Continuously fluctuating selection reveals extreme granularity and parallelism of adaptive tracking

M.C. Bitter[1], S. Berardi[2], H. Oken[2], A. Huynh[1], P. Schmidt[2]*, and D.A. Petrov[2, 3]*

1 – Department of Biology, Stanford University, Stanford CA 94305

2 – Department of Biology, University of Pennsylvania, Philadelphia, PA 19104

3 – Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

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

Abstract

Temporally fluctuating environmental conditions are a ubiquitous feature of natural habitats. Yet, how finely natural populations adaptively track fluctuating selection pressures via shifts in standing genetic variation is unknown. We generated high-frequency, genome-wide allele frequency data from a genetically diverse population of *Drosophila melanogaster* in extensively replicated field mesocosms from late June to mid-December, a period of ~12 generations. Adaptation throughout the fundamental ecological phases of population expansion, peak density, and collapse was underpinned by extremely rapid, parallel changes in genomic variation across replicates. Yet, the dominant direction of selection fluctuated repeatedly, even within each of these ecological phases. Comparing patterns of allele frequency change to an independent dataset procured from the same experimental system demonstrated that the targets of selection are predictable across years. In concert, our results reveal fitness-relevance of standing variation that is likely to be masked by inference

approaches based on static population sampling, or insufficiently resolved time-series data. We propose such fine-scaled temporally fluctuating selection may be an important force maintaining functional genetic variation in natural populations and an important stochastic force affecting levels of standing genetic variation genome-wide.

Main Text

Adaptation is a fundamental driver of evolutionary change in natural populations. The field of population genetics has studied this process through the development of models and empirical research that largely assume that most fitness-associated traits are under stabilizing selection and slowly evolve to track an optimal phenotype that changes incrementally through time (1). However, a ubiquitous feature of the habitats in which populations persist is sharp temporal fluctuations in the selective environment. Life has evolved distinct mechanisms to cope with these fluctuations.

At the individual-level, organisms alter their phenotype to track predictable changes in the environment via phenotypic plasticity, or produce a fixed phenotype that minimizes fitness variance amidst unpredictable environmental shifts via bet-hedging (2–5). At the population-level, populations can adapt to different states of the environment via shifts in standing genetic variation, a process termed adaptive tracking (3). What remains unknown is the temporal scale over which populations have the capacity to adaptively track fluctuations in the selective environment. Even in the face of strong, rapidly fluctuating selective pressures, adaptive tracking may be limited if a population lacks genetic variants with sufficiently large selection coefficients to drive differential survival and reproduction at the same pace as the shifting environmental conditions (6–8). Determining whether such conditions are met in natural populations hinges upon well-resolved, time-series genomic data. Procuring such data may ultimately reveal a functional relevance of standing genetic variation that remains hidden

from inference approaches using static snapshots of population genomic variation (9–13). Furthermore, such approaches would help clarify the extent to which temporally fluctuating selection could act to maintain adaptive variation in natural populations, an issue of longstanding theoretical and empirical debate (14–19).

Seasonal environments, which fluctuate cyclically and on a timescale that lends to repeated empirical observation, are well-suited to quantify how finely populations adaptively track fluctuating selection pressures. For example, populations of *Drosophila* spp. inhabiting temperate environments expand exponentially from spring through summer as resources become abundant, then crash during the onset of winter as abiotic conditions deteriorate and resources become scarce (20). These boom-and-bust dynamics are a dominant feature of the demographic trends in a range of species, occur on scales shorter and longer than the seasonal scales (21–23), and have been shown to drive the evolution of life-history, reproductive, and stress tolerance traits (24–26).

Investigation into the population genetics of adaptation in seasonally fluctuating environments has a long history (27–30). Recent studies that have leveraged genome-wide sequencing of *D. melanogaster* populations have shown how adaptive tracking keeps pace with fluctuating selection on seasonal timescales (31–33). Notably, Rudman et al. (2022) established that parallel, genome-wide signatures of fluctuating selection can be observed on monthly intervals (~1-3 generations) across ten replicate populations, suggesting that changes in the direction of selection may occur due to more fine-scale shifts in selective pressures than the most conspicuous environmental fluctuations across seasons (33). Still, the underlying granularity of this process remains unknown: do alleles responding to fluctuating selection over monthly intervals exhibit parallel, deterministic changes on even finer scales in response to the exceptionally dynamic environments in the system, which shift daily across multiple biotic and abiotic axes? Resolving the ultimate limit of adaptive tracking in this system would

76 in turn lend broader insight into the power of temporally fluctuating selection as a process that
77 could preclude the fixation of adaptive alleles, and thus help to maintain variation, in natural
78 populations. Furthermore, repeated observation in this system may be leveraged to probe the
79 predictability of the targets of selection across independent bouts of adaptive tracking, an
80 outstanding issue of widespread practical importance (34–36).

81 Here, we aimed to resolve the scale and pervasiveness of fluctuating selection and the
82 predictability of adaptive tracking via high-frequency monitoring of a genetically diverse
83 population of *D. melanogaster* in an extensively replicated field mesocosm system (Fig. 1).
84 Specifically, we seeded twelve replicate outdoor cages with a genetically diverse population
85 (hereafter, baseline population), generated via four generations outbreeding a panel of inbred
86 reference lines that were originally collected in local Pennsylvania orchards. The experiment
87 spanned from early summer through fall, a period of 10-12 generations (25), during which
88 time the replicate populations were subject to the same general shifts in abiotic and biotic
89 selective pressures across seasons. A generic demographic trend in the system was observed,
90 whereby population sizes expanded rapidly after cage founding, stabilized at a peak density
91 by late summer, and declined during fall (20, 33). Notably, the population trajectories
92 exhibited distinguishable behavior, likely representing unique environmental exposures that
93 alter the particular selective environment within each replicate at any given point in time (Fig.
94 1). We sampled and quantified patterns of genomic variation weekly between 13 July and 7
95 September 2021, after which samples were collected on 21 September, 20 October, and 20
96 December (12 total time points; Fig. 1).

97 The sampling scheme allowed powerful inference of adaptive tracking within the
98 system. Specifically, continuous birth, aging, and mortality of flies throughout the experiment
99 induced a shifting age structure within the replicate populations. As individual viability and
100 fecundity vary as a function of age, it was critical to ensure that our quantification of allele

frequencies reflected shifts to the instantaneously reproductive genotypes, or 'leading edge', of the population (37). Accordingly, at each time point eggs were collected overnight directly from the mesocosms, and larval development was carried out in lab-based conditions. One hundred adult female flies from the resulting cohort were collected and sequenced in pools, from which we generated high-accuracy, haplotype-informed allele frequencies at 1.9 million single nucleotide polymorphisms (SNPs) (38; see Methods). The resulting data from twelve cages and time points were used to quantify parallel, deterministic shifts in allele frequencies across replicates and infer instances of fluctuations in the selective environment over a range of temporal scales. We additionally compared observed patterns of allele frequency change to an independent study using the same mesocosm system (33) to explore the predictability of the targets of selection during independent bouts of adaptive tracking.

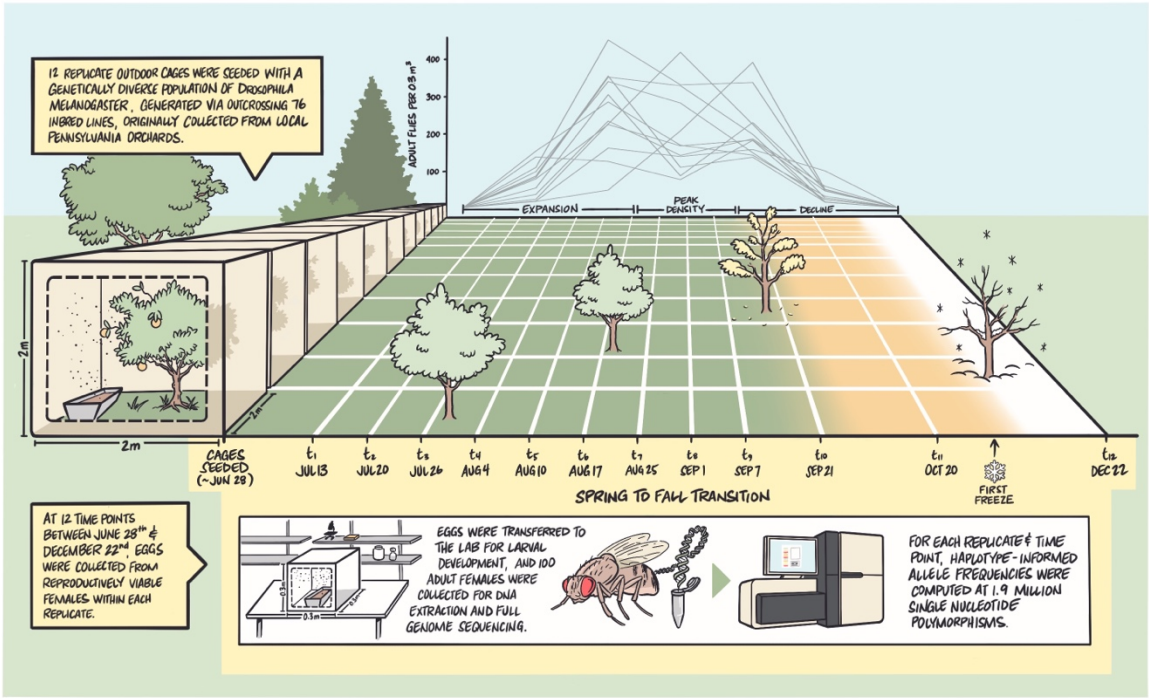


Figure 1 Experimental Schematic | Twelve replicate outdoor cages in Pennsylvania were seeded with an outbred *D. melanogaster* population, generated via outcrossing 76 inbred lines, originally collected in local orchards. The replicate populations expanded exponentially through mid-summer before declining in early fall, and ultimately

128 crashing following the first freeze in November. At twelve time points throughout the experiment, eggs from the
129 reproductively viable females within each replicate were collected, *in situ*, overnight. The eggs were then
130 transferred to controlled, lab-based conditions for development. One hundred randomly selected female flies
131 from each sample were collected and used for pooled, whole genome sequencing. The resulting sequencing data
132 from each replicate and time point was then used for haplotype-informed allele frequency estimation at 1.9 M
133 single nucleotide polymorphisms previously identified in the inbred reference panel.

