species, from soay sheep (20), to seaweed flies (21), to humans (22). It 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 (23–25). 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 (26–28). 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 (10,11,27,29–31).

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). One hypothesis for this predictability at the genomic level is that the loci responding to selection are not sensitive to specific environmental variables (which can be idiosyncratic across study locales and years), but rather respond more coarsely to the generic and repeatable demographic fluctuations in the system (i.e., shifts in population density). The observed parallel genomic signals of temporally varying selection in this system may then represent the manifestation of fundamental fitness trade-offs emerging in response to these shifting population densities. Systematically and rigorously testing this hypothesis, however, necessitates controlling for the complex suite of abiotic shifts that occur in outdoor

141 environments and isolating the relative contribution of demographic selection to patterns of
142 allele frequency change.
143 Here, we conducted two experimental evolution studies to systematically quantify how
144 predictable evolutionary trajectories arise in complex, fluctuating environments and, in effect,
145 explicitly link key components of life history evolution to patterns of fluctuating selection and
146 maintenance of variation in *D. melanogaster*. In the first of these studies, we assayed the impact
147 of demographic shifts, independent from any environmental selection. To do so, we monitored
148 patterns of genomic variation in genetically diverse, replicate populations housed in a
149 controlled, indoor environment throughout eleven discrete (non-overlapping) generations of
150 population expansion, and a single bout of population truncation (Fig. 1; Fig S1). Our replicate
151 populations were derived from a reconstituted, outbred population initiated from 145 inbred
152 (DGRP) lines (Methods; Supplementary Data File 1), and evolution proceeded in four replicate
153 cages (2 m^3). Food availability in the replicate cages was commensurate with increasing
154 population sizes during the expansion phase, generating a selective regime that eliminated the
155 impact of intraspecific competition (e.g. for food or egg laying space) and favored the
156 reproduction-associated traits expected when resources are abundant, such as fecundity and
157 developmental rate (27). Contrarily, population truncation was induced by removing all food
158 and water from our replicate cages and sampling the surviving individuals as the replicates
159 collapsed. This mirrors dynamics of a population that has exceeded carrying capacity and, in
160 effect, selects for the ability to withstand depleted resources via stress-tolerance mechanisms
161 (32), which in our experiment is most likely to be desiccation resistance (33). We hypothesized
162 that these contrasting selective regimes (hereafter, 'reproduction' and 'stress-tolerance'
163 selection) would reveal trade-offs between life-history variation, whereby antagonistically
164 pleiotropic alleles favored during sustained population expansion would become selected
165 against during truncation.
166 The second experiment leveraged an independent inbred panel to construct a genetically
167 diverse, outbred population, which was split into a set of indoor cages (N = 10) and outdoor
168 mesocosm (N = 12) replicates. Through this, we used the same mapping population (thereby
169 eliminating any confounding impacts of differences in recombination and/or linkage
170 disequilibrium) to directly test whether loci identified under selection during increasing
171 population density in a controlled, indoor environment are relevant to selection and trade-offs
172 during adaptation in an outdoor, naturally fluctuating environment. In concert, our results
173 provide evidence for: (1) strong, parallel selection elicited by changes in population density (2)
174 genome-wide tradeoffs associated with contrasting selective regimes of population expansion
175 and truncation, (3) the role of cosmopolitan inversions underpinning fitness trade-offs in the

species and (4) relevance of the loci responding to density selection in a controlled, lab-based setting to patterns of adaptation in response to natural environmental fluctuations.

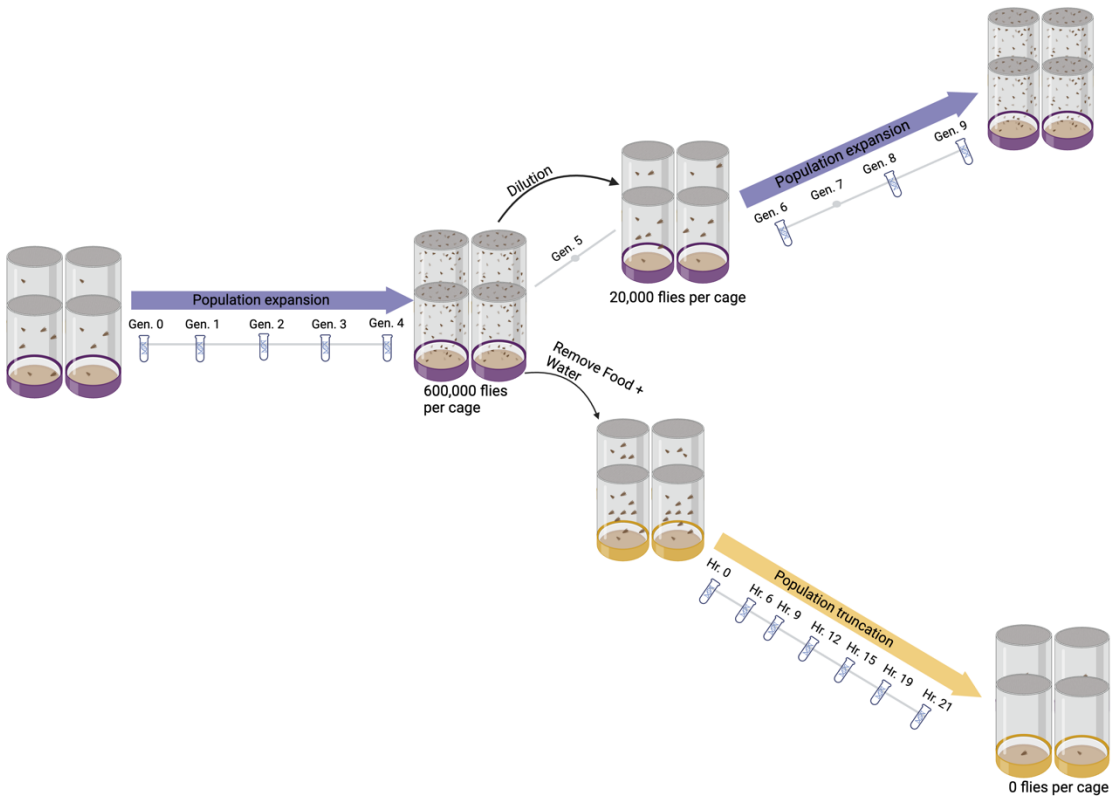

Figure 1. Schematic of lab-based, population expansion-truncation selection experiment | Four replicate, 2 m^3 cages were seeded with a genetically diverse, outbred population generated via four generations of recombination (N = 145 inbred lines). Evolution proceeded in discrete generations, whereby 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/).

195 **Results**

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

197 We observed parallel shifts in allele frequencies across the four replicate populations
198 throughout the progression of population expansion, indicating adaptive responses to the
199 shared selective pressures imposed in our experimental system. We quantified these systematic
200 allelic shifts using several approaches. We first computed genome-wide divergence as average
201 F_{ST} across all segregating sites pairwise among all expansion samples and used these data to
202 carry out several tests. F_{ST} divergence from the generation 0 populations increased steadily
203 within each cage through time, readily exceeding differentiation between biological replicates (F
204 = 172.2; p-value < 0.01; Fig. 2A & Fig. S2). We then used pairwise F_{ST} values as a distance metric
205 to create multi-dimensional scaling (MDS) plots, in which divergence between samples is
206 represented as distance among the points in a 2-D plane. Coloring samples based on the
207 replicate cage identity indicated that genome-wide allele frequencies in one replicate were
208 perturbed, shifting its points from the remaining three replicates throughout expansion (Fig.
209 2B). We hypothesize this was likely a result of potential bottlenecks during replicate cage
210 founding. However, despite this offset, samples from all cages appeared to shift across the 2-D
211 plane in the same direction through time, suggesting that the replicate populations experienced
212 parallel shifts in genome-wide allele frequencies (Fig 2C and Fig S3; a trend that was also
213 detected using principal component analysis; Fig. S4).

214 To more rigorously quantify the systematic movement of samples through time
215 observed in the two-dimensional MDS space, we translated coordinates for each cage such that
216 the centroid of all generation 0 samples was centered at the origin. We then used these
217 translated points to fit a simple linear regression model to samples from expansion generations
218 0-4 (Methods; Fig. S5). The resulting axis represents the primary axis of variation in the 2D
219 plane during early expansion for each replicate. We hypothesized that if sustained, directional
220 selection imposed by population expansion was a primary driver of patterns of genomic
221 variation, samples from the remaining expansion generations (generations 6-9) would continue
222 to proceed along the established axis of variation in the same direction. Indeed, projection of
223 late expansion samples onto this axis of variation indicated a significant correlation of sample
224 collection generation and distance along the axis (Fig 2D.; significance of correlation determined
225 via permutations). When segregating this analysis by chromosomal arm, however, this
226 parallelism was only evident on chromosomal arms 2L, 3L, and 3R ; while correlations derived
227 from SNPs on 2R and X were indistinguishable from permuted values (Fig. 2D). We re-sampled
228 SNPs across the genome iteratively to match the number present on each chromosomal arm to
229 confirm that the variation in parallelism was not a technical artifact. In each case, our sub-
230 sampling yielded correlations that matched the genome-wide trend, indicating that differences

231 observed among chromosomal arms (Fig. 2D) reflect systematic differences in the behavior of
232 loci across the genome (Fig. S6, Table S2).
233 To determine the distribution of SNPs (N = 1.7 M SNPs) contributing to the parallel
234 patterns of divergence across cages, a generalized linear model was fit to allele frequencies from
235 all expansion samples (generation 0-9). This model assessed the significance of the linear
236 relationship between allele frequency and generation of sampling across all cages, using a
237 quasibinomial error model to reduce false positive associations (15,16,34). While an association
238 between allele frequency and sampling timepoint within a single cage may represent either drift
239 or selection, a significant parallel association across all four replicate cages indicates allele
240 frequency trajectories that are both predictable over time and parallel across populations. In this
241 case, selection (or linked selection) is the more parsimonious explanation. A total of 389,588
242 SNPs (22.9 % of all sites), across all chromosomal arms, showed significant parallelism after
243 multiple testing correction (Benjamini-Hochberg false discovery rate <0.01 and allele frequency
244 change > 2%). These SNPs showed systematic, directional selection across replicates throughout
245 expansion (Fig. 2E; Fig. S7). Furthermore, the genomic distribution of these SNPs spanned all
246 five chromosomal arms and were located both within and outside the breakpoints of five major
247 cosmopolitan inversions that segregated at appreciable frequency (>4%) in our inbred reference
248 panel (Fig. 2F).
249 We next used our GLM signal to infer the minim number non-overlapping, putatively
250 unlinked loci, underpinning patterns of adaptation throughout expansion (ref; Methods). This
251 resulted in the identification of 250 total loci, ranging in size from 30 kb to 305 kb and located
252 across all chromosomal arms (Supplementary File X). Allele frequency trajectories of the most
253 significant SNPs within each clusters indicate strongly directional selection throughout
254 expansion, with median selection coefficients of 7% per generation (Fig S8). The number of
255 unlinked loci and the inferred strength of selection is on par with similar estimates obtained
256 from outdoor mesocosom and wild population sampling (14–16) While determining the
257 underlying causal regions (which could be either SNPs or larger structural variants) within
258 these unlinked loci requires further investigation, this analysis at least suggests adaptation in
259 response to increasing population densities to be underpinned by a fairly large number of
260 independent variants (i.e. on the order of hundreds), each containing alleles exhibiting
261 relatively large selection coefficients. We further note that while long-range patterns of linkage
262 could theoretically induce non-independent patterns of allele frequency change among loci on
263 the same chromosomal arm, we have previously shown that how this method, which identifies
264 independent loci based on patterns of physical linkage, produces sets of loci with independent
265 allele frequency movement throughout evolution (14).

Finally, 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.

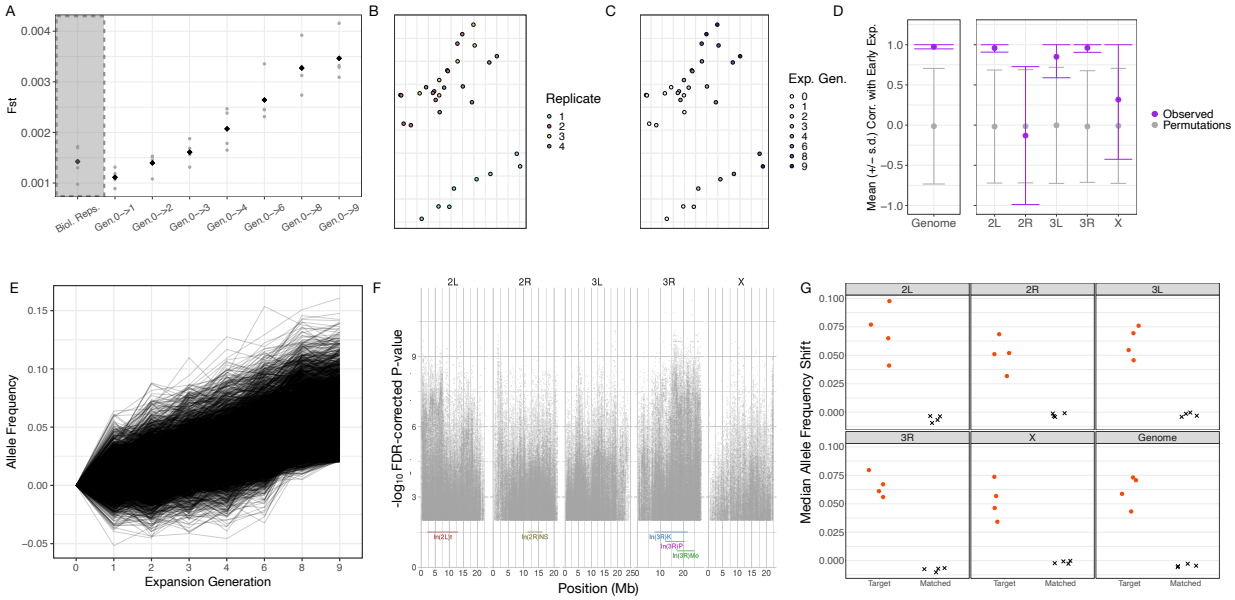

Figure 2. Adaptation under sustained population expansion elicits strong, genome-wide parallel responses | (A) Mean, genome-wide F_{ST} between biological replicates ('Biological Replicates'; same replicate/collection time point, different pooled sample/extraction of flies) and evolved replicate samples 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

291 correspond to values derived from N = 100 permutations. (E) Trajectories of rising alleles at SNPs identified via GLM
292 (FDR < 0.01 and effect size > 2%) across all nine generations of expansion. (F) Genomic distribution of SNPs depicted
293 in (E). The coordinates of the major cosmopolitan inversions segregating in our founding reference panel at greater
294 than 4% frequency in our founding strains are depicted above the x-axis. (G) Leave-one-out cross-validation to infer
295 replicate-specific parallelism of adaptation to sustained population expansion. Portrayed is the median shift of the
296 rising allele for sets of target SNPs (points) identified via GLM (FDR < 0.01 and effect size > 2%) in 3 of 4 cages,
297 relative to medians derived from a matched control SNP set (X's), in the left-out replicate. Target SNP medians are
298 colored red if the distribution of phased allele frequency shifts was significantly greater than that of matched control
299 sites (two-tailed t-test, FDR < 0.05).
300
301 *Genomic evidence of trade-offs induced by fluctuating selection across expansion and truncation*
302 The parallel frequency shifts observed throughout expansion may be the product of
303 three, not mutually exclusive, evolutionary dynamics: (1) directional selection in response to
304 sustained fecundity selection, (2) adaptation to the lab environment, and/or (3) the purging of
305 recessive deleterious alleles (i.e., negative selection) in outbred population. To disentangle these
306 various dynamics and, in turn, identify the presence of fitness-tradeoffs at putatively selected
307 alleles, we leveraged our samples collected throughout truncation selection. Alleles identified
308 during expansion that were a product of consistent lab selection and/or negative selection
309 against unconditionally deleterious recessives should continue to show systematic, directional
310 change throughout truncation (as their effect is not conditional on the specific treatment).
311 However, alleles with treatment-specific behavior during expansion and truncation (i.e.,
312 moving in the opposite direction) represent those that likely underpin trade-offs between
313 fecundity and stress tolerance selection.
314 We tested for evidence of context-specific behavior and trade-offs of genome-wide SNPs
315 by re-conducting our MDS analysis of pairwise divergence values, this time including all
316 samples collected throughout expansion and truncation. If the dominant direction of allele
317 frequency change was sustained across both expansion and truncation (indicating sustained lab
318 selection or purging of deleterious mutations), truncation samples would be ordered from early
319 to late in a parallel manner as the expansion samples (Fig 2B). Instead, we observed that
320 samples taken during truncation (initiated at expansion generation 5) shifted back towards
321 earlier expansion samples, potentially suggesting a genome-wide reversion of allele frequencies
322 (Fig. 3A; Fig. S9). We quantified these trends as above, translating the MDS coordinates of each
323 truncation sample such that the centroid of all hour 0 samples was centered at the origin and
324 then projecting them onto a the single axis linear regression model derived from early
325 expansion points (Fig. 2D; Fig. S5). As hypothesized, correlations between collection time and
326 distance along this axis were, genome-wide, significantly negative (Fig. 3B; Table S1).
327 Segregating this analysis by chromosomal arm yielded more nuanced dynamics whereby 3L,

328 3R, and X yielded evidence of anti-parallel/reversions in allele frequencies, while 2L and 2R
329 exhibited correlations that were insignificant relative to our permuted distributions (Fig. 3B).
330 Again, as above, we validated that these differences among chromosomal arms were not simply
331 a reflection of variation in number of SNPs available for measurement and reflected systematic
332 differences in allele frequency dynamics across the genome (Fig. S10).
333
334 We garnered further evidence of trade-offs across fecundity and desiccation selection
335 using allele frequency trajectories. Specifically, we selected those SNPs with evidence of
336 systematic movement across replicates throughout expansion (GLM FDR < 0.01 and allele
337 frequency shift > 2%) and measured the average frequency shifts of the rising allele across cages
338 throughout truncation. If patterns of fluctuating selection and antagonistic pleiotropy
339 dominated the data, we expected trajectories to be enriched for SNPs moving in the opposite
340 (negative) direction during truncation. Indeed, across all expansion-identified SNPs, the ratio
341 was significantly skewed towards those with exhibiting fluctuating selection, as opposed to
342 sustained directional selection, across expansion and truncation selection ($\chi^2 = 1123.6$; P-value <
343 0.001). These trajectories are depicted in Figure 3C, in which colored trajectories indicate those
344 expansion-identified SNPs with additional, independent evidence of systematic movement
345 across cages during truncation (GLM FDR < 0.05 and effect size > 1%): purple for those SNPs
346 with consistent directional selection across selection regimes, and orange for those with anti-
347 parallel movement across regimes. As expected, this subset of SNPs (N = 24,814) exhibited even
348 greater evidence of fluctuating selection and were 3.5 times more likely to fluctuate directions
349 across treatments than exhibit sustained directional selection (19,158 relative to 5,656 SNPs,
350 respectively). We assessed how consistent this dynamic was across the genome by quantifying
351 the ratio of parallel vs. fluctuating selection genome-wide, and separately for each chromosomal
352 arm. As expected based on the F_{ST} MDS analysis described above, there was strong evidence of
353 enrichment for fluctuating selection on chromosomal arms 3L, 3R, and X. However, this
354 analysis revealed an enrichment of 2L SNPs with sustained parallel selection across selection
355 regimes, which was an undetected in our F_{ST} MDS analysis. We explored this dynamic further,
356 using the underlying allele frequency shift distributions to compute empirical cumulative
357 distribution functions, separately for alleles with sustained directional selection and for those
358 with evidence of fluctuating selection (See Supp. Mat.). This verified a genome-wide enrichment
359 of alleles exhibiting fluctuating selection (driven by chromosomal arms 3L, 3R, and X), and
360 evidence of sustained directional selection on 2L (Fig S11). This signal of fluctuating selection
361 was further corroborated by an assessment of the behavior of unlinked loci identified during
362 reproduction selection throughout truncation (Supplementary Data File 2, Figure S12). In sum,

we find overwhelming 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 allele frequency data from which the trajectories were computed are entirely independent across expansion and truncation (i.e., have no overlapping time points).

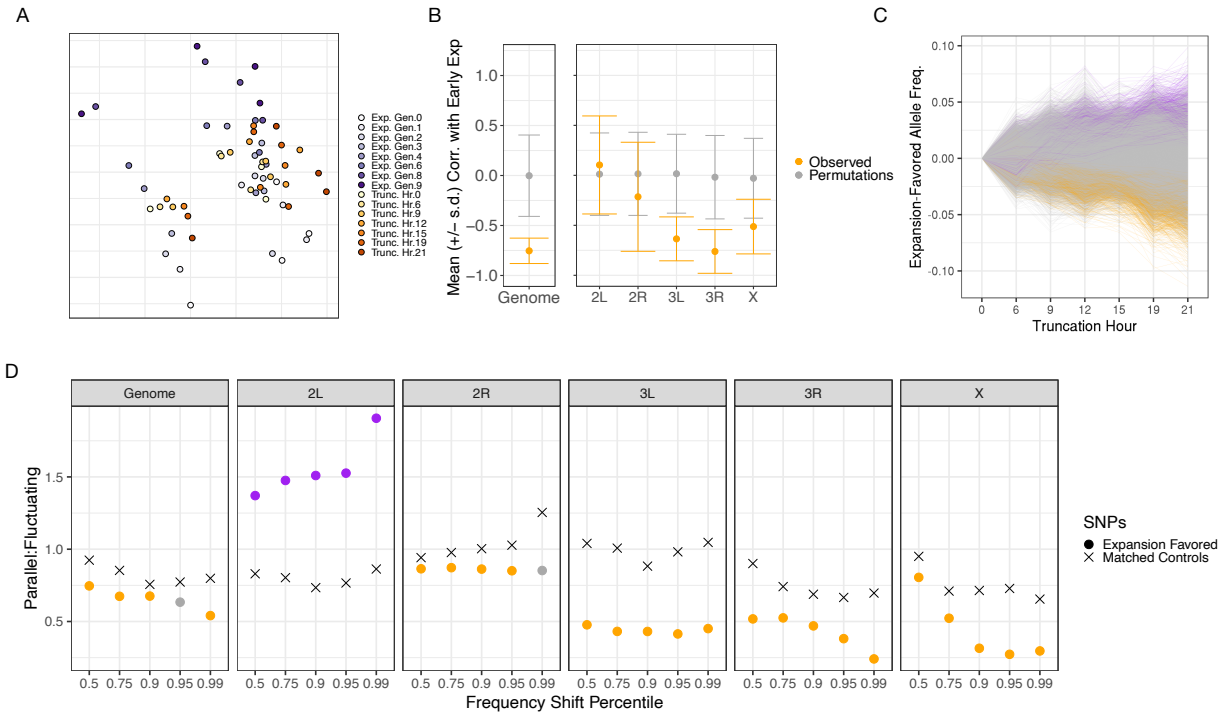

Figure 3. Genome-wide signal of fitness trade-offs between population expansion and truncation | (A) MDS of pairwise F_{ST} values across all samples collected throughout expansion (purple-hue points) and truncation (orange-hue points), shaded according to collection time point (darker hues indicate 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

386 frequency shift. Points are further colored if the ratio of parallel to fluctuating behavior differed significantly from
387 that expected based on a matched control set (χ^2 test P-value < 0.05).
388
389
390
391 *One and two locus simulations of linkage, selection, and antagonistic pleiotropy*
392 We explored whether patterns of linkage between loci, antagonistic pleiotropy at a
393 single SNP, and/or haplotype-level antagonistic pleiotropy could generate the patterns of allele
394 frequency change observed in our data using Wrigh-Fisher simulations in SLiM (35). Briefly, we
395 first simulated a scenario in which a single SNP had an advantageous effect on fitness during
396 expansion, which was countered by a disadvantageous effect during truncation. As expected,
397 this produced frequency trajectories across reproduction and stress tolerance selection that
398 recapitulated the patterns observed in our empirical data (Fig S13).
399 Next, we focused on a series of two-locus scenarios in which SNP 'A' was directionally
400 favored during expansion with neutral fitness effects during truncation, while SNP 'B' had
401 neutral fitness effects during expansion and then came under directional selection during
402 truncation. We then computed fitness values (empirical selection coefficients) during the first
403 generation of truncation (which represents the single generation (no-recombination) of stress
404 tolerance selection in in our experiment) to infer how the relationship between the two SNPs
405 may influence their behavior across phases. Under scenarios in which the two loci were
406 unlinked, their frequency trajectories were largely independent across phases: SNP 'A'
407 increased systematically during expansion with neutral dynamics during truncation, while SNP
408 'B' exhibited neutral dynamics during expansion and then systematic increases during
409 truncation (Fig S14-S16). Under a scenario of high linkage and a net 'attraction' between the
410 favored allele at each SNP (i.e., the favored allele at SNP 'A' was non-independently assorted
411 with the favored allele at SNP 'B'), trajectories were largely correlated. For example, SNP 'A'
412 increased in frequency systematically across both expansion and truncation, causing empirical
413 selection coefficients to indicate a selective advantage during truncation even though the locus
414 was in fact neutral during this phase (Fig S14; Fig S17-18). Alternatively, under a scenario of
415 high linkage and net 'repulsion' between the favored alleles at each SNP (i.e., the favored allele
416 at SNP 'A' was non-independently assorted with the disadvantageous allele at SNP 'B'), SNP
417 'A' exhibited a trajectory of fluctuating selection whereby it systematically increased in
418 frequency during expansion and then declined during truncation (Fig S14; Fig S19-20). This
419 scenario thus represents a form of 'haplotype-level' antagonistic pleiotropy driven by the
420 orientation of causal alleles and the underlying haplotype structure in the population. Finally,
421 under a scenario in which the attracted and repulsed states of the selected alleles occur at equal

422 probability (i.e., random assortment of alleles between loci), the net-effect of possible linkage
423 patterns produces allele frequency trajectories and empirical fitness values that recapitulate the
424 dynamics observed in the 'unlinked' scenario (Fig S14; Fig S21-22).
425 As we have no prior knowledge of a non-independent assortment of selected alleles in
426 our focal population, the final simulated scenario (independent assortment of advantageous
427 alleles between locus 'A' and 'B') likely represents the most plausible dynamic occurring in our
428 experiment, should two linked SNPs be under differential selection across selection regimes.
429 Therefore, within the confines of the current data (i.e., not yet knowing the underlying causal
430 loci), we conservatively interpret our observations of reversions in allele frequency trajectories
431 as antagonistic pleiotropy at a single SNP, though recognize the possibility that it could in fact
432 represent antagonistic pleiotropy at the haplotype-level. Understanding how such haplotype
433 structure could be maintained over long time periods, via either strong selection for the
434 'repulsed' linkage between selected alleles or some form of recombination suppression (e.g., an
435 inversion), will ultimately hinge upon resolving the causal loci in this system.
436
437 *The role of inversions in reproduction/stress-tolerance trade-offs*
438 Theoretical, and emerging empirical, research suggests a dominant role of chromosomal
439 inversions in adaptation in natural populations (36). Accordingly, we quantified whether five
440 major cosmopolitan inversions in *D. melanogaster* (*In(2L)t*, *In(2R)NS*, *In(3R)K*, *In(3R)P*,
441 *In(3R)Mo*), which all occurred at starting frequency > 4% in our inbred reference panel,
442 exhibited dynamics consistent with adaptation and, if so, trade-offs between population
443 expansion and truncation (Supplementary Data File 3). Generalized linear regression of
444 inversion frequencies yielded evidence that all inversions shifted systematically across
445 replicates during expansion, with magnitude of frequency shifts ranging between 1 and 7%
446 (Fig. 4A-E; Table S2). *In(3R)K* and *In(3R)P* were the only two inversions that exhibited evidence
447 of trade-offs across selection regimes, whereby each inversion was systematically favored
448 during expansion and then became selected against during truncation (Fig. 4A-B).
449 Given the appreciable effect of inversions on patterns of adaptation throughout the
450 experiment, we next explored the extent to which our inference of trade-offs was solely a
451 function of variation within, or in tight linkage between, these structural variants. Specifically,
452 we reconducted F_{ST}-based MDS analysis on a subset of SNPs, excluding all those within, or 100
453 kb away from, inversion breakpoints. Through this, we observed that the signature of trade-offs
454 was maintained on 3L and X, but eliminated on 3R. Still, the genome-wide signal of this
455 analysis continued to indicate parallel movement across early and late-expansion samples, and
456 that a signal of fluctuating selection dominates patterns of genomic variation between

expansion and truncation (Fig. S23). 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. S24).

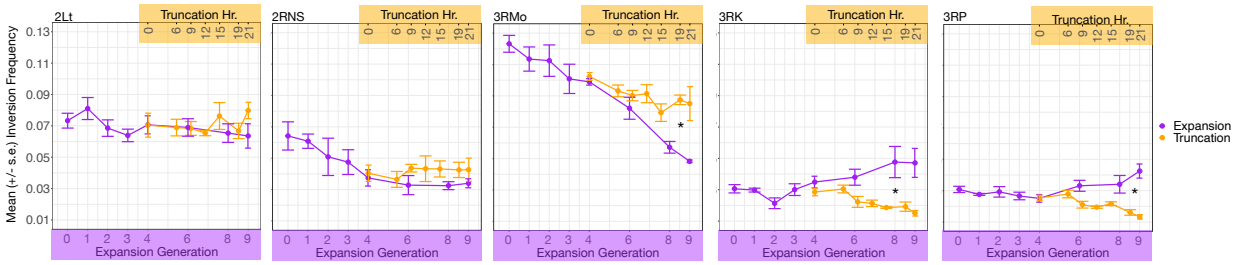

Figure 4. Role of inversions in reproduction and stress-tolerance trade-offs | Mean inversion frequency changes (+/- s.e.) across expansion generations (purple trajectories) and truncation hours (orange trajectories). Asterisks indicate inversions with evidence of fluctuating selection across expansion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

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. S25). 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 4). 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. S25; 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). Interestingly, we failed to detect such significant statistical overlap using the set of loci identified during the reproduction selection experiment described above, which used a different inbred reference panel (see Methods). Next, we asked whether the dominant direction

of selection was parallel across the indoor and outdoor mesocosm 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 significant genome-wide reversion in allele frequencies in the outdoor cages throughout this period, a dynamic most primarily driven by patterns of variation on 2L and 3R (Fig. 5C; Table S6). This shift in the dominant direction of indoor-identified SNPs indicates that their signatures of selection are not simply driven by the purging of deleterious recessives in our reference panel, but rather exhibit context-dependent fitness effects across population expansion and truncation.

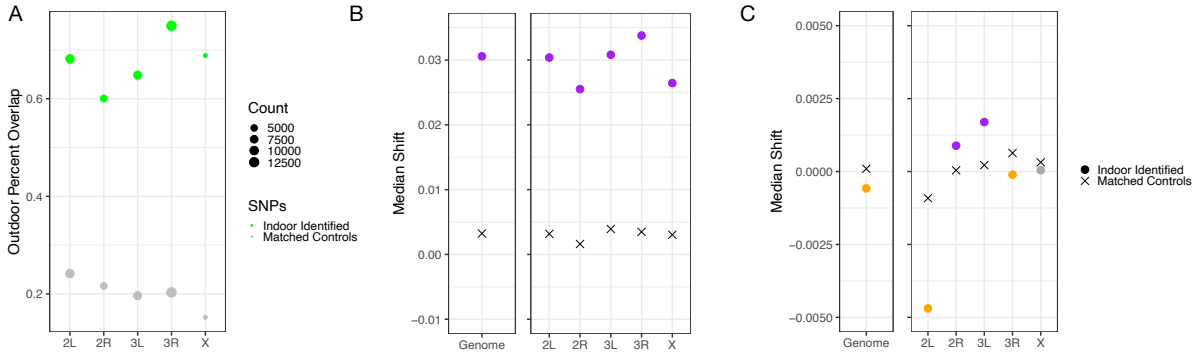

Figure 5. Selection on reproduction under controlled, lab-based conditions predicts patterns of adaptation and trade-offs in outdoor mesocosms | (A) Observed vs. expected overlap in SNPs with systematic allele frequency movement (GLM FDR < 0.01; allele frequency change > 2%) in a paired indoor and outdoor 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).

516
517
518 **Discussion**

519 Here, we used experimental evolution in large, genetically diverse populations of *D.*
520 *melanagoster* to quantify genome-wide evidence of fitness trade-offs in response to selection
521 elicited in the context of ecologically realistic population boom and bust dynamics. We first
522 manipulated reproduction selection via nine generations of sustained population expansion in
523 the absence of density regulation mechanisms. This process induced parallel, genome-wide
524 shifts in allele frequencies across four, independent replicate populations. The strong selection
525 coefficients quantified throughout this period (~7% per generation) are comparable to those
526 quantified during sampling between summer and fall in wild populations and outdoor
527 mesocosms (13–16). Thus, while it has been previously speculated that such patterns of
528 evolution of *D. melanogaster* across seasons may be dominated by adaptation to shifting
529 temperatures (e.g., Boulétreau-Merle, Fouillet, and Terrier 1987), our lab-based manipulation
530 here provides direct evidence that the sole impact of shifting population densities can drive
531 rapid adaptation over the course of several generations.

532 By imposing a bout of truncation selection midway through population expansion,
533 during which time individual survival and associated patterns of allele frequency change
534 revealed differences in relative stress tolerance (e.g., desiccation resistance) among genotypes,
535 we aimed to explore evidence of fitness-tradeoffs in the system. We quantified a pervasive,
536 genome-wide signal of fluctuating selection, whereby patterns of allele frequency movement
537 reversed direction across selection regimes. These patterns are indicative of antagonistic
538 pleiotropy, whereby selected alleles conveying advantageous trait values for the phenotypes
539 favored during reproduction selection convey disadvantageous trait values for the suite of
540 phenotypes favored during truncation (7,37). While this dynamic was *a priori* expected based on
541 the negative genetic correlations between traits likely under selection in our experiment (10,29–
542 31), few data exist directly observing the manifestation of such trade-offs, particularly over the
543 ecologically-relevant timescales assayed in our experiment. The key implication of this
544 antagonistically pleiotropic behavior is that this process may, in effect, maintain variation in the
545 population (7,8). Specifically, mutations with unconditionally advantageous effects will
546 ultimately be driven to fixation, while those with context-specific behavior can in principle be
547 maintained for longer periods of time in the presence of fluctuating selective pressures (38,39).
548 The key challenge for future work will be to discern whether these dynamics indeed balance
549 alleles over long time-periods, which will ultimately be aided by refining the causal loci driving
550 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 and, more recently, in trade-offs (21). Mechanistically, suppressed recombination within the breakpoints of relatively young inversions can give rise to the accumulation of putatively adaptive alleles (40–43), and over long periods maintain allelic structure within populations that generate haplotype-level, antagonistically pleiotropic fitness effects (see *One and two locus simulations of linkage, selection, and antagonistic pleiotropy,* and Fig. S14). In natural populations of *D. melanogaster*, inversions have indeed been 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 (15,44–46). 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 demonstrated that inverted karyotypes of *In(3R)P* are shorter-lived and less stress resistant than non-inverted karyotypes, as would be predicted based on the frequency trajectories of this inversion in our experiment (47). Still, resolving the functional relevance of each of these major inversions is nascent and provides a fruitful avenue of future research.

It is noteworthy that the trade-offs quantified here manifested during a single bout of truncation selection, which proceeded over the course of just 24 hours. This illustrates how finely adaptation via standing variation can track shifting environmental conditions, and that selection can be detected on temporal scales shorter than a single generation. Indeed, mounting research has shown how both anomalous (e.g., heat wave, hurricane, or poaching) and non-anomalous (e.g., weekly to seasonal shifts in abiotic and biotic conditions) environmental perturbations can elicit patterns of adaptation on similar timescales in natural populations (13–16,46,48–51). This growing body of research increasingly suggests that the selection coefficients underpinning the standing, functional variation in natural populations may be substantially larger than previously recognized, and ultimately augment adaptation to the accelerated shifts in mean environments associated with global climate change.