134 *Parallel allele frequency change across summer to fall transition*

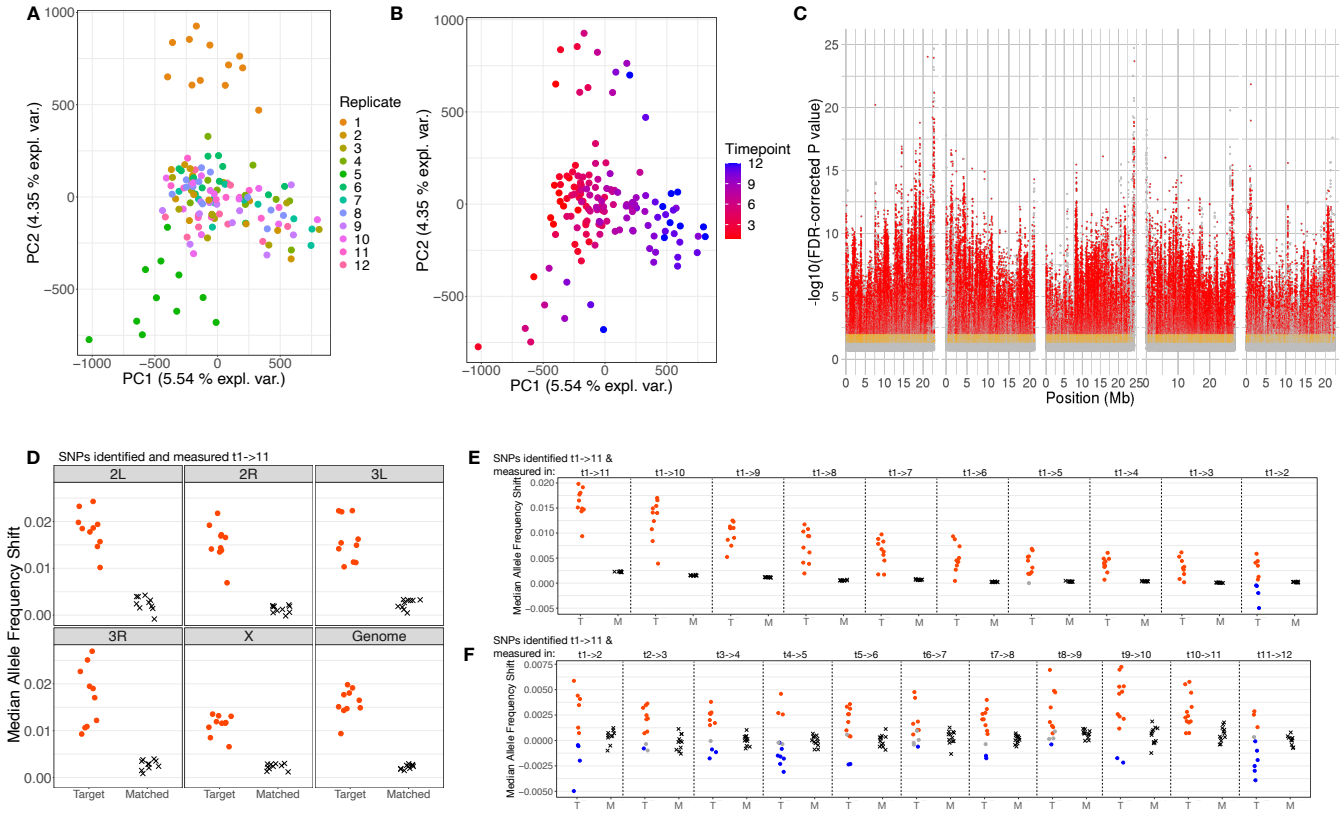
135 We aimed to detect whether amidst the inherent complexity of our outdoor study
136 system (Fig. 1), there existed evidence of parallel shifts in genomic variation in response to the
137 shared selective pressures experienced across replicates. We first computed genome-wide F_{ST}
138 between the replicate populations and the baseline to ensure that the evolution of allele
139 frequencies quantified throughout the experiment exceeded sources of biological and technical
140 noise (Fig. S1). Next, we used principal component analysis (PCA) to assess the impact of two
141 key experimental factors, replicate cage and collection time, on variance in allele frequencies
142 throughout the experiment. Segregation of samples along the first principal component (5.54 %
143 variance explained) indicated that variation in allele frequencies across all samples was
144 predominantly driven by collection time point (Fig. 2A-B), a pattern maintained when the
145 analysis is run separately across each chromosomal arm (Fig. S2). The segregation of two
146 replicate cages along the second principal component (Fig. 2A) indicates that background
147 allele frequencies across replicates were potentially perturbed by population bottlenecks
148 during cage founding. Yet, these two cages maintain the temporal trend along PC1, further
149 indicating parallel shifts in allele frequencies throughout the experiment (Fig. 2B). An
150 unsupervised clustering (K-means) of the PCA results further validated, over a range of cluster
151 sizes, that the predominant axes of variation in the data were most strongly associated with
152 sample collection time (Fig. S3). In concert, these results suggest that the shared selective

153 pressures experienced by the replicates induced concordant changes in genomic variation
154 throughout the experiment.

155 We next sought to identify the number and genomic distribution of SNPs underpinning
156 parallel changes in genomic variation. Accordingly, we fit allele frequencies to a generalized
157 linear model (GLM; with quasibinomial error variance) to identify SNPs moving
158 systematically across replicates throughout the entire spring to fall transition (time points 1-
159 >11; before the total crash of the adult population occurring between time points 11 and 12).
160 This yielded pervasive, genome-wide signal, whereby 28.5% of all SNPs exhibited a Benjamini-
161 Hochberg FDR-corrected P-value < 0.05 (Fig. 2C). Using an algorithm to identify clusters of
162 unlinked loci enriched for significant SNPs, we identified 264 loci underpinning this signal,
163 located across all chromosomal arms and inside and outside major cosmopolitan inversions
164 (33; see Methods) (Supplementary Data File 1). This suggests an oligogenic architecture of
165 adaptive tracking in the system, driven by alleles of relatively large effect size. Specifically, the
166 range of allele frequency shifts for the top SNP of each unlinked cluster (defined as lowest P-
167 value SNP) was 2-11% (median 5.5%), values on par with previous documentation of seasonal
168 allele frequency change in the species (31, 32).

169 We quantified the degree of parallel, neutral, and/or anti-parallel evolutionary change
170 for individual replicate cages by re-implementing our GLM analysis using a leave-one-out
171 cross-validation approach: sets of GLM significant SNPs were identified iteratively in n -1
172 'training' cages, and the frequency dynamics of the rising alleles (FDR < 0.2, allele frequency
173 change > 0.5 %) were then quantified in the left-out, 'test', cage. Each target SNP was matched
174 to a control SNP on the same chromosomal arm and with comparable baseline frequency,
175 recombination rate, and inversion status (see Methods). The distributions of target and control
176 shifts for each cage were then compared to: 1) determine whether the magnitude of target
177 shifts exceeded background allele frequency shifts and therefore the target SNPs exhibited

178 non-neutral dynamics, and 2) infer whether the direction of non-neutral shifts was in a
179 predominately concordant/parallel direction as that observed in the training cages. This
180 revealed a striking degree of parallelism across all replicates, whereby the median shift of
181 target SNPs within left-out cages was significantly greater than allele frequency change at
182 matched control sites (Fig. 2D). These patterns suggest that even amidst the variation in the
183 genomic composition driven by early founder effects (Fig. 2A), genome-wide allele frequency
184 change in the system exhibited detectably parallel behavior throughout the summer to fall
185 transition.
186



187

188 **Figure 2 Genome-wide patterns of parallel allele frequency change throughout spring to fall transition** |
189 (A/B) Principal component analysis of normalized allele frequency data (1.9 M SNPs) across all replicates and
190 time points. Sample scores are projected onto the first two principal components and colored by replicate cage

191 number (A) and collection time point (B). (C) Manhattan plot depicting SNPs identified as significant using a
192 generalized linear model and regression of allele frequencies across the entire summer to fall transition (t1->11).
193 Significant SNPs (points) are visualized as a function of genomic position and FDR-corrected P-value, and colored
194 as follows: FDR < 0.2 and effect size > 0.5% (grey); FDR < 0.05 and effect size > 2% (orange); FDR < 0.01 and effect
195 size > 2% (red). (D) Leave-one-out cross-validation of SNPs depicted in (C) whereby the phased frequency
196 dynamics of target SNPs for each left-out replicate are compared to a set of matched control SNPs. Portrayed is
197 the median shift of target and matched control SNPs for each cage, separately for each chromosomal arm and
198 genome-wide. Target site shifts are colored red if the distribution of phased allele frequency shifts was
199 significantly greater, and blue if significantly less, than that of matched control sites (paired t-test, FDR < 0.05).
200 (E/F) Median allele frequency shifts at target (T) and matched (M) SNPs depicted in (D), but measured in
201 sequentially shorter intervals (E) and across all pairs of single time points (F). Coloring scheme same to that used
202 in (D).

203 *Fluctuating selection revealed by higher resolution measurement of allele frequencies*

204 Are alleles exhibiting parallel and systematic changes throughout the summer to fall
205 transition subject to sustained, directional selection throughout this period? Or are fluctuations
206 in the selective environment sufficiently strong to drive reversals in the direction of frequency
207 change on finer timescales? To explore if the direction of selection on t1->11 SNPs was
208 constant throughout the experiment, we first quantified allele frequency shifts of each set of t1-
209 >11 target SNPs (identified above; Fig. 2C/D) in sequentially shorter intervals. In this way we
210 explored whether there exists a limit at which replicates no longer exhibited deterministic
211 shifts in allele frequency (the magnitude of phased allele frequency shifts no longer differed
212 from matched control shifts) and/or whether evidence of anti-parallel behavior emerged on
213 shorter intervals (significantly negative, median phased shifts at target SNPs), indicating
214 fluctuating selection for the variants that systematically moved in a consistent direction from
215 the beginning to the end of the experiment. As the dynamics of each set of target SNPs were
216 measured using the left-out cage, inferred changes in the direction of SNP movement are
217 insensitive to statistical artifacts, such as regression to the mean.

218 We observed persistently deterministic, and parallel behavior of target SNPs from the
219 longest interval (t1-> 11; 15 weeks), to the two-week interval (t1->3) (Fig. 2E). The magnitude
220 of the allele frequency change declined monotonically with test interval length. Strikingly, our
221 shortest interval length (1 week, t1->2) continues to indicate predominantly deterministic
222 shifts at target SNPs, but also provides us with the first evidence of anti-parallel behavior
223 across replicates (both red and blue points observed within measurement interval; Fig. 2E).
224 Figure 2F expands the set of test intervals to include all pairs of single time points (t1->2, t2-
225 >3…t11->12), the first nine of which are weekly intervals. Concordant with the t1->2
226 comparison, we see largely deterministic shifts at target sites in left-out cages, yet both parallel
227 and anti-parallel behavior. This indicates that even those SNPs inferred to shift systematically
228 and in parallel throughout the summer to fall transition (Fig. 2D), were in fact subject to
229 numerous bouts of fluctuating selection on much finer temporal scales.

230 *Continuously fluctuating selection underpins rapid adaptive tracking*

231 How many times does the direction of selection fluctuate throughout the experiment?
232 Most cages experienced at least one reversal in the direction of selection on SNPs identified
233 from t1->11 (Fig. 2E). We hypothesized that sets of alleles identified on shorter intervals, with
234 negligible net change throughout the experiment, may reveal even greater incidence of
235 fluctuating selection in the system. We first defined the shortest interval length over which we
236 had power to detect SNPs shifting systematically, and in parallel, across replicate cages using
237 GLM (Fig. S4). The models regressing allele frequencies across five time point intervals (e.g. t1-
238 >6, t2->7…t7->12) represented the shortest interval length in which all comparisons yielded
239 pervasive, genome-wide significant and large effect size variants across all chromosomal arms
240 (Fig S5-13). We thus focused on this interval length to explore the complex selective
241 environment observed in Figure 2F. For this analysis, we discarded those SNPs identified as
242 strongly parallel (GLM FDR < 0.05) across the t1->11 interval, as these exhibit predominantly

243 directional change (Fig. 2E) and may mask the finer-scale fluctuations in the selective
244 environment we ultimately aimed to uncover.

245 We first used leave-one-out to quantify the behavior of SNPs discovered for all five time
246 point intervals across the time points over which they were identified. In accordance with
247 SNPs identified from t1->11, those identified on five time point intervals exhibited largely
248 parallel shifts across replicates: for the majority of left-out cages and intervals, the magnitude
249 of frequency shifts at target SNPs were significantly greater than those at matched control
250 SNPs (red points, Fig. 3A). Notably, during the first two intervals (t1->6 and t2->7) two cages
251 (replicates 1 and 5) exhibited predominately anti-parallel allele frequency change (blue dots).
252 These are the same cages offset along PC2 in Figure 2A/B, and this dynamic likely provides
253 further illustration of how bottlenecks associated with cage founding can drive idiosyncratic
254 behavior across replicates in the early season that are ultimately overcome by parallel shifts in
255 genomic variation once population sizes become sufficiently large (Fig. 1).