586 Our paired indoor cage and outdoor mesocosm experiment demonstrated how selection
587 induced by population expansion under controlled, indoor conditions can identify alleles that
588 are relevant to adaptation and trade-offs in response to natural environmental fluctuations. The
589 conditional benefit of the indoor-identified loci in the outdoor mesocosm (i.e., their genome-
590 wide advantageousness during population expansion and genome-wide disadvantageousness
591 during truncation; Fig. 5) indicates that these the shared signatures of selection across
592 environments do not simply represent a set of loci under negative selection (e.g., purging of
593 deleterious recessives), but rather subject to some form of balancing selection imposed by the
594 temporally fluctuating environment. In turn, this finding indicates that the repeated signals of
595 fluctuating selection quantified in *D. melanogaster*, both across populations and through time
596 (13–16), may be underpinned by fundamental life-history trade-offs that emerge as a result of
597 shifts in the selective environment induced by population boom-bust demographic dynamics.
598 These boom-bust demographic dynamics are a generic feature of populations across taxa, and
599 this phenomenon may thus be more widespread than previously recognized and provide a key
600 instance of the interplay of ecological and evolutionary forces in natural populations (52–55).
601 Furthermore, given the dramatic differences in indoor and outdoor environments, the shared
602 allele frequency patterns observed here suggest that core alleles may underpin adaptation
603 during generic boom and bust cycles, regardless of specific abiotic condition. This bolsters the
604 hypothesis that the predictability of adaptation in this system is augmented by the nature of the
605 underlying genetic variation: loci responding to rapidly fluctuating selection are not specific in
606 their response to particular abiotic parameters, but rather respond more coarsely to generic and
607 repeatable features to the selective environments, in this case the population density
608 fluctuations that occur yearly in the system. Such a 'coarse-graining' of the architecture of the
609 adaptive response would position life-history trade-offs and fluctuating selection as a key force
610 in maintaining variation in natural populations. Ultimately, fine-mapping the specific causal
611 loci underpinning these trade-offs and signatures of fluctuating selection is a key area of future
612 research that, as the data presented here suggest, may even be feasible within the context of
613 highly replicated indoor experiments. Finally, we note the possibility that a unique set of loci
614 become fitness-relevant under specific abiotic contexts, our goal here was to simply query
615 whether a common set of loci underpinned trade-offs and responded generically to fluctuating
616 selection across abiotic contexts, for which we found substantial evidence.
617 It is notable that we observed a putative difference in the architecture of trade-offs
618 between the indoor population expansion/truncation selection experiment, and the paired
619 indoor-outdoor mesocosm study. For example, chromosomal arm 2L displayed evidence of
620 sustained selection across reproduction and stress tolerance selection in the first experiment, but
621 exhibited strong evidence of antagonistic pleiotropy between population expansion and

622 collapse during our outdoor mesocosm experiment. Furthermore, we did not detect any shared
623 enrichment of SNPs with evidence of linked selection across these experiments. While there
624 were distinct methodological and environmental differences that may have played a role in this
625 discordance, a more salient possibility is the use of a different set of inbred lines to generate the
626 outbred mapping population for each experiment. As our analyses ultimately identify sets of
627 SNPs in tight linkage to an underlying causal locus, differences in patterns of linkage across
628 mapping populations could, in effect, lead to dramatic differences in the relative effect size of
629 marker alleles across studies. This process is analogous to the oftentimes poor portability of
630 association studies across human populations (56,57), and is an important consideration for
631 future research using experimental evolution approaches to characterize the architecture of
632 complex trait adaptation.
633 In conclusion, the set of experiments conducted here demonstrate how adaptation
634 associated with population boom-bust ecology impose selection on a core set of loci, eliciting
635 trade-offs and driving predictable evolutionary dynamics in populations of *D. melanogaster*. This
636 dynamic, in turn, likely acts as a key force maintaining variation in the species, and potentially a
637 broader range of taxa experiencing similar boom-bust demographic dynamics. Finally, this
638 study demonstrates how well-resolved time-series genomic data can reveal the presence of, and
639 genomic architecture underlying, fitness trade-offs.
640
641 **Methods**
642 *Population construction and replicate cage seeding for expansion/truncation selection experiment*
643 We constructed a genetically diverse founder population via outbreeding 145 lines of the
644 Drosophila melanogaster Genetic Reference Panel (DGRP; Mackay et al. 2012). Ten mated
645 females per DGRP inbred line, each of the same age cohort and from density-controlled line
646 cultures, were pooled into a single 0.3m x 0.6m x 0.3m cage (P0 generation). Over the course of
647 4d, eggs were collected in 64 culture bottles (P1 generation) using cornmeal molasses medium
648 (16 bottles added per day, then capped, removed, and 16 new bottles added); the 64 bottles
649 were then randomly assigned to 1 of the 4 replicate cages; from this point forward, all replicates
650 were cultured independently. Once flies eclosed, they were again released into 4 replicate 0.3m
651 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;
652 1L of Drosophila cornmeal molasses medium per tray) per replicate. Once these flies eclosed
653 (the first true F1 generation), the flies were released into 4 replicate, medium sized cages (0.6m x
654 0.6m x 1.2m) and allowed to oviposit on 4 trays of culture medium over 24h (embryos were the
655 F2 generation). The F2 embryos from these trays were sealed and collected, and the adults
656 discarded. Once the F2 flies eclosed, they were then released into large, experimental cages (3
657 m^3) and given 8 trays of media for oviposition over 24h.

658 *Population expansion/truncation selection in large, indoor cages*

659 Culturing within the large, experimental cages proceeded via discrete generations,
660 throughout which the amount of food was doubled every generation (starting at 8 trays for the
661 F2). By doubling food every generation we generated sustained population expansion, thereby
662 continually selecting for the fecundity-associated traits that are advantageous in wild
663 populations when resources are abundant, most likely reproductive output and increased
664 developmental rate. Hereafter, we refer to collection timepoints in accordance with the number
665 of generations since release into large cages (the first 8 tray stage), which was two generations
666 from the actual founding of the experimental populations (i.e., the samples labeled "generation
667 0" are F2's, "generation 1" are F3's, and so forth). Once flies had laid eggs for a period of 24h,
668 the adults were removed from each cage and their volume measured to estimate census size.
669 The remaining embryos (i.e., subsequent generation of flies) were left to develop and eclose
670 within the same replicate cage. Across replicates, the population sizes rapidly expanded across
671 generations, a dynamic selecting for individuals exhibiting increased reproduction (e.g.,
672 fecundity and developmental rate) (Fig. 1).

673 We maintained food doubling until generation 4 at which point, due to logistical
674 constraints, it was not possible to continue the doubling of population size per cage beyond the
675 64 tray per cage generation. In order to continue the selection regime of population expansion
676 (and selection for early fecundity, fast developmental rate), we instituted a random dilution
677 followed by re-expansion. Specifically, the fifth generation was founded with embryos collected
678 over 24h on two food trays (containing eggs of approximately 20,000 flies), from the from the
679 fourth generation. These generation 5 flies were allowed to develop and eclose in the large
680 indoor cages, at which point 8 trays of medium were added to each replicate cage, and
681 oviposition was carried out for 24h and after which adults were discarded. Once the sixth
682 generation of flies eclosed, they were allowed to oviposit for 24h on 2 food trays. These
683 generation 7 embryos were then allowed to develop, eclose, and expanded out to 8 trays for the
684 eighth generation. The generation 8 flies were allowed to oviposit over 24h on 8 trays,
685 consistent with the previous generation (thus each generation of flies resulted from egg laying
686 over a period of 24h from the preceding generation). We collected a random sample of 100 male
687 and 100 female flies in generations 0, 1, 2, 3, 4, 6, 8, and 9, and pooled sexes separately for later
688 whole genome shotgun sequencing (see below). We took additional samples of each replicate at
689 generations 6 and 8 to use as biological replicates (i.e., same replicate/generation, different set
690 of 100 flies) to quantify noise in our allele frequency estimates (see below). Census estimates
691 (based on the total volume of dead flies) of adult flies within each cage were conducted on
692 generations 1-4. Calibration of census estimates was conducted by counting the number of flies
693 (desiccated and dried) in $1cm^3$ and then measuring the volume in a given sample.

694 Concurrent to the generation 4 dilution and re-expansion described above, we induced a
695 bout of truncation selection that segregated flies based on stress-tolerance, most likely
696 desiccation resistance as water availability is expected to drive mortality at a much faster rate
697 than starvation (33). Specifically, after collecting generation 5 embryos for continued expansion,
698 we retained the adult flies within their respective cages, removed all food and water, and
699 sampled the surviving adults at various timepoints by direct aspiration using vacuums until all
700 flies had died. The live samples were immediately sorted by sex and timepoint and preserved in
701 ethanol (as with expansion samples, 100 male and 100 female flies were isolated at each
702 collection time point for whole genome shotgun sequencing). We continued this process until
703 there were an insufficient number of flies (i.e., < 100 individuals) for sampling, resulting in
704 samples for DNA analysis collected at hours 0, 6, 9, 12, 15, 19, and 21. Samples collected
705 throughout this process were expected to reveal changes in the relative frequency of genotypes
706 of differing degrees of stress-tolerance, traits expected to trade-off with the reproductive traits
707 favored during the nine generations of expansion selection (10,29).
708 *Pooled genomic sequencing and allele frequency estimation of expansion/truncation selection experiment*
709 Genomic DNA from pools of 100 male and 100 female flies were extracted and
710 sequenced in two rounds. First, multiplexed libraries were created for 56 expansion samples
711 using Illumina Nextera DNA Prep with e-gel size selection and i7 indexing. Barcoded fragments
712 were mixed and loaded evenly into 4 lanes, then sequenced with 100 bp paired-end reads on an
713 Illumina HiSeq2000 sequencer, with target coverage of 10x per sample. After QC, resulting
714 per/sample coverage was quite variable, and 26 samples from this round with coverage <2x
715 were later re-sequenced with 150-bp dual-indexed reads on an Illumina HiSeq4000 sequencer.
716 Reads from the same sample were merged after all QC and mapping steps. A separate round of
717 sequencing was conducted for the remaining 96 samples, including remaining expansion and
718 truncation samples. These samples were sequenced with 150-bp paired-end dual-index reads on
719 a NextSeq550 high output machine, with a target coverage 7x per sample. All samples were
720 demultiplexed, adapter sequences were trimmed, and reads with any 3′ bases with quality score
721 < 20 or trim length <18 were discarded. Overlapping forward and reverse reads were
722 subsequently assembled and reads were mapped separately to the *D.mel* v5.39 reference
723 genome using bwa and default parameters (58). Aligned reads were deduplicated using Picard
724 tools (http://broadinstitute.github.io/picard/), and all reads were re-aligned around indels
725 using GATK v4 IndelRealigner (https://gatk.broadinstitute.org/).
726 We used the founder line genome sequencing data to compute haplotype-informed
727 allele frequency estimates at 2.7 M, previously identified, segregating sites using a local
728 inference method and pipeline developed and described by (59,60). We conducted haplotype
729 inference in window sizes that varied proportionally to the length of un-recombined haplotype

730 blocks expected as a function of the estimated number of generations since the construction of
731 the outbred population. We have previously provided extensive validation that our haplotype-
732 informed allele frequencies are replicable across different sets of subsampled reads from the
733 same sample, and produce an accuracy of allele frequency estimates that is comparable to deep
734 sequencing and standard methods for pooled, allele frequency estimation (14,16,60). We filtered
735 our allele frequencies to only include those sites with an average minor allele > 0.02 across all
736 samples, resulting in 1.7M SNPs for analysis. Finally, we averaged technical replicate allele
737 frequencies (i.e. the male and female pool) for each replicate and collection time point, resulting
738 in 4 cages per time point, across 8 expansion generations and 7 truncation time points (60 total
739 samples).

740 *Identifying parallel patterns of allele frequency change and genome-wide evidence of trade-offs*

741 Statistical analysis of allele frequency data was conducted using R v. 3.5.6. We first
742 explored patterns of genomic divergence between samples as average F_{ST} across all segregating
743 sites, which provides a metric of allele frequency divergence between samples that is
744 normalized by the starting frequency of the allele. We compared F_{ST} between biological
745 replicates (same replicate cage/collection time point, different pool/extraction) to those values
746 obtained for each replicate cage between its generation 0 sample and all subsequent generation
747 expansion samples. We quantified whether evolutionary divergence increased throughout the
748 course of expansion, and at what point it exceeded the sources of biological noise impacting our
749 allele frequency estimates, using a linear regression.

750 We next explored whether increasing differentiation through time was underpinned by
751 parallel allele frequency shifts across replicates. Specifically, we used our pairwise divergence
752 values (i.e., F_{ST} between samples) to create multi-dimensional scaling (MDS) plots, in which
753 divergence between samples is represented as distance between points in a 2-D plane (*cmdscale*
754 function in *stats* package R). If the observed evolutionary change across replicates was
755 dominated by parallel responses to shifting population densities, then the segregation of
756 samples across the 2-D plane should, in part, correspond to the generation from which samples
757 were derived. This analysis was conducted on mean, genome-wide F_{st} values, as well as F_{st}
758 values computed separately for SNPs on each chromosomal arm. To quantify trends visualized
759 using MDS, we translated the coordinates for each cage such that the centroid of all generation 0
760 samples was centered at the origin and then projected all points onto a single axis which was
761 constructed by fitting a simple linear regression model to samples from generations 0-4
762 (regressions were constructed separately for each replicate cage). The resulting axis represents
763 the primary axis of variation in the 2D plane during early expansion for each replicate (Fig. S5).
764 We hypothesized that if sustained, directional selection imposed by population expansion was a
765 primary driver of patterns in genomic variation, the samples from the remaining expansion

766 generations would continue to proceed along the established axis of variation in the same
767 direction. To evaluate the degree of concordance of early and late expansion generation samples
768 we computed Pearson correlations between collection time and position along this axis. The
769 significance of these correlations was determined via permutations of sample collection time
770 point ($N = 100$ permutations). We conducted this analysis across all genome-wide SNPs, as well
771 as separately for SNPs on each chromosomal arm. Finally, to disentangle whether differences in
772 dynamics observed among chromosomal arms reflected true systematic differences in allele
773 frequency behavior, as opposed technical artifacts driven by differences in the number of SNPs
774 available for analysis across arms, we re-conducted this analysis using randomly sampled
775 subsets of SNPs matching that present on each arm.
776 We quantified the extent to which individual SNPs exhibited parallel movement across
777 cages and throughout expansion by fitting a generalized linear model to allele frequencies
778 (formula: allele frequency ~ expansion generation; *glm* function in base R) (34). Allele
779 frequencies were weighted by the total number of chromosomes sequenced (N = 200) and depth
780 per sample (60). This model assessed the significance of the linear relationship between allele
781 frequency and generation of sampling across all cages, using a logistic link function and
782 quasibinomial error model to reduce false positive associations (15,16,34). While an association
783 between allele frequency and sampling timepoint within a single cage may represent either drift
784 or selection, a significant association across all four replicate cages indicates allele frequency
785 trajectories that are both predictable over time and parallel across populations. In this case,
786 selection (or linked selection) is the more parsimonious explanation. P-values were adjusted
787 using the Benjamini-Hochberg false discovery rate (FDR) correction (*p.adjust* package in base R).
788 We considered a SNP significant if it exhibited an FDR < 0.01 and effect size > 2%. We further
789 validated the parallelism inferred by the GLM by implementing a leave-one-out cross
790 validation. Specifically, we iteratively identified sets of parallel SNPs (FDR < 0.01 and effect size
791 > 2%) across expansion samples using a GLM and allele frequency data from three of the four
792 replicate cages. We then quantified the frequency shifts of the favored allele at significant sites
793 in the left-out cage between generation 0 and 9 of expansion. For sets of parallel SNPs identified
794 via each iteration of the leave-one-out validation, we matched each parallel SNP to a control
795 SNP based on chromosomal arm, starting frequency (within 5% of generation 0 frequency),
796 inversion status, and recombination rate. We compared the distribution of allele frequency
797 shifts at parallel vs. matched control sites (using a two-tailed, paired t-test) for each left out cage
798 to infer if the magnitude of allele frequency change exceeded background allele frequency
799 movement, and whether the dominant direction of allele frequency change was in a direction
800 concordant with the three training cages. We conducted this analysis for SNPs identified
801 genome-wide, as well as separately for each chromosomal arm.

802 Finally, we leveraged the genomic distribution of GLM significant SNPs to disentangle
803 the minimum number of independently selected loci from the coordinated signal across SNPs
804 expected as a result of linkage and genetic draft. Specifically, we implemented a method
805 developed and described by (16). In brief, this method computes scores of GLM significance in
806 sliding windows of 500 SNPs, using a 100 SNP step size. An empirical FDR for each window is
807 generated by comparing observed scores to those generated for permuted windows. Windows
808 with an empirical FDR < 0.05 are defined as significantly enriched and overlapping enriched
809 windows are merged into a single, enriched window. A set of independently segregating loci is
810 the determined by computing average SNP-pair linkage (using the squared correlation of
811 founder genotypes) between all non-overlapping windows within 1 Mb. Windows with average
812 SNP-pair linkage is > 0.03 are merged (this value was empirically determined via chromosome-
813 wide patterns of linkage between SNPs (14, 16)), and this process continues iteratively until no
814 windows with high linkage remain. We have previously shown that this method of identifying
815 independent loci based on patterns of physical linkage indeed produces sets of loci with
816 independent allele frequency movement throughout evolution (14). Source code for this
817 analysis is provided at https://github.com/greensii/dros-adaptive-tracking). We used the
818 most significant SNP within each unlinked locus to compute selection acting across
819 reproduction selection. Specifically, we used the observed allele frequency change to quantify
820 selection coefficients as: $s = \Delta p/(p \times (1-p))$ where p corresponds to the allele frequency of the
821 focal marker SNP. We note this derivation assumes no impact of genetic drift on allele
822 frequencies and a constant selection coefficient throughout expansion, reasonable assumptions
823 in the context of the large populations and short timescale over which we conducted the
824 experiment and the controlled, isolated selection on reproductive output throughout this
825 period.
826
827 *Quantifying evidence of fecundity and desiccation tolerance trade-offs*
828 We next used samples collected throughout truncation, during which patterns of
829 survival and selection were most likely underpinned by differences in desiccation resistance
830 among genotypes, to quantify the presence of fitness-tradeoffs at those alleles systematically
831 favored during expansion. First, we re-conducted our MDS analysis of pairwise divergence
832 values, this time including all samples collected across both phases of the experiment. While
833 general qualitative inference may be obtained from the visualization of this analysis, we
834 explicitly quantified the relative movement of expansion and truncation samples across this 2D
835 plane to provide a more rigorous investigation into the relative direction of allele frequency
836 movement across phases. Specifically, as above, we translated MDS coordinates for each cage
837 such that the centroid of all hour 0 truncation samples was centered at the origin, and then

838 projected onto a single axis which, as above, was constructed by fitting a simple linear
839 regression model to samples from expansion generations 0-4. We hypothesized that if our
840 contrasting reproduction/stress-tolerance selection environments were inducing genome-wide
841 reversions in allele frequency movement, then the projection of samples collected throughout
842 truncation would regress in the opposite direction as that observed for the expansion samples.
843 The degree of concordance/discordance for truncation samples were quantified as the Pearson
844 Correlation between collection time and position along this axis, and the significance of these
845 correlations was determined via permutations of sample collection time point ($N = 100$
846 permutations). This analysis was also conducted both genome-wide and separately for each
847 chromosomal arm and, as above, we used genome-wide subsampling to infer whether variation
848 in dynamics among chromosomal arms was a technical artifact.

849 Finally, we aimed to corroborate patterns observed in the dimensionality reduction of
850 genome and chromosome-wide F_{ST} values described above by accruing additional evidence of
851 genome-wide trade-offs using allele frequency trajectories. We first isolated those SNPs with
852 strong evidence of linkage to a selected locus throughout expansion (GLM FDR < 0.01 and allele
853 frequency shift > 2%). We then quantified the behavior of the rising allele at each SNP
854 throughout the progression of truncation, and evaluated whether these alleles were more likely
855 to exhibit a reversion in trajectory direction, or continue in the same direction, using a Chi-
856 squared test and set of control SNPs (matched on chromosome, starting frequency,
857 recombination rate, and inversion status). We further assessed if and how the results of this
858 analysis varied when only evaluating trajectories of expansion-identified SNPs that also
859 exhibited independent evidence of linkage to a selected locus (GLM FDR < 0.05 and allele
860 frequency shift > 1%) during truncation. Finally, we further quantified the distribution of
861 expansion-identified SNPs, genome-wide and per chromosomal arm, using empirical
862 cumulative distribution functions (eCDF). The eCDFs for SNPs with sustained directional
863 movement across treatments and evidence of fluctuating selection were evaluated
864 independently via comparison to eCDFs generated via a set of matched control SNPs, and using
865 a Kolmorgorov-Smirnov test (*stats* package in R). We then used the set of 250 unlinked loci
866 identified throughout to further quantify the signal of fluctuating selection in our data.
867 Specifically, for each locus, we computed the mean frequency shift of expansion favored alleles
868 (GLM FDR < 0.05 effect size > 2%) during truncation. We then generated a set of allele
869 frequency shifts using matched control SNPs for each locus (matched on starting frequency,
870 chromosome, inversion status, and recombination rate) and a paired, two-sided t-test to infer
871 whether the locus continued in a consistent direction during truncation (positive median shift;

872 t-test FDR < 0.05), was neutral, or moved significantly in the opposite direction (negative
873 median shift; t-test FDR < 0.05).

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

875 From the known cosmopolitan inversions found in the DGRP lines (62), we analyzed
876 those that occurred in at least 10 of our founding strains (i.e., > 4%). We used chromosome-level
877 haplotype frequencies to compute the frequencies of each of these inversions independently
878 during expansion and truncation. Next, we regressed inversion frequencies through time using
879 a generalized linear model (logistic link function, and quasibinomial error variance) to classify
880 inversions moving systematically across replicates throughout each phase (Benjamini-Hochberg
881 corrected P-value < 0.1). For those inversions displaying significant frequency change across
882 both expansion and truncation, we characterized those that exhibited thus fluctuating selection
883 and trade-offs (i.e. switched directions) across phases.

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

892

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

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

906 The replicate cages for both the outdoor mesocosms and indoor cage study were seeded
907 with a genetically diverse, outbred population, derived from a panel of 76 inbred strains
908 originally collected wild from Linvilla Orchards, Media, PA (Supplementary Data File 2).
909 Constat food was supplied to the replicates within each environment, whereby four hundred ml
910 of Drosophila media ('Spradling cornmeal recipe') was provided in 900 cm^3 aluminum loaf pans
911 within each cage three times per week. These pans provided the only source of food and egg
912 laying substrate, and egg laying upon each loaf pan was carried out for two days, after which a
913 new pan was added, and the original pan was covered with a mesh lid to prevent any further
914 laying. The lids were removed after eclosure was first observed, causing each replicate to
915 experience a near continual input of new flies that evolved with overlapping generations. The
916 census size of the adult population in the outdoor mesocosms were estimated five times
917 throughout the progression of the experiment (19 July, 5 August, 20 August, 17 September, and
918 20 October) (14), while the final census in the indoor cages were estimated at the end of the
919 monitoring period via collection and volumetric quantification of all flies remaining in each
920 replicate. The outdoor mesocosms were sampled to quantify patterns of genomic variation
921 weekly for 9 weeks (13 July – 7 September), after which samples were collected on 21
922 September, 20 October, and following the first freeze and population crash on 20 December (12
923 total time points). The indoor cages were sampled at the first, second, eighth, and eleventh time
924 points. The overlapping monitoring period of the outdoor and indoor cages encompassed the
925 rapid population expansion and ultimate stabilization of cage densities as the replicates in each
926 environment presumably reached carrying capacity (Fig S25) (14). The selective landscape
927 experienced during this shared monitoring period probably mirrors that experienced under the
928 sustained population expansion imposed by population expansion portion of the first
929 experiment of this study, described above. The notable differences between these experiments,
930 however, is that the first experiment increased food availability as a function of population
931 density throughout expansion, while our latter experiment maintained a constant food
932 substrate throughout the study period. Thus, this paired, indoor-outdoor study likely also
933 imposed selection on traits impacted by density regulation (e.g., larval competition). Analogous
934 to the truncation portion of the first experiment, the final collection interval of the outdoor
935 mesocosm study (time point 11->12) encompassed the onset of winter, during which time a
936 total crash of the adult populations in the cages was observed. Selection during this phase of the
937 experiment was thus hypothesized to favor increased in stress tolerance traits associated with
938 the deteriorating abiotic conditions.
939 Methods for sampling flies for sequencing from the outdoor mesocosms and indoor
940 cages were similar to those described above: eggs were collected overnight directly from each
941 replicate cage, and larval development was carried out in the indoor environment where eggs

942 developed and eclosed to F1 adults in 30 cm^3 cages. A random set of 100 females were sampled
943 from each cage 3-5 days post-eclosure and preserved in 99% ethanol at -20° C (14). Genomic
944 DNA was extracted from pools of flies for each replicate and time point using the Monarch
945 Genomic DNA Purification Kit (New England Biolabs). Libraries were nonstructured using the
946 Illumina DNA Prep Tagmentation Kit and all samples were sequenced on Illumina Novaseq
947 6000 flow cells using 150 bp, paired end reads. Raw sequencing reads from each sample were
948 trimmed of adapter sequences and bases with quality score < 20, and aligned to the *Drosophila*
949 *melanogaster* v5.39 reference genome using bwa and default parameters(58). Aligned reads were
950 deduplicated using Picard tools (http://broadinstitute.github.io/picard/) and the final set of
951 reads for each sample was down-sampled to obtain an equivalent, genome-wide coverage of 8x
952 across all samples (>100x effective coverage) (14,60). Haplotype informed allele frequencies
953 from reads from each replicate and time point were then generated using the local inference
954 pipeline described above (see *Allele Frequency Calculation*) (59,60). The final set of allele
955 frequencies only included those sites with an average minor allele frequency > 0.02 in the
956 baseline population, and present in at least one evolved sample at a MAF > 0.01, ultimately
957 yielding 1.9 M SNPs (14).
958 We quantified the extent to which evolution in the outdoor environment is driven by the
959 selective pressures solely associated with increasing population density (e.g., selection for
960 increased fecundity, faster developmental rate, increased competitive ability). Specifically, we
961 identified alleles with systematic behavior independently in the indoor and outdoor cages using
962 GLM (FDR < 0.01; allele frequency change > 2%), and compared the observed overlap between
963 these lists using a Chi squared test and with null expectations generated via a set of matched
964 control SNP set. Next, we queried whether the dominant direction of selection throughout the
965 shared monitoring period of the indoor and outdoor environment was consistent. Specifically,
966 we quantified the mean frequency shift of the rising allele at SNPs identified in the indoor
967 environment via GLM, in the outdoor cages during the same time interval. We determined
968 whether the magnitude of frequency shifts exceeded background allele frequency movement
969 expected via matched control SNPs, and whether the dominant direction was parallel or anti-
970 parallel given the sign of observed median shift (analysis conducted both genome-wide and
971 separately for each chromosomal arm using a paired t-test and sign test). Finally, to test
972 whether the total population crash in the outdoor mesocosms elicited evidence of the trade-offs,
973 we quantified the behavior of the indoor identified alleles throughout this period. Through this,
974 we used a common set of alleles, determined as linked to an advantageous locus during
975 increasing population density in a controlled, lab-based setting, to infer the presence of trade-
976 offs during a bout of population expansion and collapse in a semi-natural setting.
977

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.

Financial Disclosures: 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. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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 [cited 2023 Jul 24]. (Cambridge Studies in Mathematical Biology). Available from: https://www.cambridge.org/core/books/evolution-in-agestructured-populations/48C9E8DE72A6FAD03938E586641BA620

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. Nordborg M, Wittkopp PJ, Nordborg M, editors. 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. p. 2024.06.14.599085. 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. Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM, et al. Life history trade-offs at a single locus maintain sexually selected genetic variation. Nature. 2013 Oct;502(7469):93–5.

21. Mérot C, Llaurens V, Normandeau E, Bernatchez L, Wellenreuther M. Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. Nat Commun. 2020 Feb 3;11(1):670.

1046 22. Cole JM, Scott CB, Johnson MM, Golightly PR, Carlson J, Ming MJ, et al. The battle of the
1047 sexes in humans is highly polygenic. Proceedings of the National Academy of Sciences.
1048 2024 Sep 24;121(39):e2412315121.

1049 23. Ives PT. THE GENETIC STRUCTURE OF AMERICAN POPULATIONS OF
1050 DROSOPHILA MELANOGASTER. Genetics. 1945 Mar 15;30(2):167–96.

1051 24. Band HT, Ives PT. Correlated changes in environment and lethal frequency in a natural
1052 population of drosophila melanogaster. Proceedings of the National Academy of Sciences.
1053 1961 Feb;47(2):180–5.

1054 25. Gleason JM, Roy PR, Everman ER, Gleason TC, Morgan TJ. Phenology of Drosophila
1055 species across a temperate growing season and implications for behavior. PLOS ONE. 2019
1056 May 16;14(5):e0216601.

1057 26. Boulétreau-Merle J, Fouillet P, Terrier O. Seasonal variations and balanced polymorphisms
1058 in the reproductive potential of temperate D. Melanogaster populations. Entomologia
1059 Experimentalis et Applicata. 1987;43(1):39–48.

1060 27. Behrman EL, Watson SS, O'Brien KR, Heschel MS, Schmidt PS. Seasonal variation in life
1061 history traits in two Drosophila species. Journal of Evolutionary Biology. 2015;28(9):1691–
1062 704.

1063 28. Behrman EL, Schmidt P. How predictable is rapid evolution?. bioRxiv; 2022. p.
1064 2022.10.27.514123. Available from:
1065 https://www.biorxiv.org/content/10.1101/2022.10.27.514123v1

1066 29. Hoffmann AA, Parsons PA. An integrated approach to environmental stress tolerance and
1067 life-history variation: desiccation tolerance in Drosophila. Biological Journal of the Linnean
1068 Society. 1989;37(1–2):117–36.

1069 30. Service PM. Physiological Mechanisms of Increased Stress Resistance in Drosophila
1070 melanogaster Selected for Postponed Senescence. Physiological Zoology. 1987
1071 May;60(3):321–6.

1072 31. Zera AJ, Harshman LG. The Physiology of Life History Trade-Offs in Animals. Annual
1073 Review of Ecology, Evolution, and Systematics. 2001 Nov 1;32(Volume 32, 2001):95–126.

1074 32. Del Monte-Luna P, Brook BW, Zetina-Rejón MJ, Cruz-Escalona VH. The carrying capacity
1075 of ecosystems. Global Ecology and Biogeography. 2004;13(6):485–95.

1076 33. Da Lage JL, Capy P, David JR. Starvation and desiccation tolerance in *Drosophila*
1077 *melanogaster* adults: Effects of environmental temperature. Journal of Insect Physiology.
1078 1989 Jan 1;35(6):453–7.

1079 34. Wiberg RAW, Gaggiotti OE, Morrissey MB, Ritchie MG. Identifying consistent allele
1080 frequency differences in studies of stratified populations. Methods in Ecology and Evolution.
1081 2017;8(12):1899–909.

35. Haller BC, Messer PW. SLiM 4: Multispecies Eco-Evolutionary Modeling. The American Naturalist. 2023 May;201(5):E127–39.

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

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

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

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

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

41. Sturtevant AH, Beadle GW. THE RELATIONS OF INVERSIONS IN THE X CHROMOSOME OF DROSOPHILA MELANOGASTER TO CROSSING OVER AND DISJUNCTION. Genetics. 1936 Sep 1;21(5):554–604.

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

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

44. 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.

45. 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.

46. 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.

47. 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.

1115 48. Campbell-Staton SC, Arnold BJ, Gonçalves D, Granli P, Poole J, Long RA, et al. Ivory
1116 poaching and the rapid evolution of tusklessness in African elephants. Science. 2021 Oct
1117 22;374(6566):483–7.

1118 49. Campbell-Staton SC, Cheviron ZA, Rochette N, Catchen J, Losos JB, Edwards SV. Winter
1119 storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard.
1120 Science. 2017 Aug 4;357(6350):495–8.

1121 50. Donihue CM, Herrel A, Fabre AC, Kamath A, Geneva AJ, Schoener TW, et al. Hurricane-
1122 induced selection on the morphology of an island lizard. Nature. 2018 Aug;560(7716):88–
1123 91.

1124 51. Rodríguez-Trelles F, Tarrío R, Santos M. Genome-wide evolutionary response to a heat
1125 wave in Drosophila. Biol Lett. 2013 Aug 23;9(4):20130228.

1126 52. Boor GKH, Schultz CB, Crone EE, Morris WF. Mechanism matters: the cause of
1127 fluctuations in boom—bust populations governs optimal habitat restoration strategy.
1128 Ecological Applications. 2018;28(2):356–72.

1129 53. Carroll SP, Hendry AP, Reznick DN, Fox CW. Evolution on ecological time-scales.
1130 Functional Ecology. 2007;21(3):387–93.

1131 54. Hendry AP. Eco-evolutionary Dynamics [Internet]. Princeton University Press; 2016.

1132 55. Kendall BE, Briggs CJ, Murdoch WW, Turchin P, Ellner SP, McCauley E, et al. Why Do
1133 Populations Cycle? A Synthesis of Statistical and Mechanistic Modeling Approaches.
1134 Ecology. 1999;80(6):1789–805.

1135 56. de Bakker PIW, Burtt NP, Graham RR, Guiducci C, Yelensky R, Drake JA, et al.
1136 Transferability of tag SNPs in genetic association studies in multiple populations. Nat Genet.
1137 2006 Nov;38(11):1298–303.

1138 57. Martinez KL, Klein A, Martin JR, Sampson CU, Giles JB, Beck ML, et al. Disparities in
1139 ABO blood type determination across diverse ancestries: a systematic review and validation
1140 in the All of Us Research Program. Journal of the American Medical Informatics
1141 Association. 2024 Jun 25;ocae161.

1142 58. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform.
1143 Bioinformatics. 2009 Jul 15;25(14):1754–60.

1144 59. Kessner D, Turner TL, Novembre J. Maximum Likelihood Estimation of Frequencies of
1145 Known Haplotypes from Pooled Sequence Data. Mol Biol Evol. 2013 May 1;30(5):1145–58.

1146 60. Tilk S, Bergland A, Goodman A, Schmidt P, Petrov D, Greenblum S. Accurate Allele
1147 Frequencies from Ultra-low Coverage Pool-Seq Samples in Evolve-and-Resequence
1148 Experiments. G3 (Bethesda). 2019 Oct 21;9(12):4159–68.

61. 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.

Supplementary Information for Bitter *et al.* (2024)

Supplementary Methods

To evaluate whether patterns of linkage between loci, antagonistic pleiotropy at a single SNP, and/or haplotype-level antagonistic pleiotropy could generate the patterns of allele frequency change observed in our data we ran Wright-Fisher simulations with SLiM (1). First, we simulated the behavior of a single SNP with antagonistically pleiotropic behavior evolving under fluctuating selection. Specifically, in a population of size N=1000 diploid individuals, we introduced a target SNP 'A' at a range of initial, intermediate starting frequencies ($p_0 \in$ 33-64%). We traced the frequency of the favored allele over 13 generations: the first 9 generations corresponded to the expansion phase of our experimental evolution, and the fitness of A during expansion was labeled $s_{expansion}$. The last 4 generations corresponded to the truncation phase, with the fitness of 'A' labeled $s_{truncation}$. In our simulations, we considered every possible combination of fitness values ($s_{expansion}$, $s_{truncation}$), where $s_{expansion}$ was varied between 1.0 (neutral) and 1.2 (highly advantageous), and $s_{truncation}$ was varied between 1.0 (neutral) and 0.8 (highly deleterious). This range of selection coefficients was based on the empirical selection coefficient values we computed from the allele frequency data in our reproduction/stress-tolerance selection experiment (**Fig. S8**). In all simulations, here and below, we assumed no dominance (h=0.5). In total, we ran 250 replicate simulations for each combination of two selection parameters ($s_{expansion}$, $s_{truncation}$).

We then ran a set of simulations in which two separate SNPs, 'A' and 'B', are under selection during expansion or truncation, respectively. The two SNPs are located a set distance, *d*, nucleotides apart. We seeded the genotypes of individuals in the initial generation by fixing initial frequencies of alleles, $p_0^{(A)}$ and $p_0^{(B)}$, as well as their linkage state. The linkage was determined by the magnitude of the LD measure r^2 between A and B, with an $r^2 <$ 0.01 corresponding to an effectively "unlinked" state between A and B, and r^2>0.05 corresponding to A and B being "highly linked". We further distinguished between the two different ways in which A and B could be highly linked by using the sign of r: the "raw" correlation between 'A' and 'B' that, when squared, yields r^2. When $r > 0$ and $r^2 > 0.05$, we consider A and B to be both highly linked and in "attraction", meaning that favored alleles (those that are assigned fitness effects in the simulations) at both loci are assorted together. Conversely, when $r < 0$ and $r^2 > 0.05$, 'A' and 'B' are highly linked and in a state of "repulsion", whereby the favored allele of 'A' is assorted together with the disadvantageous allele of 'B', and vice versa. To summarize, we have distinguished three linkage states: "unlinked", "highly linked in attraction", and "highly linked in repulsion".

As the range of possible values of *r* and r^2 depends on the initial frequencies of A and B (2), for each simulation we sampled the value of *r* uniformly from its range of possible values.

For example, when simulating loci in attraction, we chose r uniformly in the interval ($\sqrt{0.05}$, r_{max} ($p_0^{(A)}$, $p_0^{(B)}$)). When simulating loci in repulsion, we chose r uniformly in the interval (r_{min} ($p_0^{(A)}$, $p_0^{(B)}$), $-\sqrt{0.05}$). Finally, for unlinked loci, we chose r randomly in the interval ($-\sqrt{0.05}$, $\sqrt{0.05}$). The functions r_{min} and r_{max} were derived following the approach of VanLiere and Rosenberg (2008). After r was chosen, we computed the numbers of individuals possessing haplotypes AB as:

$$p_0^{(AB)} = p_0^{(A)} p_0^{(B)} + r\sqrt{p_0^{(A)} p_0^{(B)} (1 - p_0^{(A)})(1 - p_0^{(B)})}$$

and used this value in setting the initial generation of the simulation.

After seeding the simulation by fixing initial frequencies of A and B and their linkage state in the first generation, we ran the simulation for 13 generations as above, with the first 9 generations corresponding to "expansion" and then 4 generations of "truncation". In all simulations we set $s_{expansion}^{(B)} = s_{truncation}^{(A)} = 0$. The recombination rate ϱ was equal to ϱ=2.39×10−8 per generation per base pair during expansion, and to ϱ=0 during truncation. The initial frequencies, $p_0^{(A)}$ and $p_0^{(B)}$, were chosen randomly in each simulation from the interval (0.34, 0.66). For the rest of the parameters, we formed a parameter grid and considered every possible combination of parameter values (linkage state, d, $s_{exp}^{(A)}$, $s_{trunc}^{(B)}$) where d ∈ {0.5, 1, 2, 4, 20} measured in megabases, $s_{exp}^{(A)}$ ∈ {1.0, 1.05, 1.1, 1.15, 1.2}, $s_{trunc}^{(B)}$ ∈ {1.0, 0.95, 0.9, 0.85, 0.8}, and linkage state was one of "unlinked", "highly linked in attraction", and "highly linked in repulsion". We ran 1000 replicate simulations for each combination of parameter values. The seeding procedure described above ensures that we are quantifying average behavior of alleles in each linkage category with varying values of r^2 and initial frequency. To summarize the results of the simulation, we computed maximum-likelihood estimates of "empirically observed" fitness values for A and B directly from the allele frequency changes (eq. (4) in (3)). For the expansion stage, we used all generations, while fitness during truncation was estimated only from the first generation. The estimate of fitness during the first generation of truncation represents the stress tolerance selection period of our experiment and explicitly quantifies the linkage effect of the B SNP on the A SNP's allele frequency trajectory during this period.