256 Next, we quantified the number of times fluctuating selection produced changes in the
257 inferred direction of selection acting on replicate populations throughout the experimental
258 period. Specifically, we measured the shifts of target SNPs and matched controls identified on
259 five time point intervals across all pairs of single time points (i.e., t1->2, t2->3…t11->12; as in
260 Figure 2F). Figure 3B-C depicts the results of this analysis for sets of SNPs identified from t4-
261 >9: points are colored corresponding to replicate cage and represent the median shift of test
262 SNPs (B) or matched control SNPs (C), with grey lines connecting the shifts for each cage
263 across measurement windows. Across replicates, the median number of flips in the dominant
264 direction of selection was five (flips inferred when shifts at test SNPs changed from being
265 significantly greater to significantly less than matched controls, or vice versa). This indicates
266 that selection coefficients of selected alleles repeatedly flip in sign from positive to negative
267 over weekly timescales, and that the replicate populations adaptively track fluctuations in the

environment that occur over much finer timescales than the dominant ecological phases (e.g., population expansion, peak density, or decline; Fig. 1). The underlying magnitude of median allele frequency change at target SNPs throughout these flips oftentimes exceeds 1% (Fig. 3D), corresponding to selection coefficients that exceed 10 % per generation and indicating exceptionally strong fluctuating selection on this timescale. Our leave-one-out approach and comparison to matched controls ensures these fluctuations are not a product of regression to the mean, or that the identified SNPs simply exhibit stochastic/noisy behavior due to their intrinsic characteristics (e.g., starting frequency, recombination rate, or inversion status). The supplement provides results of this analysis across all interval lengths and possible comparisons, yielding qualitatively similar dynamics (Fig. S16-23).

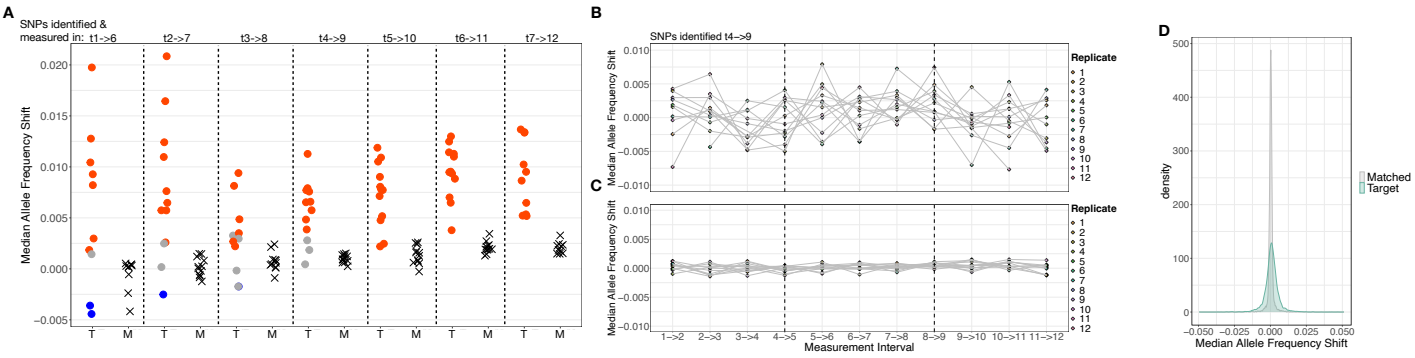


Figure 3 Pervasiveness of fluctuating selection | (A) Median allele frequency shift for target (T) and matched control SNPs (M) for each left-out cage, using sets of SNPs identified and measured within each five time point interval. Median shifts for target SNPs are colored red if the distribution of phased allele frequency shifts was significantly greater, and blue if significantly less, than that for matched control sites (paired T-test, FDR < 0.05). (B/C) Median allele frequency shifts in each left-out cage (as in A) measured within each pair of single time points for sets target SNPs (B) and matched control SNPs (C) identified t4->9. (D) The distribution of median allele frequency shifts for sets of target (green bars) and matched control SNPs (grey bars) derived from all five time point intervals and measured across pairs of single time points (variance of distributions is significantly different; F-test, F = 0.17, P value < 0.001).

289 *Predictable patterns of genomic variation underpin adaptive tracking*

290 Our results thus far indicate that even amidst differences in starting allele frequencies
291 (Fig. 2A/B), adaptive tracking across the twelve replicate cages proceeded via detectably
292 parallel shifts in genome-wide variation (Fig. 2D). Similarly, using the same mesocosm
293 system, a largely overlapping set of founder lines, and monthly allele frequency estimation
294 between July and November of 2014, Rudman et al. (2022) identified parallel allele frequency
295 change across ten replicate populations. This independent dataset from the same system
296 provides the unique opportunity to ask: how predictable are the genomic underpinnings of
297 adaptive tracking across independent bouts of evolution in cyclically fluctuating
298 environments? Across both experimental years, the replicate cages exhibited boom-and-bust
299 dynamics and were exposed to the same generic seasonal shifts in abiotic conditions and
300 showed parallel genomic shifts across the cages (33). However, it is possible that differences in
301 starting allele frequencies, in combination with year-specific environmental conditions, drive
302 largely unpredictable patterns of genomic variation from year to year. While there is
303 substantial evidence for predictable phenotypic trajectories in response to repeated selective
304 regimes in this and other systems (25, 39), evidence of predictability at the genetic level is
305 largely confined to traits with simple genetic bases (40–42) (but see (43)).

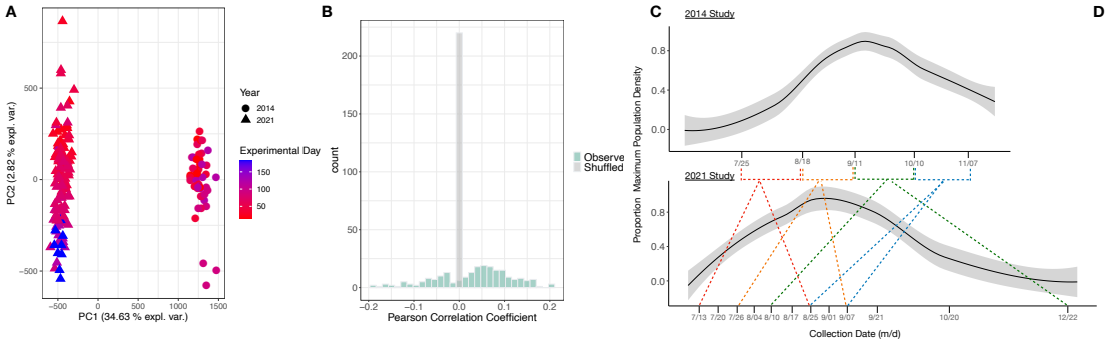
306 We tested whether independent bouts of adaptive tracking were predictable at the SNP,
307 gene, and pathway-level by comparing our results to those reported by Rudman et al (2022).
308 We first explored genome-wide patterns of differentiation across all samples collected by each
309 study using PCA. While the clustering of samples by year along the first principal component
310 indicated strong differentiation of genome-wide allele frequencies between studies, the
311 separation of early from late season samples along the second principal component suggested
312 that changes in allele frequencies through time may have been similar across years (Figure
313 4A). To test this systematically, we computed Pearson's correlations between genome wide

314 allele frequency shifts from all possible intervals of the current study (all two to eleven time
315 point intervals) and each monthly interval of the Rudman et al. study. These correlations
316 provided a directional metric of concordance, whereby positive correlations indicated that
317 genome-wide allele frequency changes were predominantly in the same direction between the
318 tested 2014 and 2021 intervals, and negative correlations to genome-wide SNPs exhibiting
319 predominantly anti-parallel behavior across compared intervals. To infer whether the
320 magnitude of observed correlations exceeded those expected by chance, we computed a null
321 distribution of correlations by randomly shuffling the frequency changes for each comparison
322 across SNPs (values shuffled between SNPs matched by chromosomal arm, starting frequency,
323 recombination rate, and inversion status) and recomputed each comparison's correlation. The
324 distributions of empirical correlations exhibited an enrichment of comparisons that were both
325 significantly correlated and anti-correlated between years (Fig. 4B; F-test, $F = 3821$, $P << 0.001$),
326 a pattern that is maintained when running this analysis separately across each chromosomal
327 arm (Fig. S24). This enrichment of elevated correlations indicates that the same alleles are
328 repeatedly subject to selection across years.

329 To explore whether those intervals with heightened allele frequency concordance across
330 years aligned with conspicuous shifts in the selective environment, we identified the interval
331 of the current study that was best matched (i.e., highest genome-wide correlation of allele
332 frequency shifts) to each of the four monthly intervals of the 2014 experiment. Strikingly, the
333 first three monthly intervals of the 2014 experiment were best matched to the t1->7, t3->9, t5-
334 >12 intervals of this study, generally overlapping with the periods of population expansion,
335 peak density, and decline (Fig. 4C). The final monthly interval was best matched to T7->9,
336 another mid-late season period, but we note that the correlation coefficient for this matching
337 was an order of magnitude less than those for the first three monthly intervals (though still
338 significant based on shuffled values). Ultimately, these matchings suggest that amidst

339 stochasticity in year-to-year abiotic conditions, the repeatable boom and bust dynamics of the
340 populations induce a repeatable ecological filter that shapes patterns of selection that
341 recurrently affects the same set of loci. Such patterns generally conform to long-standing
342 ecological hypotheses regarding the distinct selective environments associated with
343 population expansion and contraction and corroborate mounting evidence of pervasive,
344 genome-wide selection throughout the progression of seasonal adaptation (21, 27–29, 31–33).
345 The enrichment of SNP-level parallel dynamics across years encouraged a more
346 thorough investigation into the parallel molecular and phenotypic targets of selection. We first
347 explored whether the same genes were repeatedly selected for during each independent bout
348 of adaptive tracking. Specifically, Rudman et al. developed an algorithm to identify clusters of
349 unliked loci enriched for GLM signal across timepoint comparisons (introduced above, see
350 Methods). They then identified the gene containing the most significant SNP for each unlinked
351 cluster, identifying a total of 115 genes. We compared this list of genes to the top genes
352 identified throughout the spring to fall transition (t1->11), as well as to all genes identified in
353 the intervals best matched to the Rudman et al. monthly intervals (T1->7; T3->9; T5->12; T7-
354 >9) (480 total genes). We observed an overlap of 20 genes between the 2021 and 214
355 experiments, significantly more than expected by chance (3 expected; permutation P value <
356 0.001). Rudman et al. used the top genes identified to infer the predominant targets of
357 pathways under selection via a phenotype enrichment analysis (modPHEA; Weng and Liao,
358 2017). Their analysis suggested genes involved in sensory perception, synaptic functioning,
359 behavior, cell-to-cell signaling, and neuronal function to drive patterns of adaptation across
360 seasons. The same analysis conducted on the genes identified here yielded strikingly similar
361 results with a repeated overrepresentation of pathways associated with organismal
362 environmental sensing (Fig. 4D). Re-running this analysis, but excluding those genes identified
363 as overlapping between study years, suggested the same pathways dominating selection in

this study, suggesting that this result is not solely driven by a small set of overlapping genes with large effect size (Table S1).