Finally, to assess the net effect of linkage on the behavior of the allele frequency of SNP 'A' (that favored during expansion) during the first generation of truncation, we combined replicate simulation runs for all three linkage categories, and computed fitness of SNP A during the first generation of truncation from all combined simulations in which the attraction and repulsion states were equally likely to be observed. The resulting 'net' effect of linkage states on

trajectories observed via this scenario then represents independent assortment of the 'A' and 'B' alleles.

Supplementary Figures

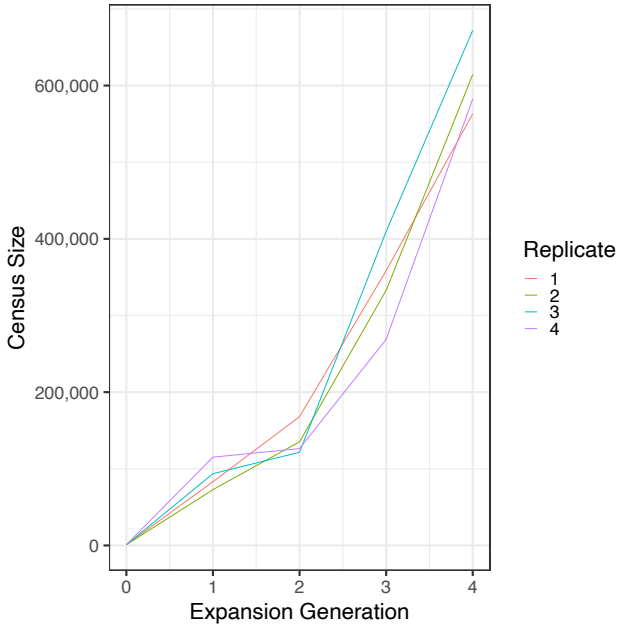

Figure S1. Census estimates of indoor replicate cages during the first four generations of population expansion.

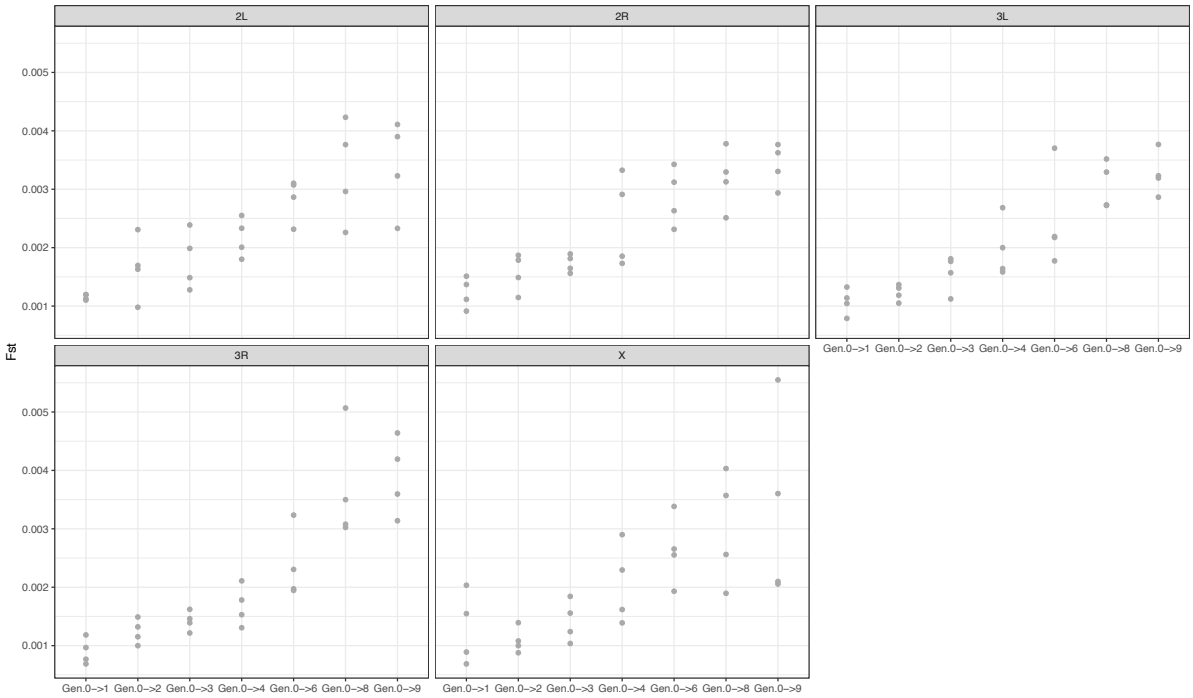

Figure S2. Evolution of allele frequencies across expansion generations | (A) Mean, genome-wide Fst between each evolved replicate and its generation 0 sample (Gen.0->n), segregated by chromosomal arm. Each grey dot corresponds to an individual replicate.

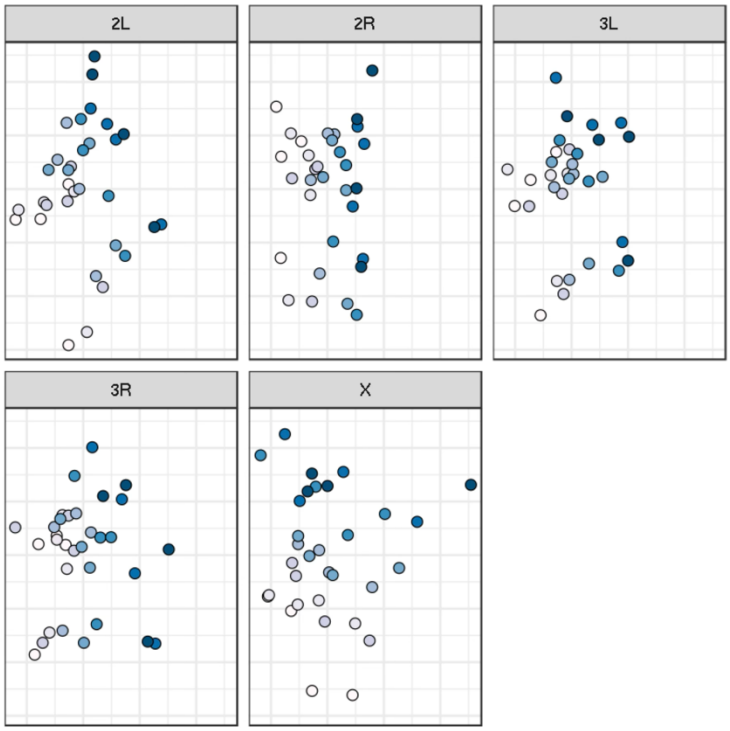

Figure S3. MDS of pairwise Fst values computed per chromosomal arm and across all expansion samples. Point Color corresponds to sample expansion collection generation.

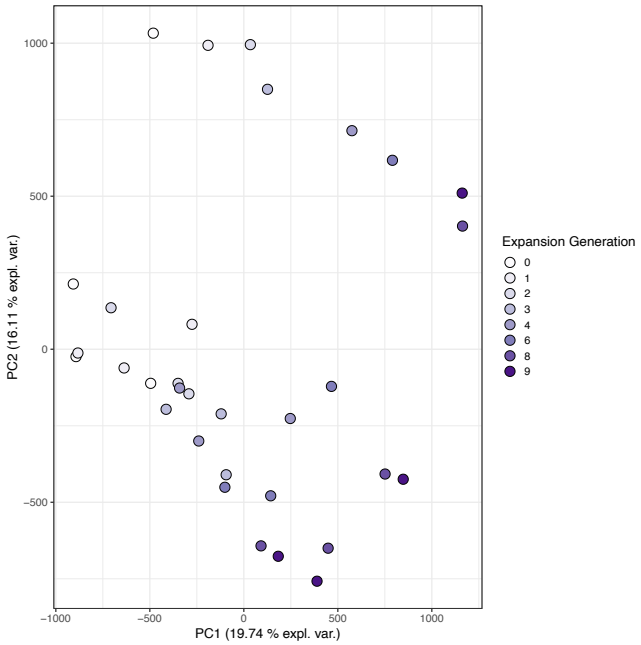

Figure S4. Principal component analysis of samples collected throughout expansion using allele frequency data across all 1.7 M SNPs. Samples are projected onto the first two principal components and colored in accordance with expansion generation collection time point.

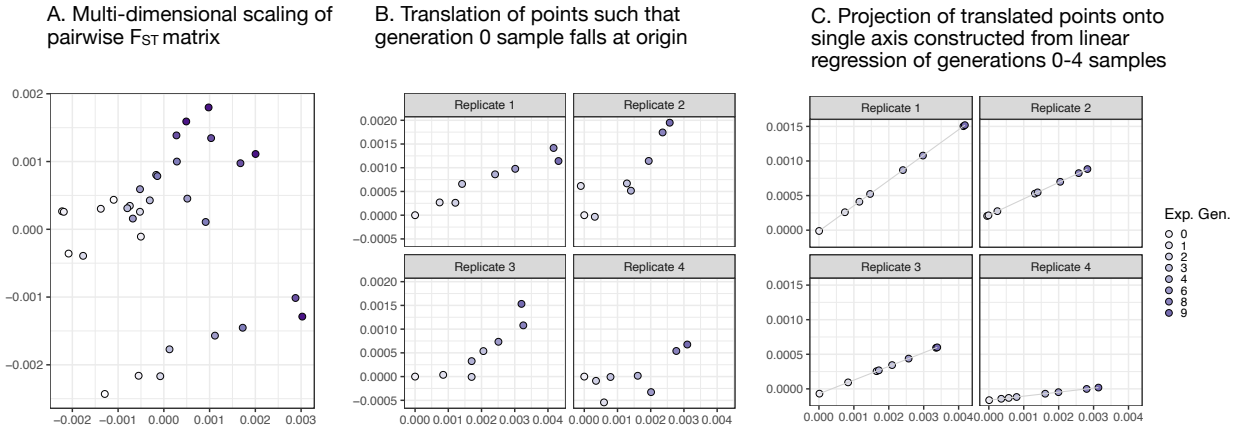

Figure S5. F_{ST}-based MDS analysis workflow. (A) Multi-dimensional scaling of expansion sample pairwise F_{ST} matrix as in main figure 2. (B) Translation of expansion points such that, for each replicate, the generation 0 sample falls at the plot origin. (C) Projection of translated points (B) onto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

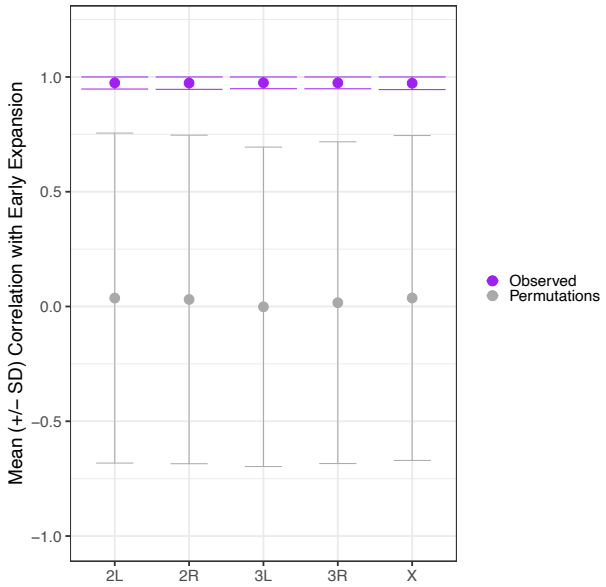

Figure S6. Average Pearson Correlation across replicates (+/- standard deviation) between sample expansion generation and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples. Purple points and error bars correspond to empirical values, 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.

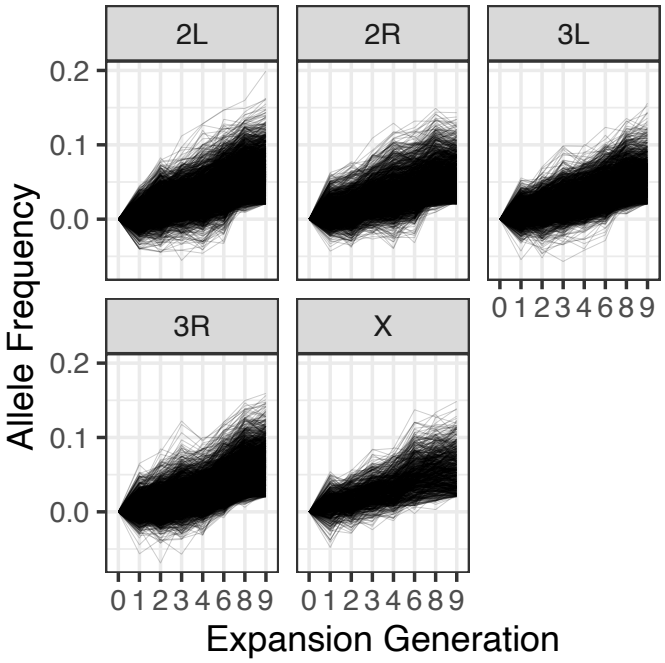

Figure S7. Trajectories of SNPs identified via GLM as exhibiting parallel frequency shifts across replicates throughout expansion (GLM FDR < 0.01 and effect size > 2%), segregated by chromosomal arm.

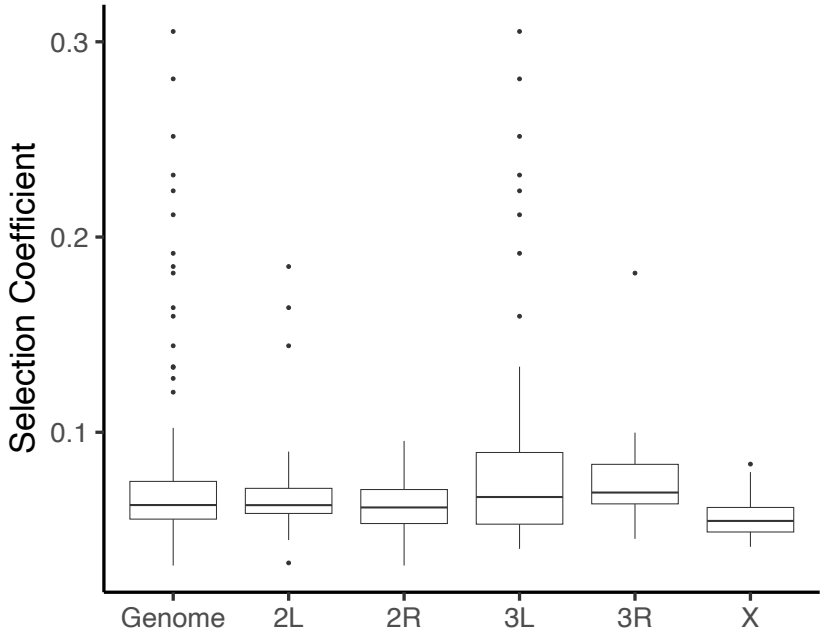

Figure S8. Per-generation selection coefficients computed using top marker SNP within each of the 250 unlinked loci identified throughout reproduction selection. Box plots span the 25th -75th percentile of distribution and horizontal bars indicate the median value (outliers depicted as dots).

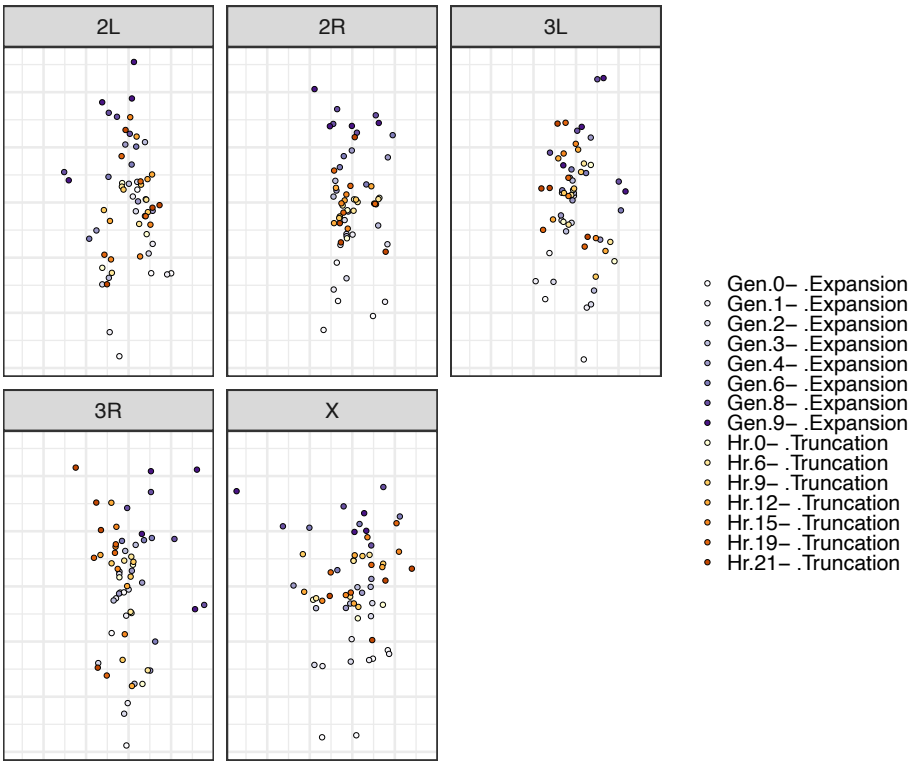

Figure S9. MDS of pairwise Fst values for each chromosomal arm and across all samples collected throughout expansion (purple-hue points) and truncation (orange-hue points). Points are shaded according to collection time point (darker hues indicate later expansion or truncation sampling generation/hour).

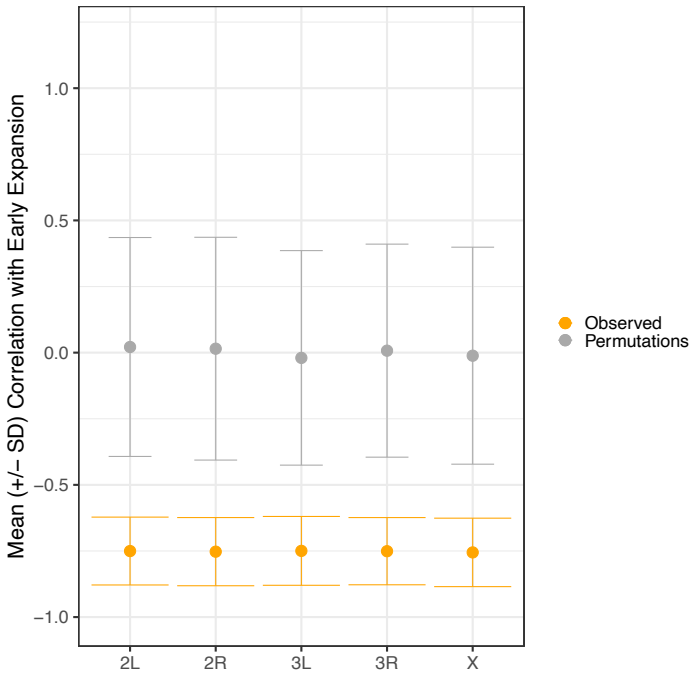

Figure S10. Average Pearson Correlation across replicates (+/- standard deviation) between sample truncation hour and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples. Orange points and error bars correspond to empirical values, 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.

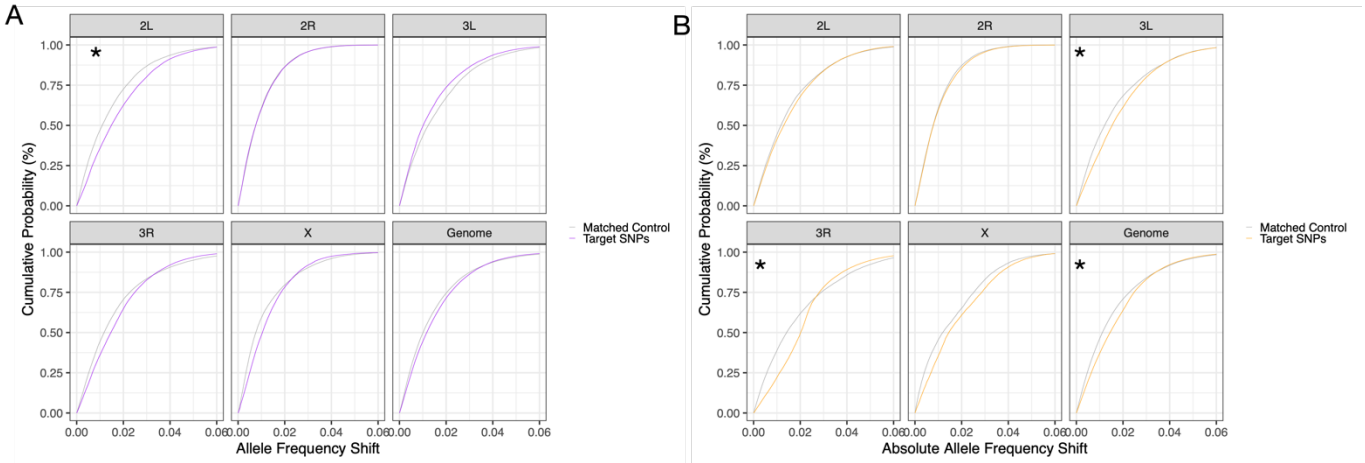

Figure S11. eCDFs of expansion-favored SNPs during truncation, segregated by those with sustained direction of movement (A) or reversals in direction (B). Asterisks correspond to instances in which the

observed eCDF differed significantly from that generated via matched control SNPs (grey lines) (one-tailed Kolmogorov-Smirnov test, p-value < 0.05).

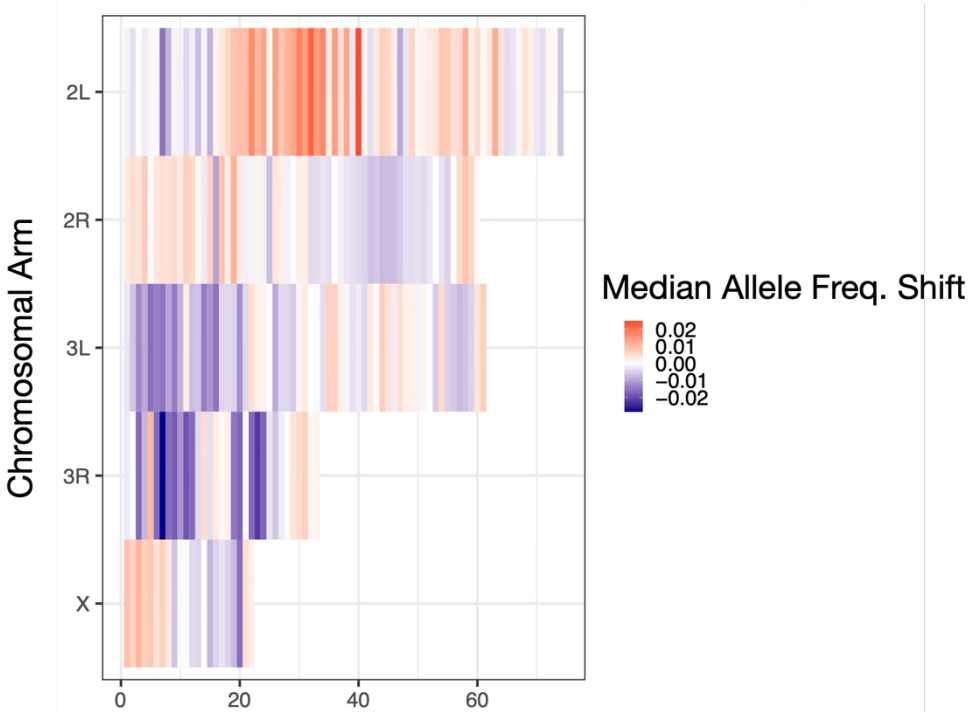

Figure S12. Expansion-favored locus trajectories (N = 250) throughout truncation. Loci are ordered according to chromosomal arm (y-axis) and position along chromosome (x-axis). X-axis ticks refer to the total number of loci identified on each arm. Tile colors correspond to the median allele frequency shift of each locus during survival stress tolerance selection. In total, 90 loci were inferred neutral during stress tolerance selection, while 90 exhibited signals of fluctuating selection (negative median allele frequency shift and paired T-test FDR < 0.05) and 70 under sustained directional selection during truncation (positive median allele frequency shift and paired T-test FDR < 0.05).

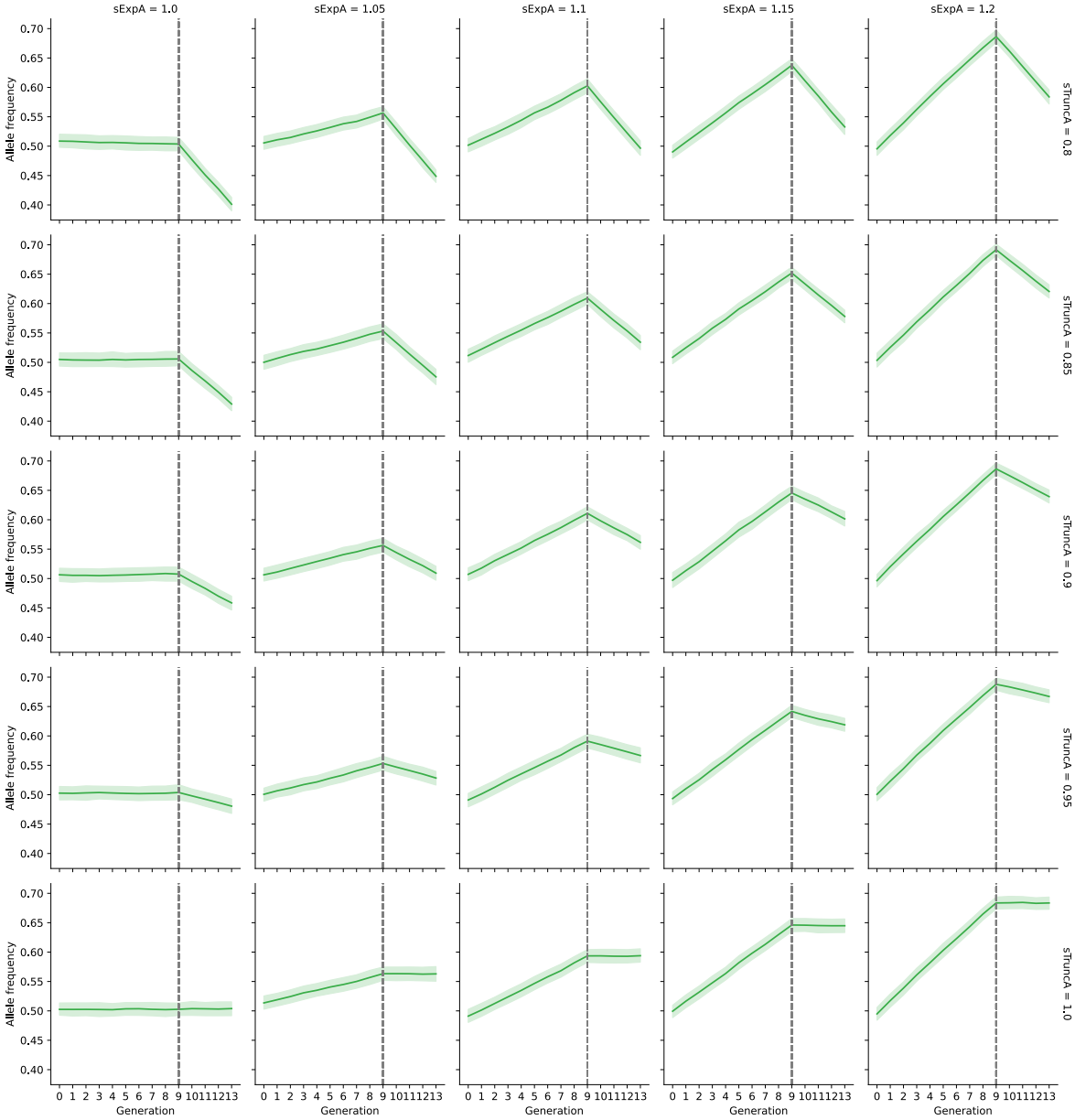

Figure S13. Simulated allele frequency trajectories of a SNP ('A') under a scenario of antagonistic pleiotropy across expansion (generation 0-9) and truncation (10-13). The selection coefficient of 'A' is ranged between 1.0 and 1.2 during expansion (columns) and (0.8 and 1.0) during truncation (rows). The solid line corresponds to the mean value across simulations and shaded area to the 95% confidence intervals (N = 250 total simulations per set of parameter values).

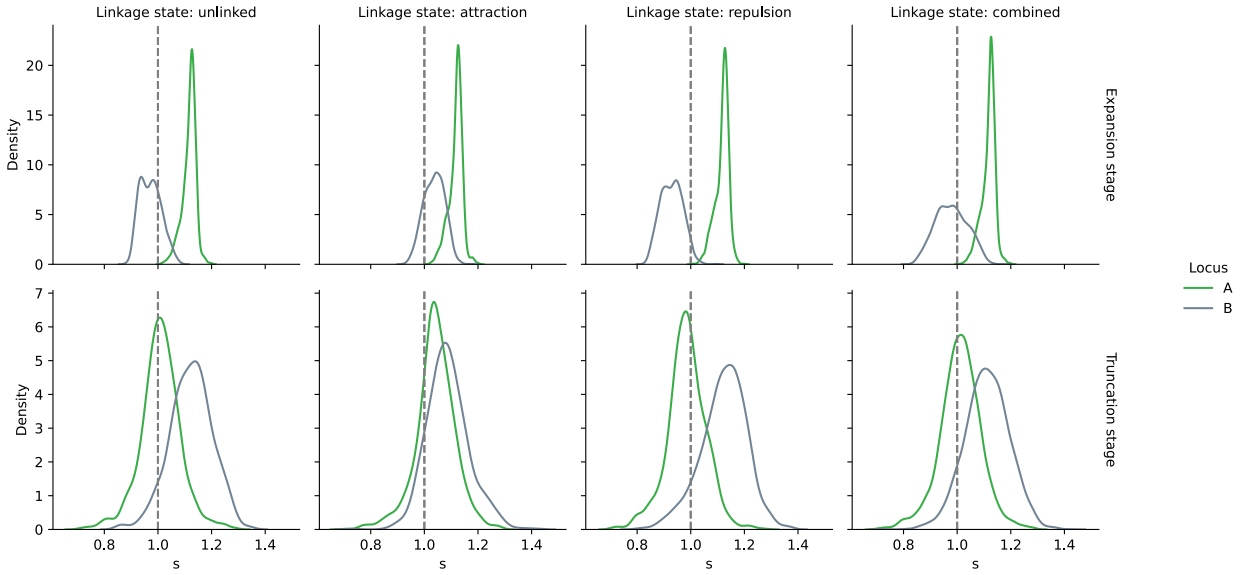

Figure S14. Kernal density estimates of empirical selection coefficients (s) computed during the first generation of truncation selection in two-locus SLiM simulations (N = 1000 simulations per scenario). In each case, the A and B loci are 1 Mb away, with locus A selected during expansion (fitness of favored allele during expansion = 1.1), but neutral during truncation, while B is neutral during expansion and favored throughout truncation (fitness of favored allele during truncation = 1.1). The columns represent different scenarios of linkage between the two loci, whereby 'attraction' represents a scenario in which the favored alleles at each locus are non-independently assorted with each other, and a state of 'repulsion' represents a scenario in which the favored allele at each locus was non-independently assorted with the disadvantageous allele at the other locus. The 'combined' scenario represents the average behavior across linkage scenarios.

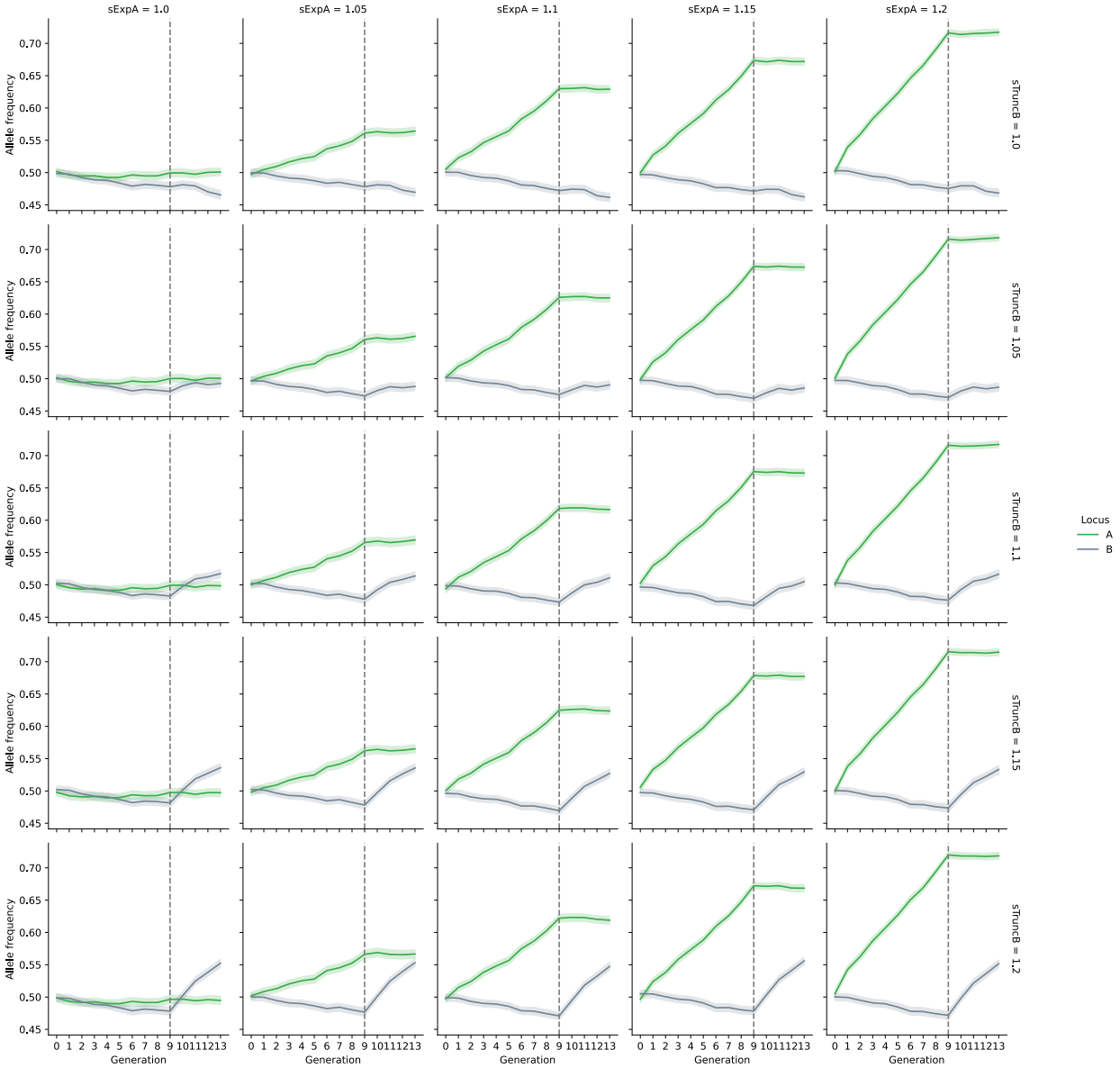

Figure S15. Simulated allele frequency trajectories for two unlinked SNPs under differential selection across expansion and truncation. The 'A' allele is under directional selection during the expansion period (generations 0-9) and neutral during the truncation phase (generations 10-13). The 'B' allele is neutral during expansion, then comes under directional selection during truncation. The selection coefficient of 'A' is ranged between 1.0 and 1.2 during expansion (columns) and remains 1.0 during truncation. The selection coefficient of 'B' is constant at 1.0 during expansion and ranged between 1.0 and 1.2 during truncation (rows). The solid line corresponds to the mean frequency value of each allele across simulations and shaded area to the 95% confidence intervals (N = 250 total simulations per set of parameter values).