2014 Enriched Phenotypes	2021 Enriched Phenotypes
bouton	**nervous system**
synapse	**neuron**
imaginal disc	**peripheral nervous system**
adult central nervous system	cuticle
neuron	**head**
central nervous system	muscle system
ganglion	**central nervous system**
ostium	**embryonic/larval nervous system**
fused midgut primordium	embryonic/larval tracheal system
midgut constriction	**adult sensillum**
stage 16 embryo	**adult head**
embryonic/larval alimentary canal	**ganglion**
embryonic/larval gut	**adult nervous system**
adult brain	embryonic/larval muscle system
adult dorsal vessel	tracheal system
denticle	**head sensillum**
embryonic/larval tracheal system	cuticular specialization
embryonic/larval digestive system	**sclerite**
adult head	appendage
adult circulatory system	organism

Figure 4 Predictable patterns of selection across independent bouts of adaptive tracking | (A) PCA of all samples collected during monitoring of adaptive tracking within the mesocosms in 2014 and 2021, colored by experimental day (days since replicate cage founding). (B) Histogram distribution of Pearson correlations computed between genome-wide allele frequency shifts for each monthly interval of the 2014 experiment (N = 4), and each interval of the 2021 experiment (N = 56). Bar color correspond to correlation coefficients generated via site information permutations (grey) and observed values (green). (C) Mesocosm-wide population demographic dynamics for the 2014 and 2021 experiments, whereby mean population size is standardized to maximum observed population density. Dashed lines connect each monthly interval of the 2014 experiment to the best matched interval of the current study, based on genome-wide Pearson correlation between allele frequency shifts. (D) Overrepresented phenotypic categories identified using top genes underpinning evolution in the 2014 and 2021 experiments. Phenotypes putatively associated with environmental sensing are bolded.

Discussion

The pervasive fluctuating selection we report here reveals key insights into the dynamics of adaptive tracking, with implications for patterns of genome-wide diversity and the maintenance of genetic variation in natural populations. Specifically, we identified changes in the dominant direction of selection repeatedly over multiple weekly intervals throughout the course of the experiment (Fig. 3). This indicates that our outbred population harbors alleles

385 at relatively high frequency that exhibit large selection coefficients that flip sign in accordance
386 with fluctuations in environmental conditions over exceptionally short timescales. More
387 specifically, across the time points over which each instance of fluctuating selection was
388 inferred, the environment must have changed in such a way to induce trade-offs in the
389 combination of traits underpinned by these alleles. Thus, even within each of the dominant
390 ecological phases (e.g., population expansion), the adaptive direction of multi-variate
391 phenotypic change likely shifts constantly as changing population densities interact with
392 specific combinations of abiotic conditions. Furthermore, it is likely that those alleles with
393 repeated sign flips exhibit pleiotropic effects, whereby they confer advantageous values for the
394 specific traits most readily impacting fitness during intervals in which they are favored, but
395 induce disadvantageous trait values for the sets of traits subject to selection during intervals in
396 which they are selected against (45–47). As these alleles exhibited undetectable systematic
397 change throughout the experiment, they would likely be inferred to be neutral using inference
398 approaches based on coarse temporal sampling or static snapshots of population genomic
399 data. Thus, our high frequency sampling revealed a functional relevance to standing variation
400 that would have remained hidden, bolstering recent calls for the procurement of well-
401 resolved, time series genomic sampling (9, 10, 12, 13, 31–33).
402 The patterns of fluctuating selection we observed are likely to have appreciable impacts
403 on genome-wide patterns of diversity within the species. Specifically, selection in various
404 forms (e.g. background selection, selective sweeps, fluctuating selection) have been shown to
405 broadly reduce diversity at linked neutral sites (48–59). In the case of temporally varying
406 selection, it has been shown that even modest fluctuations can have substantial impacts on
407 genome-wide patterns of diversity (59). Thus, the strong, rapid, and genome-wide signals of
408 fluctuating selection we quantified are likely to alter patterns of diversity in a pronounced

manner. This holds implications for recent debate as to whether elevated selection in large population-size species can help resolve 'Lewontin's Paradox' - the puzzling observation that neutral genetic variation does not scale with census population size as expected (59–63). While some studies have inferred that selection alone is insufficient to explain the narrow range of polymorphism observed across species (60, 61), the exceptionally strong selection quantified here calls this assertion into question and warrants increased effort towards the procurement of well-resolved time-series genomic data in other systems.

Our results may additionally lend insight to the mechanisms maintaining variation in natural populations. For example, a large body of theoretical models has suggested that temporally fluctuating selection pressures are not likely to maintain balanced polymorphisms for genotypes specialized to different states of the environment, but rather drive the fixation of optimally plastic genotypes that are able to alter their phenotypes to match the shifting conditions (3, 17, 64, 65). Such research has led to the viewpoint that spatially varying selection is the dominant driver of balanced polymorphisms, and thus the maintenance of variation, across species' ranges (18, 66). Yet, more recent models have begun to incorporate various forms of multi-locus dominance, particularly environmentally-induced reversals of dominance, and demonstrated that polymorphism at a large number of loci can be maintained in temporally fluctuating environments without imposing unsustainable genetic load on the population (15, 19, 45, 67). Our observations add credence to these models, demonstrating how rapidly fluctuating selection could limit the fixation of adaptive alleles in a population. But whether it indeed is the key force maintaining variation across species ranges is an outstanding empirical issue.

It is critical to note that our detection of rapidly fluctuating selection was likely contingent upon our strategy for quantifying changes in genomic variation in the system. Specifically, changes in allele frequencies over our shortest sampling intervals likely represent

434 subtle, yet significant, shifts in the relative contribution of specific genotypes to the next cohort
435 of flies. Our detection of such fine-scale shifts hinged upon our highly replicated system and
436 standardized sampling of the instantaneously reproductive population at each collection
437 point. Furthermore, as demonstrated in Figure 2D/E, the resolution of sampling throughout
438 the season dictates our ability to detect instances of fluctuating selection, which are averaged
439 out and remain hidden when evolutionary change is quantified over longer timescales (13, 68,
440 69). We thus conclude that the relatively sparse empirical evidence of fluctuating selection
441 across systems may be due to an ascertainment bias of quantifying evolutionary change on
442 longer timescales (33, 70). Indeed, research into patterns of spatially varying selection has
443 indicated that signals of balancing selection over exceptionally fine spatial scales
444 ('microgeographic' adaptation) can be masked by coarse sampling (71). Thus, this pattern may
445 be recapitulated under instances of temporally varying selection, though the ubiquity of this
446 dynamic across study systems and species remains to be determined.

447 Our findings additionally add to a growing body of research into whether adaptation
448 across independently evolving replicate populations proceeds via predictable shifts in
449 genomic variation. While numerous observations of predictable phenotypic evolution are
450 reported in this and other systems (25, 39), signals of genomic predictability are largely
451 confined to traits with relatively simple genetic bases (e.g., 40–42). Here, we detect a large
452 degree of within-year parallelism and across-year predictability, despite hundreds of
453 putatively unlinked loci driving patterns of adaptive tracking in the system (Fig. 2C).
454 Strikingly, this parallelism and predictability is evident despite substantial variance in starting
455 allele frequencies across replicate cages within and across years (Fig. 2A/B; Fig. 4). This
456 demonstrates the oligogenic architecture of adaptive tracking in the system is resilient to
457 perturbations of background/starting allele frequencies, likely a result of the additivity of the
458 marginal effects of alleles underpinning the evolving traits and the very strong selective

459 pressures in the system (72). The predictability of evolution in this system is underpinned by
460 repeated selection on sensory pathways across study years (Fig. 4D). This finding bolsters the
461 hypothesis that adaptive tracking in the system is driven by high-level sensory pathways that
462 regulate individual perception and plastic responses to the shifting environmental conditions
463 (as originally hypothesized in 33).

464 The repeated signals of selection on SNPs, genes, and phenotypic pathways across
465 study years further suggests that despite variation in specific environmental conditions, there
466 exists a generality to the suite of selective pressures imposed on the populations. We
467 hypothesize this generality is in large part driven by the boom-and-bust ecological dynamics
468 observed in our mesocosm populations, which is ubiquitous to natural populations of
469 *Drosophila* spp. as well as range of species with differing life-history characteristics (20–23).
470 This hypothesis is bolstered by the fact that the most concordant periods of allele frequency
471 change across years were those when the population sizes were in similar phases of the boom-
472 and-bust cycle (Fig. 4D), as well as previous studies reporting concordant allele frequency
473 shifts between the spring and fall over multiple years (31) and across widely dispersed *D.*
474 *melanogaster* populations (32). It is not feasible to disentangle the relative contribution of these
475 biotic forces from covarying abiotic selective pressures (e.g., temperature) with the data
476 collected here, though may be possible in the future through a systematic manipulation of
477 these variables, such as a staggered initiation of the replicate cages and population trajectories.

478 In conclusion, the allele frequency change we observed illustrate an instance of adaptive
479 tracking underpinned by strong, rapidly fluctuating selection that induces parallel allele
480 frequency change across twelve independent replicates. Such dynamic patterns of selection are
481 likely to have pronounced impacts on patterns of genetic diversity within the species and may
482 act to maintain variation across the species' range. Furthermore, given that the observed
483 adaptive tracking is likely driven by continual selection and reshuffling of standing variation,

484 it is not implausible that we could detect evolutionary change over weekly, sub-generational,
485 intervals. Indeed, previous studies have quantified selection in the wild occurring on similar
486 timescales, though these are largely confined to instances in which the evolutionary change
487 was a response to a pulse environmental perturbation, such as hurricanes (73), winter storms
488 (74), harvest and poaching (75), or heat waves (76). While this may be the first instance of
489 sustained evolutionary change at this tempo, the phenomenon may in fact be widespread in
490 populations evolving with overlapping generations and thus revealed by sufficiently resolved
491 time-series genomic data. Finally, the dynamic nature of adaptive tracking observed lends
492 promise towards the capacity for species to adaptively track the changes in the environment
493 induced by global change. Still, an important consideration is how dependent such fine-
494 grained adaptive tracking is on the total amount of genetic variation present in populations,
495 which itself may be reduced as global change progresses (77, 78). Ultimately, resolving the
496 underlying causal genes underpinning adaptive tracking will lend insights into the total
497 number of loci necessary for fine-grained, rapid adaptation in complex selective regimes.

498 **Methods**

499 *Study population and mesocosm system*

500 We quantified and tracked shifts in genomic variation of a genetically diverse *Drosophila*
501 *melanogaster* population over a six-month period in a highly replicated, field mesocosm in
502 Philadelphia, Pennsylvania. The study population was derived via outbreeding a panel of 76
503 inbred strains originally collected wild from Linvilla Orchards, Media, PA (24). Each line had
504 been previously sequenced to a minimum genome-wide coverage of 50x, following 20
505 generations of full-sib mating (as described in (38)). An outbred, baseline population was
506 initiated via combining 10 males and 10 females from each line into each of four, 30 cm^3 indoor
507 cages. Each subpopulation was subject to four generations of unmanipulated recombination
508 and population expansion, a process aimed at purging deleterious alleles that likely

509 accumulated throughout inbreeding. Eggs from the fourth generation were collected and
510 density controlled in vials. Upon eclosure, 500 randomly selected males and females were
511 released into each of twelve two m^3, outdoor mesh cages.

512 Each replicate cage harbored a single dwarf peach tree and was exposed to natural
513 environmental conditions and insect and microbial communities (33, 79). Four hundred ml of
514 Drosophila media ('Spradling cornmeal recipe'), provided in 900 cm^3 aluminum loaf pans, was
515 placed in a shaded shelving unit within each cage three times per week, providing the only
516 source of food and egg laying substrate. Egg laying upon each loaf pan was carried out for two
517 days, after which the pan was enclosed with a mesh lid to prevent any further laying. The pan
518 was then monitored daily, and the lid removed after eclosure was first oberved. This process
519 generated a near continual input of new flies into each replicate cage, generating evolution
520 with overlapping generations and a population age-structure that varied throughout the
521 duration of the experiment. The census size of the adult population was estimated five times
522 throughout the progression of the experiment (19 July, 5 August, 20 August, 17 September,
523 and 20 October). To estimate population sizes, we photographed four, 0.3 m^3 transects on the
524 ceiling of each mesocosm within 30 minutes of sunsets (when adult flies were aggregated on
525 the mesocosm walls). The number of flies within each transect were counted using semi-
526 automated image processing, the average of which was used as the estimated number of adult
527 flies per m^3 for a given cage and time point.

528 *Sample collection, sequencing, and allele frequency estimation*

529 We conducted high-frequency sampling of the cage replicates to quantify patterns of
530 genomic variation throughout the experiment. The first sampling was conducted on 13 July
531 2021, three weeks after the 21 June seeding of the mesocosms (a period aiming to further purge
532 residual deleterious mutations and allow the lab-adapted population to experience 1-2
533 generations of field selection). Samples were collected weekly for the first 9 weeks of the

experiment (13 July – 7 September), after which samples were collected on 21 September, 20 October, and following the first freeze and population crash on 20 December (12 total time points; Fig. 1). Note that egg samples for two replicates (replicate 3 and 7) were lost during the t1 sampling point and eggs for one replicate (replicate 5) were lost during the t3 sampling point. As a consequence, allele frequency data is missing for these samples/time points.

Throughout the experiment, the replicate populations evolved with overlapping generations and thus exhibited shifting age-structure as the experiment progressed (Charlesworth 1980/1994). It was thus critical to account for this demographic structure in quantifying the ecological and genetic state of the system to ensure that our quantification of allele frequencies reflected shifts to the instantaneously reproductive genotypes, or 'leading edge', of the population (37). Accordingly, at each time point eggs were collected overnight directly from the mesocosms, and larval development was carried out in lab-based conditions, where eggs developed and eclosed to F1 adults in 30 cm^3 cages. 3-5 days post-eclosure, a random set of 100 females were sampled from each cage and preserved in 99% ethanol at -20° C for later DNA extraction.

We collected an additional, independent sample of 100 females from each cage at time points 3 (26 July) and 10 (20 October) to quantify biological variation in our allele frequency estimates. Furthermore, an additional four replicate samples of the baseline population prior to field exposure were preserved for genomic analysis. Finally, it is important to note that the final sample point (22 December) occurred following the crash of the adult population, and thus sampling methods for this time point were altered such that larvae were collected directly from the remaining loaf pans within each cage (i.e., no reproductively viable adults remained in the cage at this time point, but larvae persisting in cage loaf pans were taken into the lab and reared to adults for pooled sequencing). Additionally, two replicate mesocosms were unable to be sampled at this final time point, reducing replication. We accommodate for these

559 differences in sampling of t12 by focusing analysis on patterns of variation throughout the
560 summer to fall transition (time points 1 through 11; see below).
561 We generated high-accuracy, haplotype-informed allele frequency estimates at most of
562 the common polymorphic sites within the species. Specifically, genomic DNA was extracted
563 from each homogenized pool of 100 flies using the Monarch Genomic DNA Purification Kit
564 (New England Biolabs). DNA quality was assessed using Nanodrop and quantified with
565 Qubit, before whole-genome sequencing libraries were generated for each pooled sample
566 using the Illumina DNA Prep Tagmentation Kit. All samples were sequenced on Illumina
567 Novaseq 6000 flow cells using 150 bp, paired-end reads. A total of three separate sequencing
568 rounds were required to generate sufficient coverage across all samples, and technical
569 replication across rounds were used to quantify the negligible impact of batch effects on allele
570 frequency estimation (Fig. S1). Raw sequencing reads were trimmed of adapter sequences and
571 bases with quality score < 20, and aligned to the *Drosophila melanogaster* v5.39 reference
572 genome using bwa and default parameters (80). Aligned reads were deduplicated using Picard
573 tools (http://broadinstitute.github.io/picard/) and the final set of reads for each sample was
574 down-sampled to obtain an equivalent, genome-wide coverage of 8x across all samples.
575 We computed haplotype-derived allele frequency estimates using a local inference
576 method and pipeline developed by Tilk et al. (2019). Briefly, using an expectation-
577 maximization algorithm for estimating haplotype frequencies in a pooled sample from
578 mapped sequence reads (originally developed by (81)), we obtained allele frequencies at sites
579 previously identified in the genome sequences of each founding inbred strain. We conducted
580 haplotype inference in window sizes that varied proportionally to the length of un-
581 recombined haplotype blocks expected as a function of the estimated number of generations
582 since the original outbreeding. Allele frequencies were compiled to only include those sites
583 with an average minor allele frequency > 0.02 in the baseline population, and present in at

584 least one evolved sample at a MAF > 0.01. See Tilk et al. 2019 for validation that this method
585 produces an accuracy of allele frequencies at a genome-wide depth of 5x comparable to those
586 obtained from raw allele frequencies generated via standard pool-seq data and 100x coverage.
587 *Quantifying genome-wide patterns of variation*
588 All statistical analysis of allele frequency data was carried out in R v. 3.5.6. We used a
589 population genetic summary statistic, F_{ST}, to determine whether observed shifts in genomic
590 variation exceeded sources of technical and biological noise throughout the sampling period.
591 Specifically, we quantified genomic differentiation through time by computing F_{ST} between
592 each evolved sample and each of the four baseline samples (F_{ST} was computed separately for
593 each polymorphic site and averaged across all sites to obtain a genome-wide mean of
594 differentiation). Next, we quantified sources of biological and technical noise by computing F_{ST}
595 pairwise between technical sequencing replicates (same 100 flies and pooled DNA extract,
596 different library preparation and/or sequencing rounds; N = 20 pairwise comparisons) and
597 biological replicates (same cage/time point, different sample of 100 flies; N = 17 pairwise
598 comparisons). We tested whether evolutionary-driven changes in genomic variation exceeded
599 sources of noise using a paired t-test, and whether evolved samples became increasingly
600 differentiated from the baseline population throughout the experiment using a generalized
601 linear model (family = gaussian; link = logistic).
602 We tested whether variance in allele frequencies across all collected samples was
603 dominated by replicate cage or collection time point using principal component analysis (PCA)
604 (*prcomp* package). Allele frequencies were centered and scaled, and each sample was projected
605 upon the first two principal components. We ran this analysis both genome-wide, as well as
606 separately for SNPs on each chromosomal arm. We conducted K-means clustering of each
607 sample's PC scores (*stats* package) to, in an unsupervised manner, quantify the extent to which
608 PC clustering was predominately driven by replicate cage or collection time. Specifically, if K-

609 means clustering of sample points corresponded to sample collection time point (i.e.,
610 segregated early from late season samples), this provided a quantitative validation of trends
611 visualized along PC1 and PC2.

612 *Target SNP identification and characterization of genetic architecture*

613 We identified single nucleotide polymorphism (SNPs) shifting systematically through
614 time and across replicates by fitting allele frequencies to a generalized linear model (33, 82).
615 Allele frequencies were weighted by the total number of chromosomes sequenced (N = 200)
616 and depth per sample (38). GLM's for each site were run with a quasibinomial error model,
617 and P-values were adjusted using the Benjamini-Hochberg false discovery rate (FDR)
618 correction (*p.adjust* package). We considered a SNP significant if it exhibited an FDR < 0.2 and
619 effect size > 0.5%. We used this approach to identify SNPs shifting systematically throughout
620 the entire summer to fall transition (time points 1 to 11; before total population crash and
621 reduction in mesocosm-wide replication of time point 12). We additionally used this approach
622 to determine the interval length for which we had sufficient power to identify significant,
623 relatively large effect size SNPs (FDR < 0.05 and >2 %) across all chromosomal arms (Fig. S4).
624 Finally, we explored how SNP discovery changed when using intermediate time point data in
625 the model. For example, for regressions of allele frequencies from t1->3, we compared the
626 number of SNPs identified using data from t1, t2, and t3, as well only end time point data (i.e.,
627 just data from t1 and t3) (results in Fig. S4). The increase in SNP discovery using intermediate
628 time point data informed our decision to incorporate intermediate time points in all
629 regressions.

630 We used our GLM results to explore the genetic architecture of adaptive tracking.
631 Specifically, we implemented a method developed and described by Rudman et al. 2022 to
632 identify clusters of unliked loci enriched in significant SNPs. Briefly, this method takes in
633 scores each SNP as follows: 0: FDR > 0.2; 1: FDR < 0.2; 2: FDR < 0.05 and effect size > 2%; 3:

634 FDR < 0.01 and effect size > 2%. The average SNP score is then calculated for sliding windows
635 of 500 SNPs, using a 100 SNP step size. An empirical FDR for each window is generated by
636 comparing observed scores to those generated for permuted windows, whereby SNP site
637 information is randomly shuffled and window scores are re-computed. Windows with an
638 empirical FDR < 0.05 are defined as significantly enriched and overlapping enriched windows
639 are merged into a single, enriched windows. To winnow the list of enriched windows to a set
640 of putatively unlinked (independently segregating windows) for each comparison, the
641 squared correlation coefficient of founder genotypes is calculated between all pairs of
642 significant SNPs within 1 Mb of each other. If the average SNP-pair linkage between two
643 clusters is > 0.03, the clusters are merged. This process proceeds iteratively until all pairs of
644 clusters exhibited average linkage < 0.03 (source code provided at
645 https://github.com/greensii/dros-adaptive-tracking).

646 *Quantifying fluctuating selection using leave-one-out cross validation*

647 We quantified the degree of parallel, neutral, and/or anti-parallel evolutionary change
648 for individual replicate cages by re-implementing our GLM analysis using a leave-one-out
649 cross-validation approach: sets of GLM significant SNPs were identified iteratively in *n* -1
650 'training' cages, and the frequency dynamics of the rising alleles (FDR < 0.2, allele frequency
651 change > 0.5 %) were then quantified in the left-out, 'test', cage. A set of control SNPs were
652 selected for every set of target SNPs, whereby each control SNP was matched on the following:
653 chromosomal arm, +/- 5% baseline population frequency, =/-0.5 cM/MB recombination rate,
654 and inversion status (i.e., inside or outside of an inversion). The distributions of shifts between
655 sets of target and matched control SNPs were then compared using a paired t-test to: 1)
656 determine whether the magnitude of target shifts significantly exceeded (FDR < 0.05)
657 background/neutral allele frequency change and 2) infer whether the direction of non-neutral
658 shifts were in a predominately concordant/parallel direction as that observed in the training

659 cages (positive median allele frequency shift), or exhibited significantly anti-parallel behavior
660 (negative median allele frequency shift). For a given left-out cage and set of target SNPs,
661 changes between parallel and anti-parallel behavior across test intervals was indicative of a
662 fluctuation in the dominate direction of selection. As target SNP behavior was always
663 measured in the left-out cage, inference of fluctuating selection was insensitive to statistical
664 artifacts, such as regression to the mean.

665 We first implemented this GLM and leave-one-out approach using sets of target SNPs
666 generated across the entire summer to fall transition (t1->11). The frequency dynamics of
667 target SNPs within left-out cages were measured from t1->11 (the interval over which they
668 were identified in the training cages), separately for each chromosomal arm and genome-wide.
669 Next, to resolve the pace of adaptive tracking in the system and explore the scale at which
670 signals of fluctuating selection arise, we re-measured allele frequency shifts of t1->11 SNPs in
671 increasingly shorter intervals. Specifically, we measured SNP behavior in all intervals starting
672 at t1 and of decreasing length from t1->10 (11 weeks) through t1->2 (1 week). The sustained
673 deterministic behavior of SNPs across replicates during the t1->2 interval motivated
674 assessment of SNP behavior in each pair of single time points (t1->2, t2->3…t11->12), the first
675 nine of which are single week intervals. This final pair of single time points (t11->12) includes
676 the first observed freezing of the year (Fig. 1), and the total crash of the adult fly population
677 within each replicate.