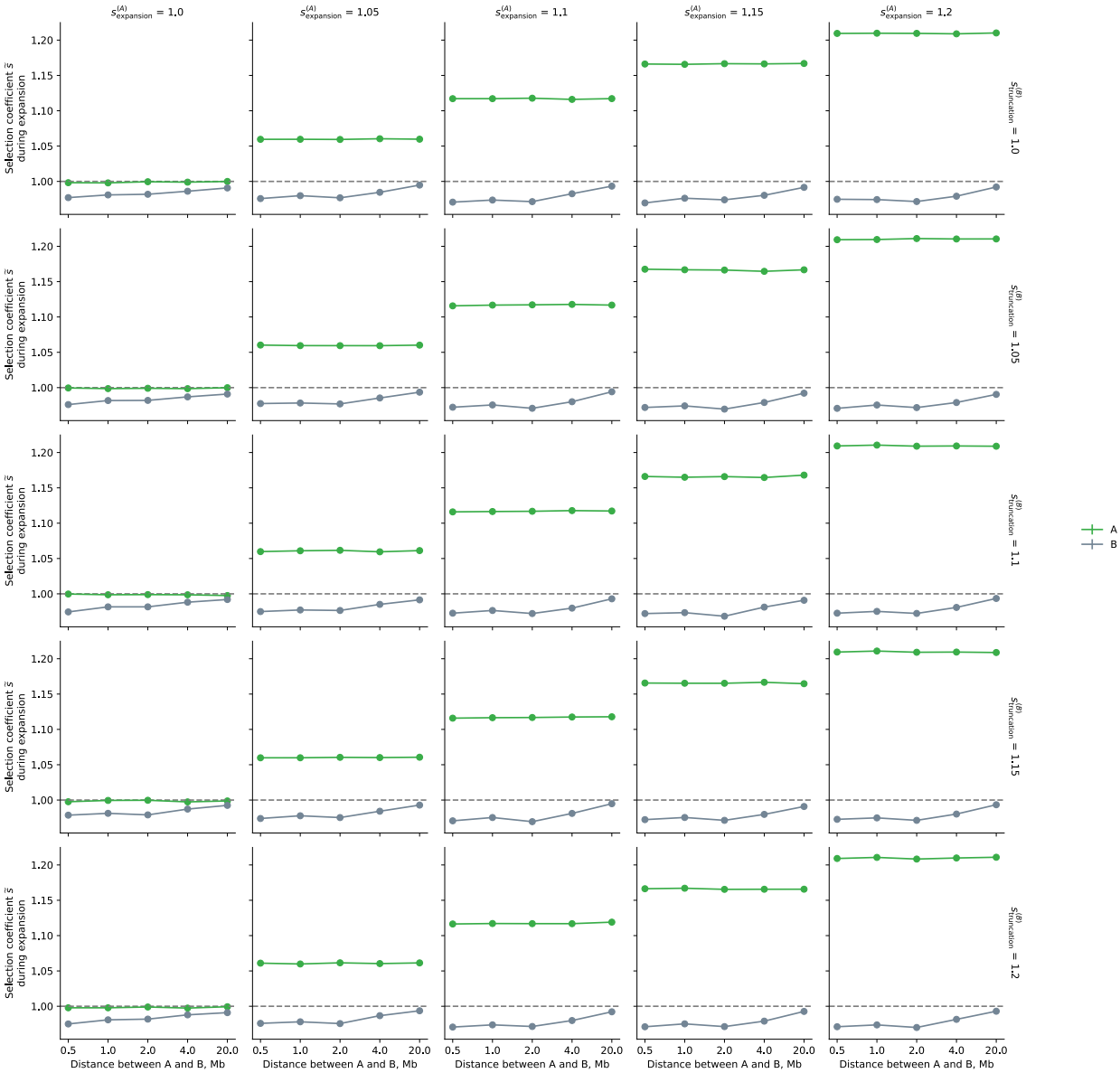

Figure S16. Fitness values (empirical selection coefficients) for two, unlinked alleles during the first generation of truncation, plotted as a function of distance between the two SNPs. Points denote the bootstrapped mean across simulations (N = 250 total simulations) with vertical lines the 95% confidence intervals (CI's often too small to be visible in plot). The 'A' allele was under directional selection during the expansion period (generations 0-9) and neutral during the truncation phase (generations 10-13). The 'B' allele was neutral during expansion, then comes under directional selection during truncation. The selection coefficient of 'A' was ranged between 1.0 and 1.2 during expansion (columns) and remains 1.0 during truncation. The selection coefficient of 'B' was constant at 1.0 during expansion, and ranged between 1.0 and 1.2 during truncation (rows).

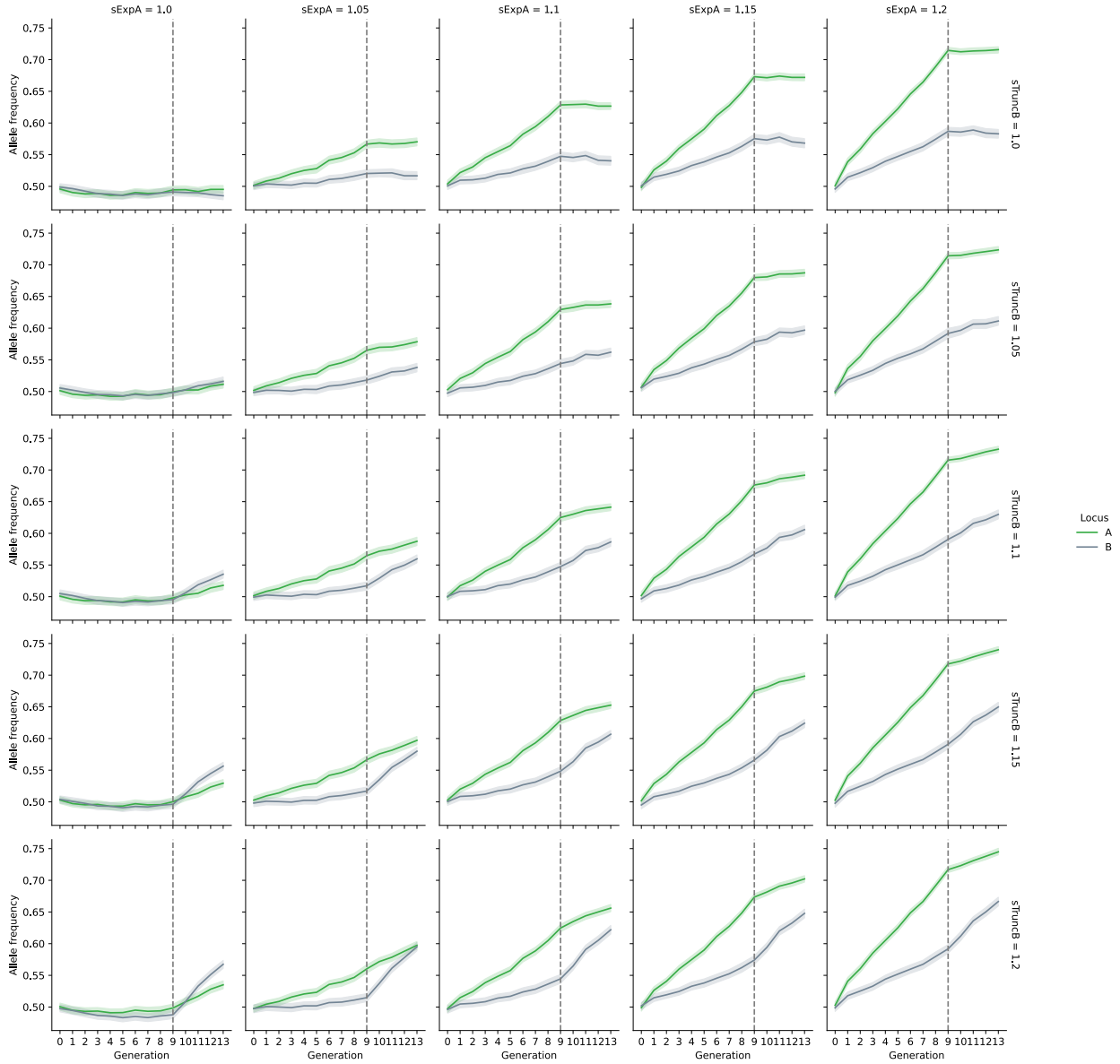

Figure S17. Simulated allele frequency trajectories for two highly linked loci in a state of attraction. The 'A' allele is encoded to be under directional selection during the expansion period (generations 0-9) and neutral during the truncation phase (generations 10-13). The 'B' allele is encoded neutral during expansion, then comes under directional selection during truncation. The selection coefficient of 'B' is constant at 1.0 during expansion and ranged between 1.0 and 1.2 during truncation (rows). The solid line corresponds to the mean frequency value of each allele across simulations and shaded area to the 95% confidence intervals (N = 250 total simulations per set of parameter values).

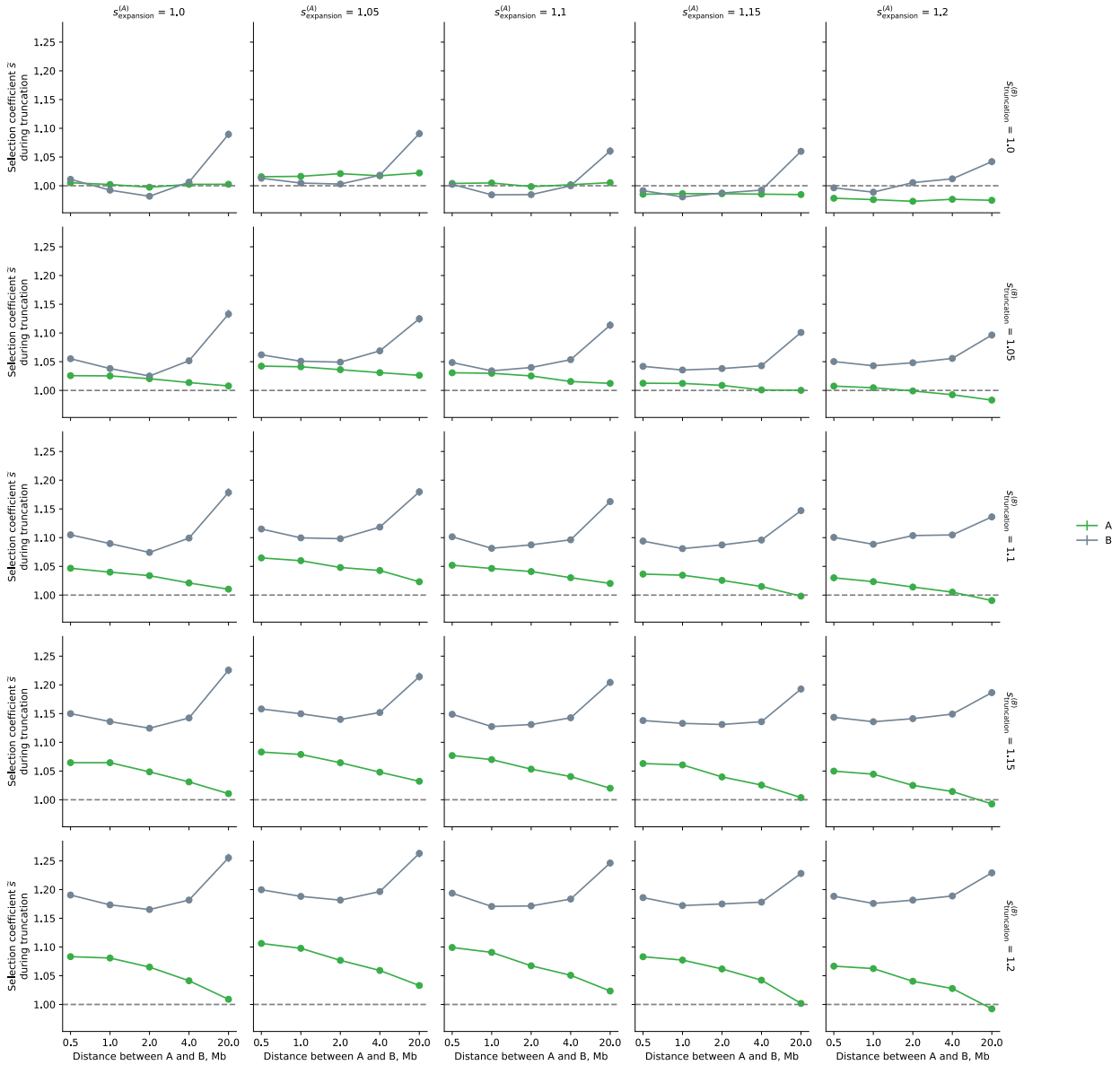

Figure S18. Fitness values (empirical selection coefficients) for two highly linked alleles in a state of attraction, computed during the first generation of truncation and plotted as a function of distance between the two SNPs. Points denote the bootstrapped mean across simulations (N = 250 total simulations) with vertical lines (often not visible in plot) the 95% confidence intervals. The 'A' allele was under directional selection during the expansion period (generations 0-9) and neutral during the truncation phase (generations 10-13). The 'B' allele was neutral during expansion, then comes under directional selection during truncation. The selection coefficient of 'A' was ranged between 1.0 and 1.2 during expansion (columns) and remains 1.0 during truncation. The selection coefficient of 'B' was constant at 1.0 during expansion, and ranged between 1.0 and 1.2 during truncation (rows).

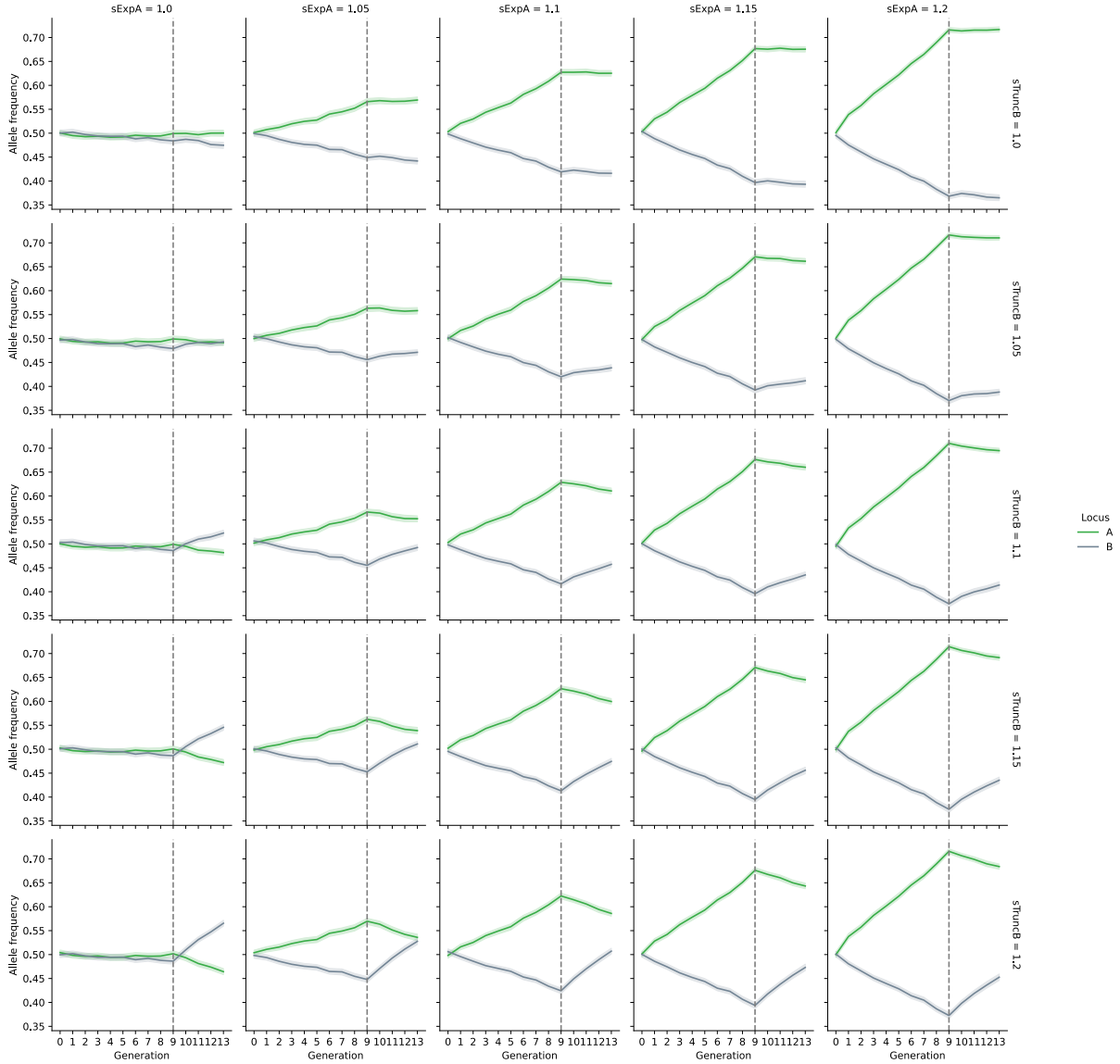

Figure S19. Simulated allele frequency trajectories for two highly linked loci in a state of repulsion. The 'A' allele is encoded to be under directional selection during the expansion period (generations 0-9) and neutral during the truncation phase (generations 10-13). The 'B' allele is encoded neutral during expansion, then comes under directional selection during truncation. The selection coefficient of 'B' is constant at 1.0 during expansion and ranged between 1.0 and 1.2 during truncation. The solid line corresponds to the mean frequency value of each allele across simulations and shaded area to the 95% confidence intervals (N = 250 total simulations per set of parameter values).

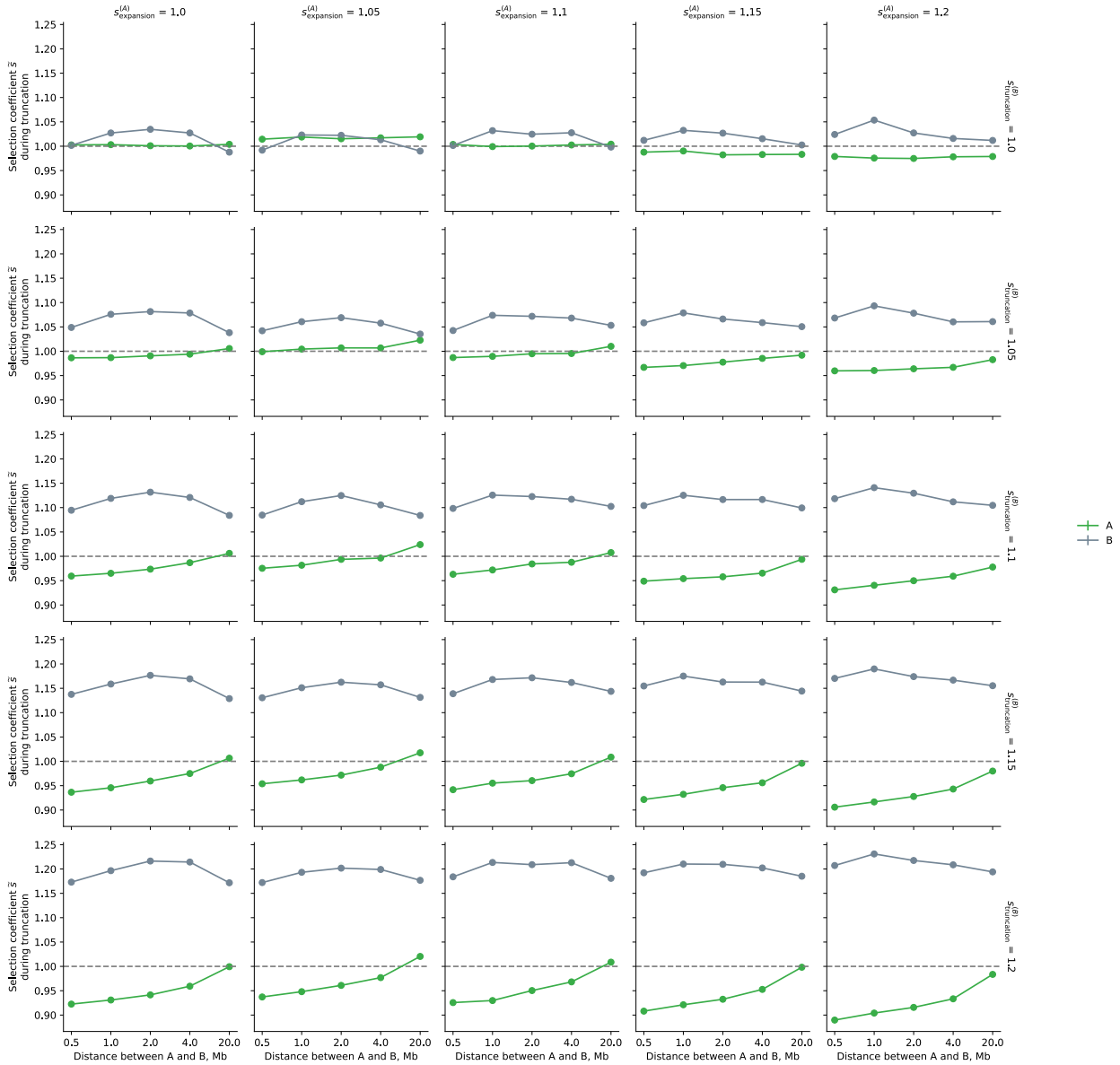

Figure S20 Fitness values (empirical selection coefficients) for two highly linked alleles in a state of repulsion, computed during the first generation of truncation and plotted as a function of distance between the two SNPs. Points denote the bootstrapped mean across simulations (N = 250 total simulations) with vertical lines the 95% confidence intervals (CI's often not visible in plot). The 'A' allele was under directional selection during the expansion period (generations 0-9) and neutral during the truncation phase (generations 10-13). The 'B' allele was neutral during expansion, then comes under directional selection during truncation. The selection coefficient of 'A' was ranged between 1.0 and 1.2 during expansion (columns) and remains 1.0 during truncation. The selection coefficient of 'B' was constant at 1.0 during expansion, and ranged between 1.0 and 1.2 during truncation (rows).

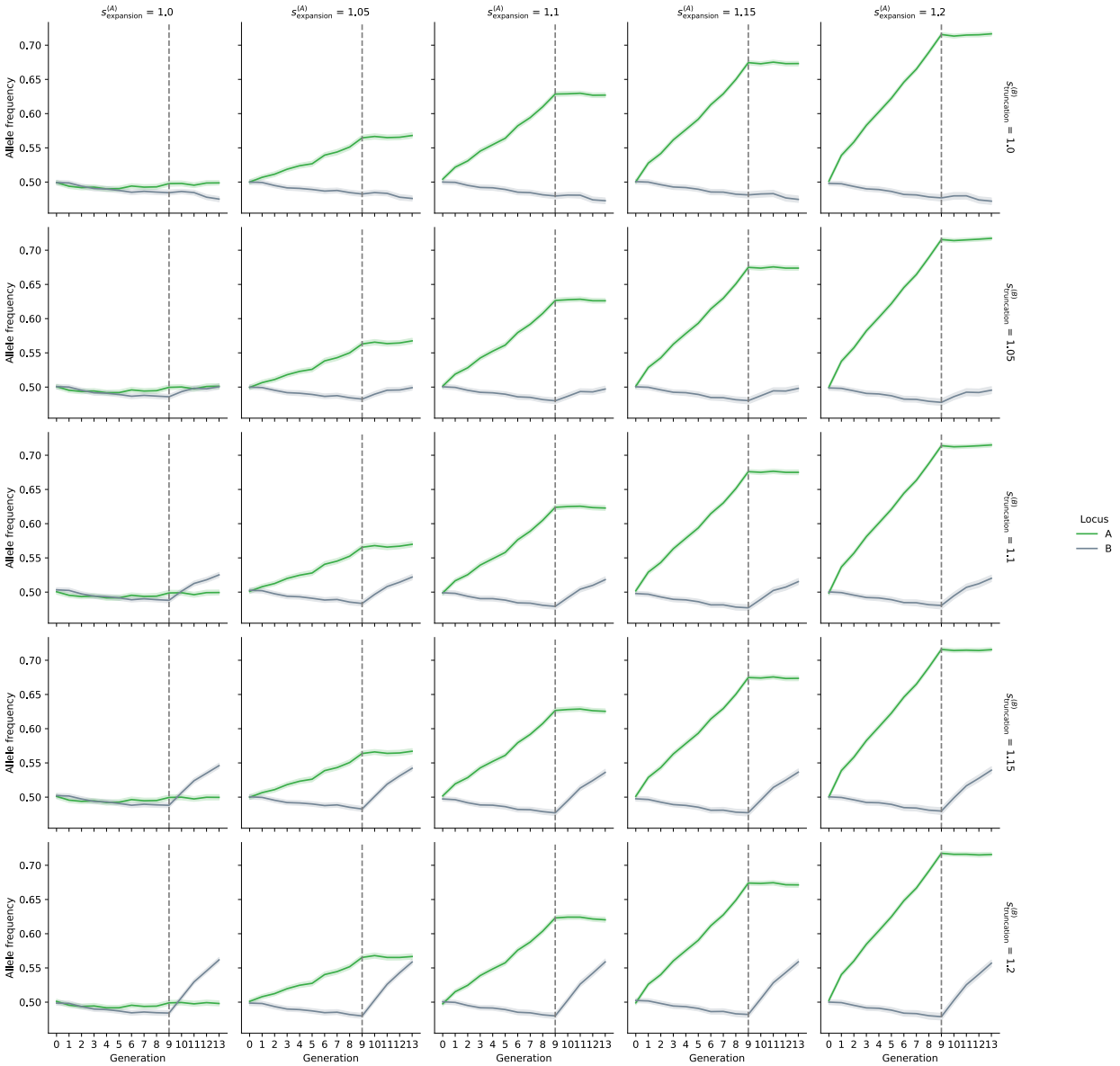

Figure S21. Allele frequency trajectories for two highly linked SNP's, with selected alleles at each SNP being independently assorted with each other. The 'A' allele is encoded to be under directional selection during the expansion period (generations 0-9) and neutral during the truncation phase (generations 10-13). The 'B' allele is encoded neutral during expansion, then comes under directional selection during truncation. The selection coefficient of 'B' is constant at 1.0 during expansion and ranged between 1.0 and 1.2 during truncation. The solid line corresponds to the mean frequency value of each allele across simulations and shaded area to the 95% confidence intervals (N = 250 total simulations per set of parameter values).

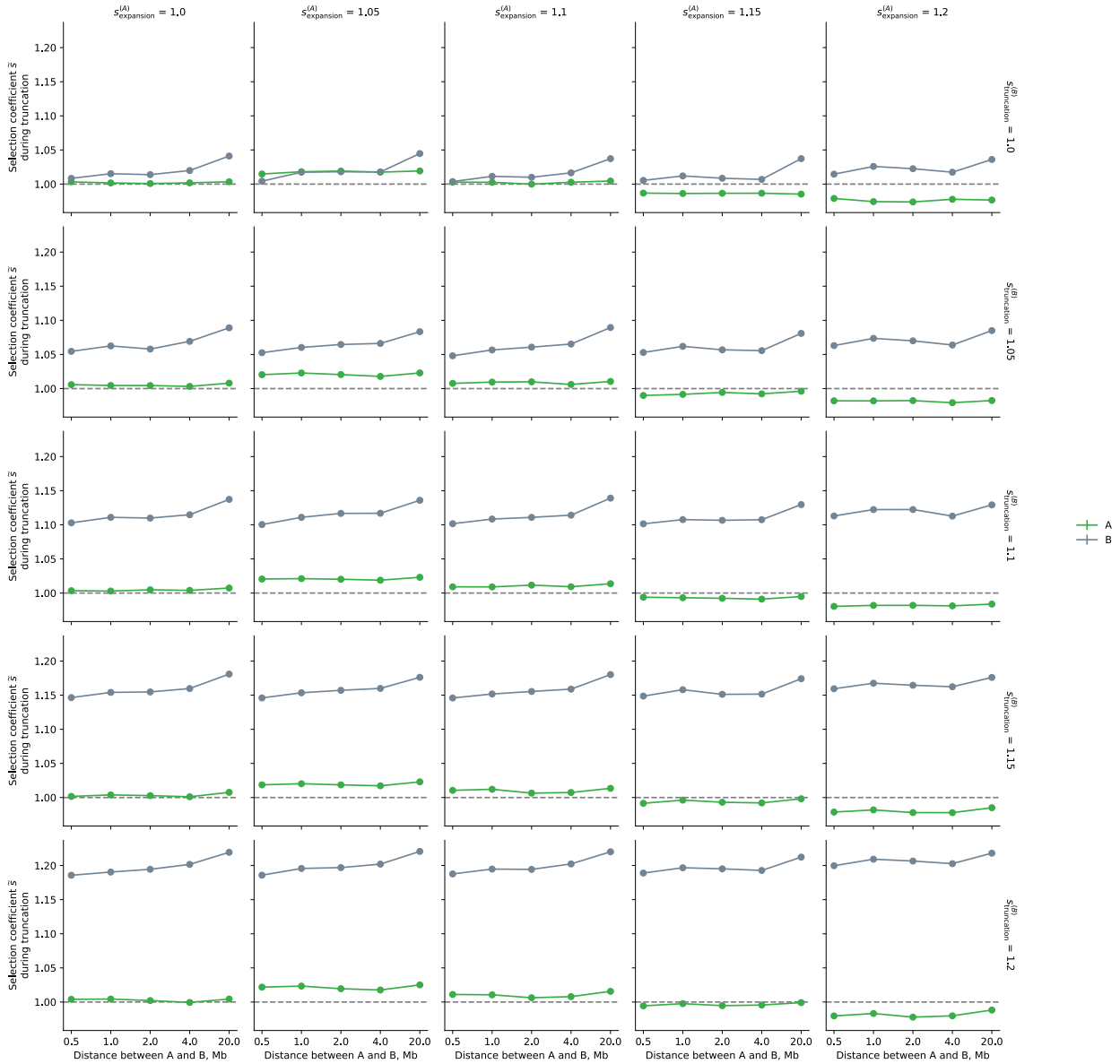

Figure S22. Fitness values (empirical selection coefficients) during the first generation of truncation for the scenario in which the advantageous alleles at linked SNPs ('A' and 'B') are independently assorted with each other. Points denote the bootstrapped mean across simulations (N = 250 total simulations) with vertical lines (often not visible in plot) the 95% confidence intervals. The 'A' allele was under directional selection during the expansion period (generations 0-9) and neutral during the truncation phase (generations 10-13). The 'B' allele was neutral during expansion, then comes under directional selection during truncation. The selection coefficient of 'A' was ranged between 1.0 and 1.2 during expansion (columns) and remains 1.0 during truncation. The selection coefficient of 'B' was constant at 1.0 during expansion, and ranged between 1.0 and 1.2 during truncation (rows).

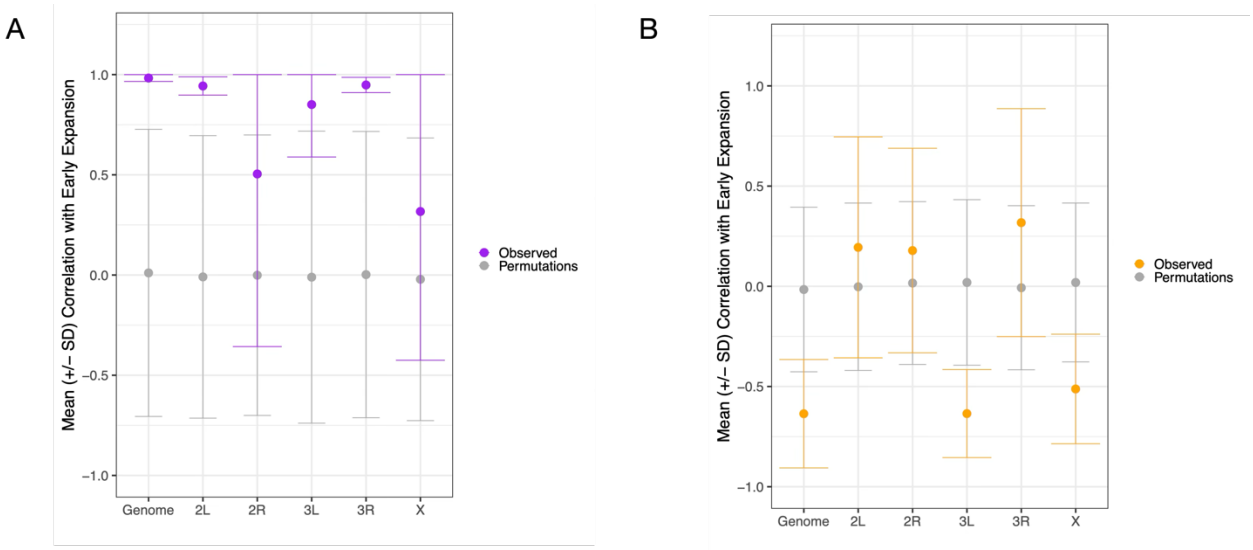

Figure S23. Average Pearson Correlation across replicates (+/- standard deviation) between sample expansion generation and distance along a one-dimensional axis constructed using F_{ST} MDS coordinates for early expansion (generation 0-4) samples. Here, Fst values were derived from SNPs excluding those within, or up to 100 Kb away from, inversion breakpoints. Purple points and error bars correspond to empirical values, while grey points and error bars correspond to values derived from empirical permutations (N = 100).

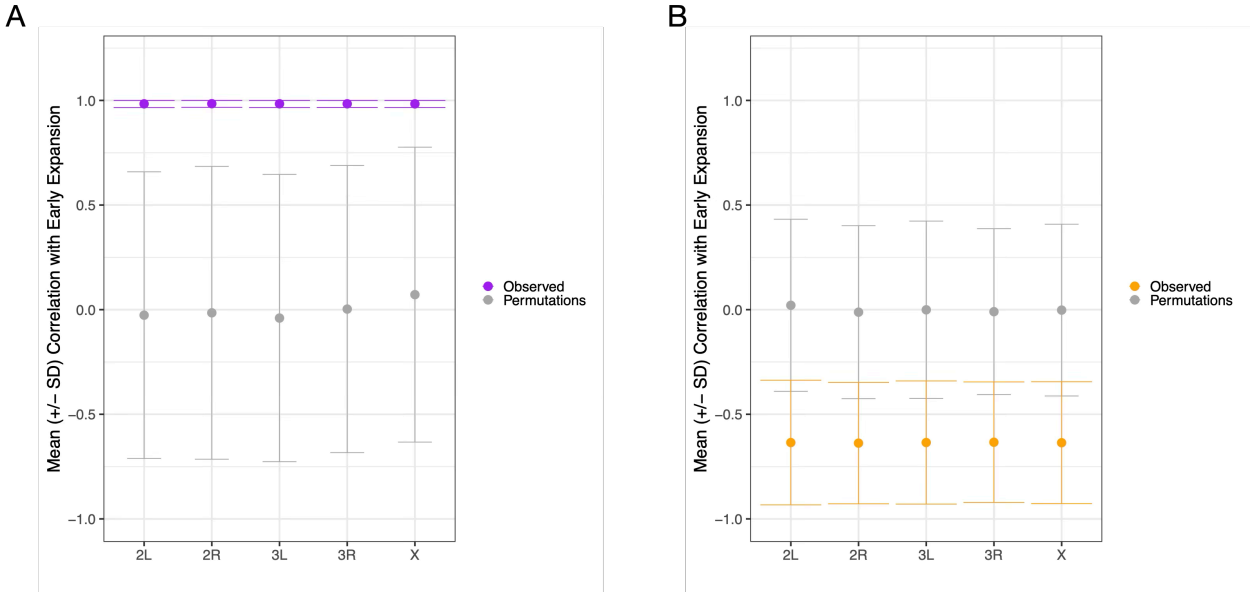

Figure S24. Fst-based MDS analysis derived from randomly sampled SNPs throughout the genome matched to the total number of SNPs present on each chromosomal arm and excluding those within, or up to 100 Kb away from, inversion breakpoints. Plotted are average Pearson Correlation across replicates (+/- 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).

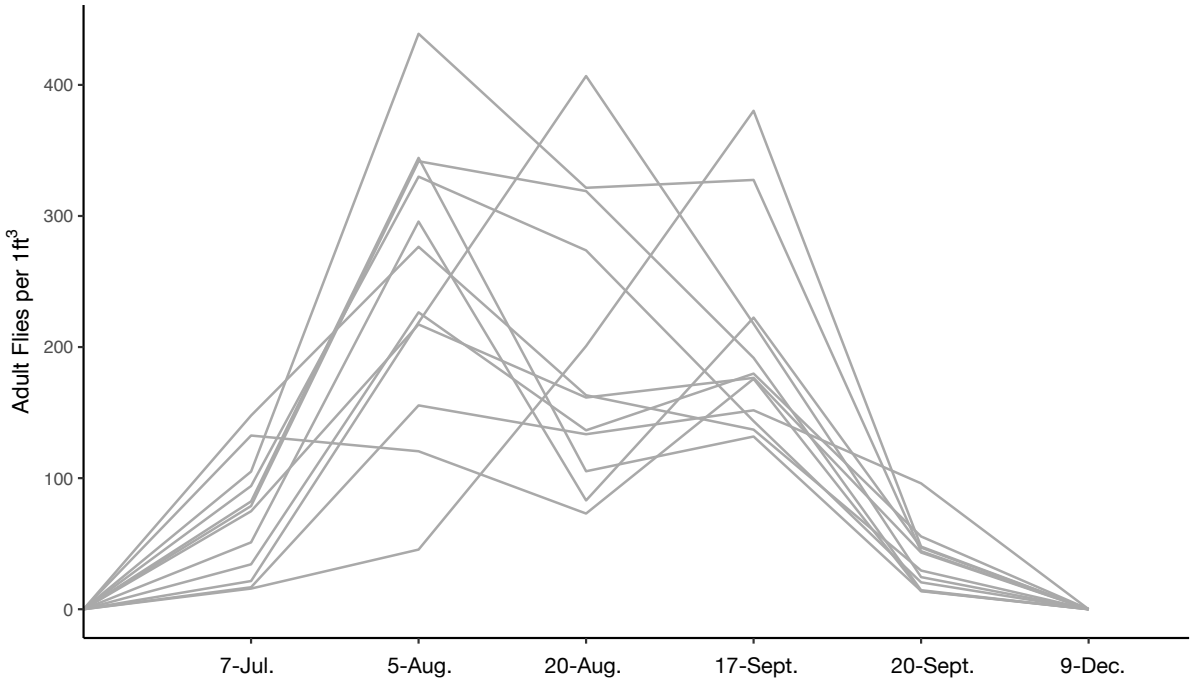

Figure S25. Census estimates of flies in the outdoor mesocosm experiment from 2021. Figure modified from (4).

Supplementary Tables

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 S2 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. 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. 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	0.975716766359953	0.07375
Late Expansion	2R	0.974541408118333	0.075
Late Expansion	3L	0.975540423627953	0.01875
Late Expansion	3R	0.9752928069539	0.02625
Late Expansion	X	0.973467774354005	0.0600000000000001
Truncation	2L	-0.751876958283634	0.02
Truncation	2R	-0.753904647531668	0.01875
Truncation	3L	-0.7502678134965	0.04125
Truncation	3R	-0.750207616089703	0.025
Truncation	X	-0.756320833115252	0.0225

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. 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. 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	0.941877272974043	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. 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. 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	0.988413197497469	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

Supplementary References

1. Haller BC, Messer PW. SLiM 4: Multispecies Eco-Evolutionary Modeling. The American Naturalist. 2023 May; 201(5):E127–39.

2. VanLiere JM, Rosenberg NA. Mathematical properties of the r2 measure of linkage disequilibrium. Theor Popul Biol. 2008 Aug; 74 (1):130–7.

3. Mathieson I, McVean G. Estimating Selection Coefficients in Spatially Structured Populations from Time Series Data of Allele Frequencies. Genetics. 2013 Mar 1; 193(3): 973–84.

4. 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.