678 We next aimed to quantify the pervasiveness of fluctuating selection throughout the
679 experimental period. Accordingly, we re-conducted our GLM and leave-one-out approach in
680 all interval lengths greater than five time points (e.g., t1-6…t7-12) (a five time point interval
681 length represented the shortest producing substantial, genome-wide GLM signal across all
682 chromosomal arms; Fig. S4). For each set of test SNPs, we measured their allele frequency
683 shifts across the time points over which they were identified, as well as all single time point

684 comparisons (e.g., for SNPs identified during t1->6, we quantified allele frequency shifts from
685 t1->6, as well as t1->2, t2->3…t11->12). For each set of target SNPs, we excluded sites
686 identified as strongly parallel (FDR < 0.05) throughout the summer to fall transition (t1->11).
687 This approach thus explicitly eliminates those sites most predominately subject to primarily
688 directional selection throughout the experimental period to explore whether finer modalities
689 of environmental change drive systematic shifts in allele frequencies across cages. Finally, we
690 quantified the changes in the dominate direction of selection throughout the experiment for
691 each individual cage by enumerating the number of instances sets of target SNPs switched
692 from being predominately favored (significantly positive median phased allele frequency
693 change) to predominately selected against (significantly positive median phased allele
694 frequency change), or vice versa, across all pairs of single time points.

695 *Quantifying the predictability of independent bouts of adaptive tracking*

696 We leveraged results reported in Rudman et al. (2022) to probe the extent to which
697 independent bouts of adaptive tracking in the system were underpinned by repeated selection
698 on functional pathways and exhibited gene and SNP-level parallelism. The Rudman et al.
699 study took place during the 2014 spring to fall transition, using the same outdoor mesocosm
700 system (with a total of 10 cage replicates cages), and a total of 81 inbred strains, 51 of which
701 were also used in this study (the 25 strains unique to this study were derived from flies
702 collected at the same local Pennsylvania orchard population as overlapping strains). Rudman
703 et al. collected samples for analysis of genomic variation monthly (five total time points,
704 spanning 25 July to 7 November) and used identical methods of haplotype-derived allele
705 frequency estimation, GLM-based identification of sites moving systematically through time,
706 and clustering of enriched regions of genomic variation.

707 We first used PCA to explore the impact of experimental year and collection time on
708 variance in allele frequencies across all samples from studies. While the experiments were

709 initiated on different calendar days and extended different periods of time, we standardized
710 collection time by days since initial founding of the mesocosm replicates. We projected each
711 sample onto the first two principal components to ask – are variance in allele frequencies
712 predominately driven by experimental year and/or collection time? Next, we tested whether,
713 despite inter-annual differences in founder population allele frequencies and environmental
714 conditions, a parallel set of SNPs underpinned seasonal adaptation across years. We compared
715 the allele frequency shifts obtained across all interval lengths and comparisons of this study
716 (e.g., all two time point to eleven time point intervals), to each of the four monthly intervals of
717 the Rudman et al. study (270 total comparisons). For each comparison, we computed Pearson
718 correlations between observed allele frequency shifts (~1.6 M SNPs with overlapping data
719 across studies). Positive correlations indicated that SNPs moved in a concordant manner
720 between the tested 2014 and 2021 intervals, and negative correlations to SNPs moving in
721 predominantly opposing directions between years. To infer whether observed correlations
722 exceeded those expected by chance, we computed a distribution of null correlations by
723 randomly shuffling the allele frequency shift values by SNP for each comparison (SNPs
724 shuffled to sites matched on chromosomal arm, starting frequency, recombination rate, and
725 inversion status). We compared the distribution of observed and shuffled correlations using an
726 F test to determine whether an excess of parallel (and anti-parallel) intervals were observed
727 between years. We re-conducted this analysis separately for each chromosomal arm. Finally, to
728 explore whether those intervals with heightened allele frequency concordance across years
729 aligned with conspicuous shifts in the selective environment, we identified the interval of the
730 current study that was best matched (via highest genome-wide correlation) to each of the four
731 monthly intervals of the 2014 experiment.

732 We next explored whether patterns of genomic variation across both years indicated
733 repeated selection on a shared set of genes and phenotypic pathways. Using the clustering

734 method described above, Rudman et al. identified 165 unlinked loci across all assessed
735 intervals (each monthly interval and the first to final time point). We compared these loci to
736 the set of loci procured throughout the summer to fall transition of this study (t1->11,
737 described above), as well clustering of GLM results from those intervals exhibiting the highest
738 Pearson correlations with each of the Rudman et al. monthly intervals. For each locus, we
739 identified the top SNP (lowest P-value) and its associated gene and then compared the number
740 of overlapping genes across studies. We computed the probability of the observed overlap by
741 generating a null distribution of the expected overlap based on random sampling (N = 5000
742 permutations). Finally, to infer whether the phenotypic targets of selection were shared across
743 years, we used the top cluster genes to conduct the same enrichment analysis as Rudman et al.
744 (modPhEA; Weng and Liao, 2017) and assessed whether overrepresented phenotypic classes
745 were concordant across studies.
746
747 **Acknowledgements:** We are grateful to members of the Petrov and Schmidt labs for
748 discussion during experimental design, lab work, and data analysis. We acknowledge
749 Marianthi Karageorgi and Katie Solari for support with sample sequencing. We thank Jess
750 Rhodes and Sharon Greenblum for feedback on statistical analysis and computing. We
751 additionally thank Moi Exposito-Alonso, Hunter Fraser, Olivia Ghosh, Sharon Greenblum,
752 James Hemker, Marianthi Karageorgi, Bernard Kim, Jess Rhodes, Katie Solari, and Sophie
753 Walton for insightful comments on drafts of this manuscript.
754 **Funding:** We are grateful to our funding organizations: the National Science Foundation (NSF
755 PRFB 2109407 to M.C.B.) and the National Institutes of Health (NIH R35GM118165 to D.A.P.
756 and NIH R01GM137430 to P.S.)
757 **Author Contributions**: The experimental was conceived by M.C.B., D.A.P, and P.S. The
758 mesocosm study was conducted by M.C.B., S.B., H.O., and P.S. Population census size

759 estimates were conducted by A.H. DNA extraction, library preparation, and sequencing was
760 conducted by M.C.B. Bioinformatic analysis and statistical analysis of allele frequency data
761 was carried out by M.C.B and D.A.P. M.C.B., D.A.P, and P.S. wrote the manuscript, with
762 feedback from all authors.

763

764

765 **References**

766 1. G. Bell, Fluctuating selection: the perpetual renewal of adaptation in variable
767 environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 87–97
768 (2010).

769 2. M. C. Bitter, *et al.*, Fluctuating selection and global change: a synthesis and review on
770 disentangling the roles of climate amplitude, predictability and novelty. *Proceedings of the*
771 *Royal Society B: Biological Sciences* **288**, 20210727 (2021).

772 3. C. A. Botero, F. J. Weissing, J. Wright, D. R. Rubenstein, Evolutionary tipping points in the
773 capacity to adapt to environmental change. *Proc Natl Acad Sci USA* **112**, 184–189 (2015).

774 4. N. A. Moran, The Evolutionary Maintenance of Alternative Phenotypes. *The American*
775 *Naturalist* **139**, 971–989 (1992).

776 5. S. M. Scheiner, Genetics and Evolution of Phenotypic Plasticity. *Annu. Rev. Ecol. Syst.* **24**,
777 35–68 (1993).

778 6. R. Bürger, M. Lynch, Evolution and Extinction in a Changing Environment: A
779 Quantitative-Genetic Analysis. *Evolution* **49**, 151–163 (1995).

780 7. L.-M. Chevin, R. Lande, G. M. Mace, Adaptation, Plasticity, and Extinction in a Changing
781 Environment: Towards a Predictive Theory. *PLOS Biology* **8**, e1000357 (2010).

782 8. A. A. Hoffmann, C. M. Sgrò, Climate change and evolutionary adaptation. *Nature* **470**, 479–
783 485 (2011).

784 9. V. Buffalo, G. Coop, Estimating the genome-wide contribution of selection to temporal
785 allele frequency change. *PNAS* **117**, 20672–20680 (2020).

786 10. R. D. Clark, *et al.*, The practice and promise of temporal genomics for measuring
787 evolutionary responses to global change. *Molecular Ecology Resources* (2023).

788 11. M. Dehasque, *et al.*, Inference of natural selection from ancient DNA. *Evolution Letters* **4**, 94–
789 108 (2020).

12. J. C. Habel, M. Husemann, A. Finger, P. D. Danley, F. E. Zachos, The relevance of time series in molecular ecology and conservation biology. *Biological Reviews* **89**, 484–492 (2014).

13. M. Lynch, W. Wei, Z. Ye, M. E. Pfrender, The Genome-wide Signature of Short-term Temporal Selection. 2023.04.28.538790 (2023).

14. N. H. Barton, M. Turelli, Evolutionary quantitative genetics: how little do we know? *Annu Rev Genet* **23**, 337–370 (1989).

15. J. Bertram, J. Masel, Different mechanisms drive the maintenance of polymorphism at loci subject to strong versus weak fluctuating selection. *Evolution* **73**, 883–896 (2019).

16. P. W. Hedrick, Genetic Polymorphism in Heterogeneous Environments: The Age of Genomics. *Annual Review of Ecology, Evolution, and Systematics* **37**, 67–93 (2006).

17. P. W. Hedrick, Genetic variation in a heterogenous environment. II. Temporal heterogeneity and directional selection. *Genetics* **84**, 145–157 (1976).

18. P. W. Messer, S. P. Ellner, N. G. Hairston, Can Population Genetics Adapt to Rapid Evolution? *Trends in Genetics* **32**, 408–418 (2016).

19. M. J. Wittmann, A. O. Bergland, M. W. Feldman, P. S. Schmidt, D. A. Petrov, Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. *Proceedings of the National Academy of Sciences* **114**, E9932–E9941 (2017).

20. J. M. Gleason, P. R. Roy, E. R. Everman, T. C. Gleason, T. J. Morgan, Phenology of Drosophila species across a temperate growing season and implications for behavior. *PLOS ONE* **14**, e0216601 (2019).

21. G. K. H. Boor, C. B. Schultz, E. E. Crone, W. F. Morris, Mechanism matters: the cause of fluctuations in boom—bust populations governs optimal habitat restoration strategy. *Ecological Applications* **28**, 356–372 (2018).

22. B. E. Kendall, *et al.*, Why Do Populations Cycle? A Synthesis of Statistical and Mechanistic Modeling Approaches. *Ecology* **80**, 1789–1805 (1999).

23. P. Turchin, A. D. Taylor, Complex Dynamics in Ecological Time Series. *Ecology* **73**, 289–305 (1992).

24. E. L. Behrman, S. S. Watson, K. R. O'Brien, M. S. Heschel, P. S. Schmidt, Seasonal variation in life history traits in two Drosophila species. *Journal of Evolutionary Biology* **28**, 1691–1704 (2015).

25. E. L. Behrman, P. Schmidt, How predictable is rapid evolution? 2022.10.27.514123 (2022).

26. P. S. Schmidt, D. R. Conde, Environmental Heterogeneity and the Maintenance of Genetic Variation for Reproductive Diapause in *Drosophila melanogaster*. *Evolution* **60**, 1602–1611 (2006).

27. H. T. Band, P. T. Ives, Correlated changes in environment and lethal frequency in a natural population of Drosophila melanogaster. *Proc Natl Acad Sci U S A* **47**, 180–185 (1961).

28. T. Dobzhansky, Genetics of natural populations. XVI. Altitudinal and seasonal changes produced by natural selection in certain populations of *Drosophila pseudoobscura and Drosophila persimilis*. *Genetics* **33**, 158–176 (1948).

29. T. Dobzhansky, Genetics of natural populations IX. Temporal changes in the composition of populations of *Drosophila pseudoobscura. Genetics* **28**, 162–186 (1943).

30. P. T. Ives, Further genetic studies of the south Amherst population of *Drosophila melanogaster*. *Evolution* **24**, 507–518 (1970).

31. A. O. Bergland, E. L. Behrman, K. R. O'Brien, P. S. Schmidt, D. A. Petrov, Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in Drosophila. *PLOS Genetics* **10**, e1004775 (2014).

32. H. E. Machado, *et al.*, Broad geographic sampling reveals the shared basis and environmental correlates of seasonal adaptation in Drosophila. *eLife* **10**, e67577 (2021).

33. S. M. Rudman, *et al.*, Direct observation of adaptive tracking on ecological time scales in Drosophila. *Science* **375**, eabj7484 (2022).

34. M. Lässig, V. Mustonen, A. M. Walczak, Predicting evolution. *Nature Ecology & Evolution* **1**, 0077 (2017).

35. P. Nosil, *et al.*, Ecology shapes epistasis in a genotype–phenotype–fitness map for stick insect colour. *Nature Ecology & Evolution*, 1–12 (2020).

36. M. T. Wortel, *et al.*, Towards evolutionary predictions: Current promises and challenges. *Evol Appl* **16**, 3–21 (2022).

37. B. Charlesworth, *Evolution in Age-Structured Populations*, 2nd Ed. (Cambridge University Press, 1994) https:/doi.org/10.1017/CBO9780511525711 (July 24, 2023).

38. S. Tilk, *et al.*, Accurate Allele Frequencies from Ultra-low Coverage Pool-Seq Samples in Evolve-and-Resequence Experiments. *G3 (Bethesda)* **9**, 4159–4168 (2019).

39. D. I. Bolnick, R. D. H. Barrett, K. B. Oke, D. J. Rennison, Y. E. Stuart, (Non)Parallel Evolution. *Annual Review of Ecology, Evolution, and Systematics* **49**, 303–330 (2018).

40. R. D. H. Barrett, *et al.*, Linking a mutation to survival in wild mice. *Science* **363**, 499–504 (2019).

41. P. F. Colosimo, *et al.*, Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. *Science* **307**, 1928–1933 (2005).

42. H. E. Hoekstra, R. J. Hirschmann, R. A. Bundey, P. A. Insel, J. P. Crossland, A Single Amino Acid Mutation Contributes to Adaptive Beach Mouse Color Pattern. *Science* **313**, 101–104 (2006).

43. S. Yeaman, *et al.*, Convergent local adaptation to climate in distantly related conifers. *Science* **353**, 1431–1433 (2016).

44. M.-P. Weng, B.-Y. Liao, modPhEA: model organism Phenotype Enrichment Analysis of eukaryotic gene sets. *Bioinformatics* **33**, 3505–3507 (2017).

45. T. Connallon, S. F. Chenoweth, Dominance reversals and the maintenance of genetic variation for fitness. *PLOS Biology* **17**, e3000118 (2019).

46. H. Levene, Genetic Equilibrium When More Than One Ecological Niche is Available. *The American Naturalist* **87**, 331–333 (1953).

47. M. Turelli, N. H. Barton, Polygenic Variation Maintained by Balancing Selection: Pleiotropy, Sex-Dependent Allelic Effects and G × E Interactions. *Genetics* **166**, 1053–1079 (2004).

48. M. Aguade, N. Miyashita, C. H. Langley, Reduced variation in the yellow-achaete-scute region in natural populations of Drosophila melanogaster. *Genetics* **122**, 607–615 (1989).

49. N. H. Barton, Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **355**, 1553–1562 (2000).

50. D. J. Begun, C. F. Aquadro, Levels of naturally occurring DNA polymorphism correlate with recombination rates in D. melanogaster. *Nature* **356**, 519–520 (1992).

51. J. J. Cai, J. M. Macpherson, G. Sella, D. A. Petrov, Pervasive Hitchhiking at Coding and Regulatory Sites in Humans. *PLOS Genetics* **5**, e1000336 (2009).

52. B. Charlesworth, Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* **10**, 195–205 (2009).

53. B. Charlesworth, M. T. Morgan, D. Charlesworth, The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303 (1993).

54. B. Charlesworth, J. D. Jensen, Effects of Selection at Linked Sites on Patterns of Genetic Variability. *Annu Rev Ecol Evol Syst* **52**, 177–197 (2021).

55. E. Elyashiv, *et al.*, A Genomic Map of the Effects of Linked Selection in Drosophila. *PLOS Genetics* **12**, e1006130 (2016).

56. J. H. Gillespie, Junk ain't what junk does: neutral alleles in a selected context. *Gene* **205**, 291–299 (1997).

57. J. M. Macpherson, G. Sella, J. C. Davis, D. A. Petrov, Genomewide Spatial Correspondence Between Nonsynonymous Divergence and Neutral Polymorphism Reveals Extensive Adaptation in Drosophila. *Genetics* **177**, 2083–2099 (2007).

58. Y. Park, Y. Kim, Partial protection from cyclical selection generates a high level of polymorphism at multiple non-neutral sites. *Evolution* **73**, 1564–1577 (2019).

59. M. J. Wittmann, S. Mousset, J. Hermisson, Modeling the genetic footprint of fluctuating balancing selection: From the local to the genomic scale. 2022.07.15.500223 (2022).

60. V. Buffalo, Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's Paradox. *eLife* **10**, e67509 (2021).

61. G. Coop, Does linked selection explain the narrow range of genetic diversity across species? 042598 (2016).

62. R. B. Corbett-Detig, D. L. Hartl, T. B. Sackton, Natural Selection Constrains Neutral Diversity across A Wide Range of Species. *PLOS Biology* **13**, e1002112 (2015).

63. R. C. Lewontin, *The Genetic Basis of Evolutionary Change* (Columbia University Press, 1974).

64. Levins, Richard, *Evolution in Changing Environments* (Princeton University Press, 1968) (September 15, 2023).

65. S. Via, R. Lande, Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity. *Evolution* **39**, 505–522 (1985).

66. L. F. Delph, J. K. Kelly, On the importance of balancing selection in plants. *New Phytologist* **201**, 45–56 (2014).

67. K. Grieshop, E. K. H. Ho, K. R. Kasimatis, Dominance reversals and the maintenance of genetic variation (2021) https:/doi.org/10.48550/arXiv.2109.01571 (October 11, 2023).

68. J. K. Kelly, The genomic scale of fluctuating selection in a natural plant population. *Evolution Letters* **n/a** (2022).

69. J. T. Stroud, M. P. Moore, R. B. Langerhans, J. B. Losos, Fluctuating selection maintains distinct species phenotypes in an ecological community in the wild. *Proceedings of the National Academy of Sciences* **120**, e2222071120 (2023).

70. P. D. Gingerich, Rates of Evolution: Effects of Time and Temporal Scaling. *Science* **222**, 159–161 (1983).

71. J. L. Richardson, M. C. Urban, D. I. Bolnick, D. K. Skelly, Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution* **29**, 165–176 (2014).

72. I. Höllinger, B. Wölfl, J. Hermisson, A theory of oligogenic adaptation of a quantitative trait. *Genetics* **225**, iyad139 (2023).

73. C. M. Donihue, *et al.*, Hurricane-induced selection on the morphology of an island lizard. *Nature* **560**, 88–91 (2018).

74. S. C. Campbell-Staton, *et al.*, Winter storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard. *Science* **357**, 495–498 (2017).

75. S. C. Campbell-Staton, *et al.*, Ivory poaching and the rapid evolution of tusklessness in African elephants. *Science* **374**, 483–487 (2021).

76. F. Rodríguez-Trelles, R. Tarrío, M. Santos, Genome-wide evolutionary response to a heat wave in Drosophila. *Biol Lett* **9**, 20130228 (2013).

77. M. Exposito-Alonso, *et al.*, Genetic diversity loss in the Anthropocene. *Science* **377**, 1431–1435 (2022).

78. S. U. Pauls, C. Nowak, M. Bálint, M. Pfenninger, The impact of global climate change on genetic diversity within populations and species. *Molecular Ecology* **22**, 925–946 (2013).

79. S. M. Rudman, *et al.*, Microbiome composition shapes rapid genomic adaptation of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **116**, 20025–20032 (2019).

80. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

81. D. Kessner, T. L. Turner, J. Novembre, Maximum Likelihood Estimation of Frequencies of Known Haplotypes from Pooled Sequence Data. *Molecular Biology and Evolution* **30**, 1145–1158 (2013).

82. R. A. W. Wiberg, O. E. Gaggiotti, M. B. Morrissey, M. G. Ritchie, Identifying consistent allele frequency differences in studies of stratified populations. *Methods in Ecology and Evolution* **8**, 1899–1909 (2017).

Supplementary Material

Bitter *et al.* (2023)

Supplementary Figures

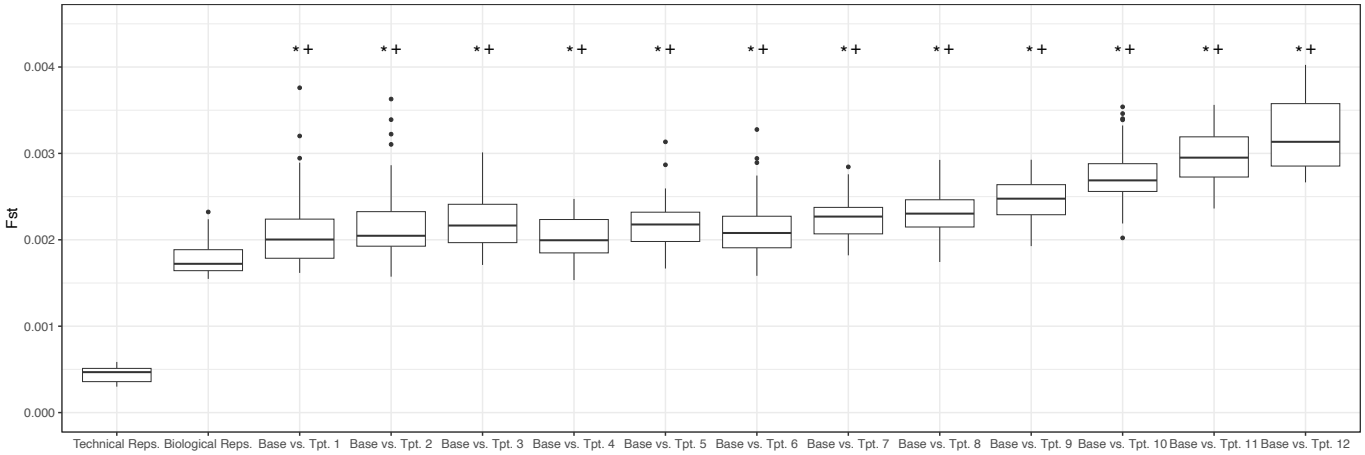


Figure S1 Genome-wide differentiation throughout experiment | Distribution of genome-wide Fst computed pairwise between all technical replicate samples (same DNA extraction, different library preparation and sequencing lane; N = 10), biological replicates (same time point and replicate cage, different pool of 100 flies; N = 17), and samples collected at each time point (N = 12) and the baseline population samples ('Base', N = 4), for each of the twelve time points. The distributions of genome-wide F_{ST} between time point samples and the baseline population was significantly greater than that observed between technical replicates (*) and biological replicates (+) (paired T test; P-value < 0.01). F_{ST} between time point samples and the baseline population increased throughout the course of the experiment (generalized linear model; P = 0.039).

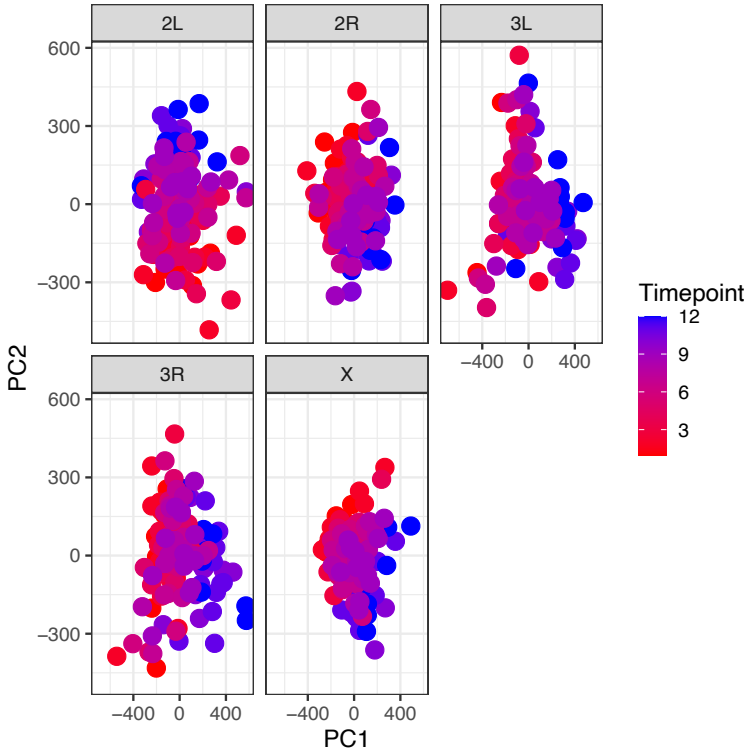


Figure S2 Principal component analysis by chromosomal arm | Principal component analysis of normalized allele frequency data across all replicates and time points, run separately on SNPs of each chromosomal arm (2L:

453,445 SNPs; 2R: 373,755 SNPs; 3L: 445,457 SNPs; 3R: 448,008 SNPs; X: 254,497. Sample scores are projected onto the first two principal components and colored by collection time point. Variance explained for PC1 and PC2 for each chromosomal arm were as follows: 2L: 7.12% and 6.33%; 2R: 5.91% and 5.04%; 3L: 7.75% and 5.60%; 3R: 7.42% and 5.82%; X: 6.71% and 5.29%.

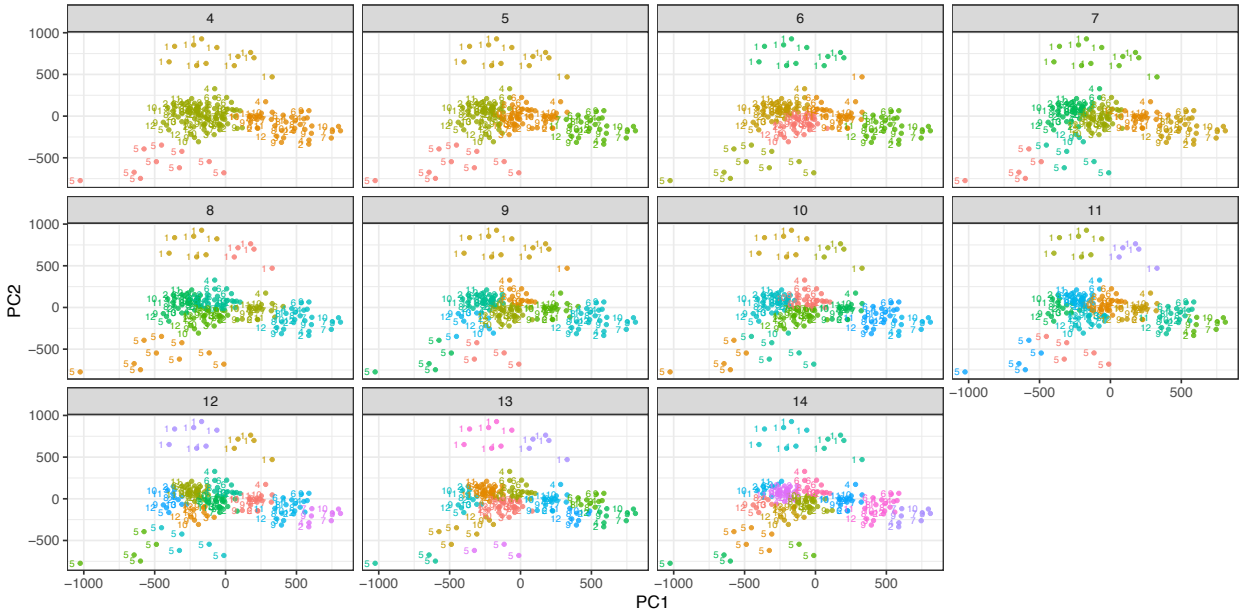


Figure S3 K-means clustering of PCA | Unsupervised K-means clustering was conducted using each sample's PC 1 and 2 scores (main text Figure 2A/B). Depicted here are results for K4-14 in which individual samples are colored by cluster membership and labelled by replicate cage.

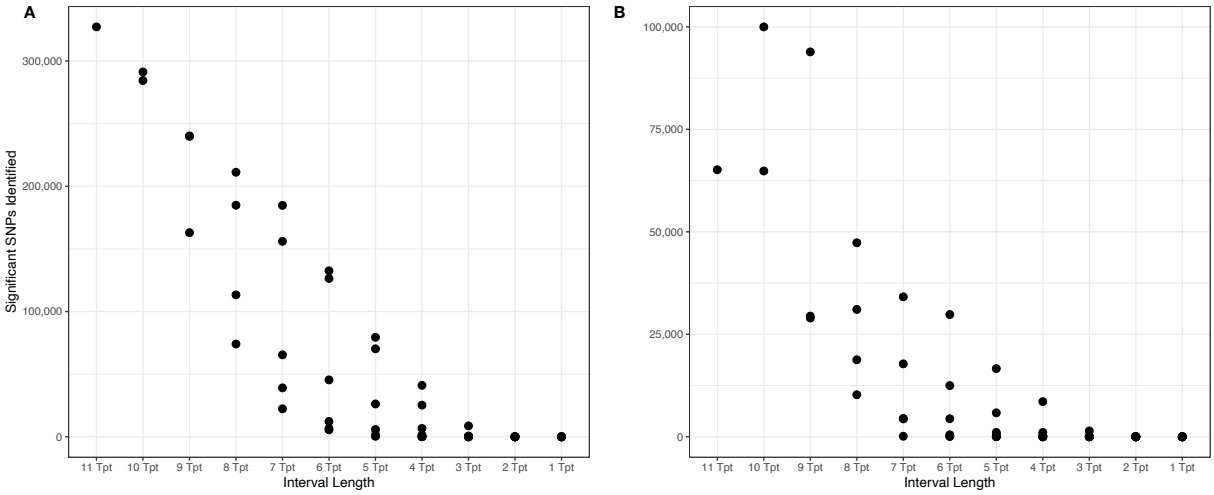


Fig. S4 SNP discovery as a function of GLM interval length | The number of SNPs with Benjamini-Hochberg FDR < 0.05 and effect size > 1% identified via regression of allele frequencies through time using all possible interval lengths. For each interval length, all possible time point comparisons were considered (each point represents a separate comparison). For example, nine time point intervals include regression of allele frequencies from t1->10, t2->11, and t3->12. (A) Depicts results for GLM's using intermediary time points within each comparison in the regression (e.g., for t1->3, model included allele frequency data from t1, t2, and t3). (B) Depicts results for GLM's in which only allele frequency data from the first and last time point of the comparison were considered in the model (e.g., for t1->3, model included allele frequency data from only t1 and t3).

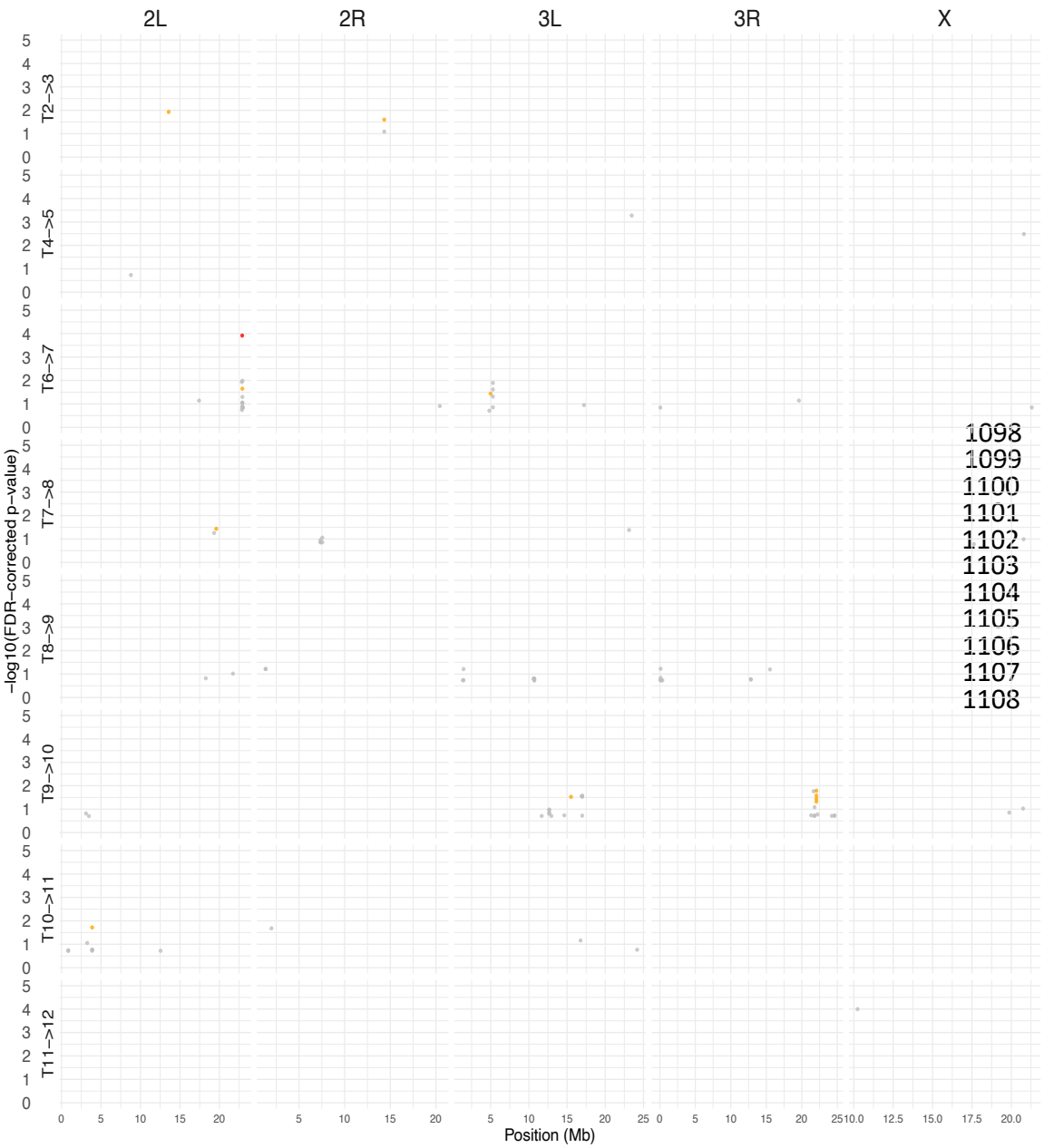


Figure S5 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all single pairs of time points | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%. Note t1->2 not shown as no significant SNPs were identified in that interval.

Figure S6 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all two time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

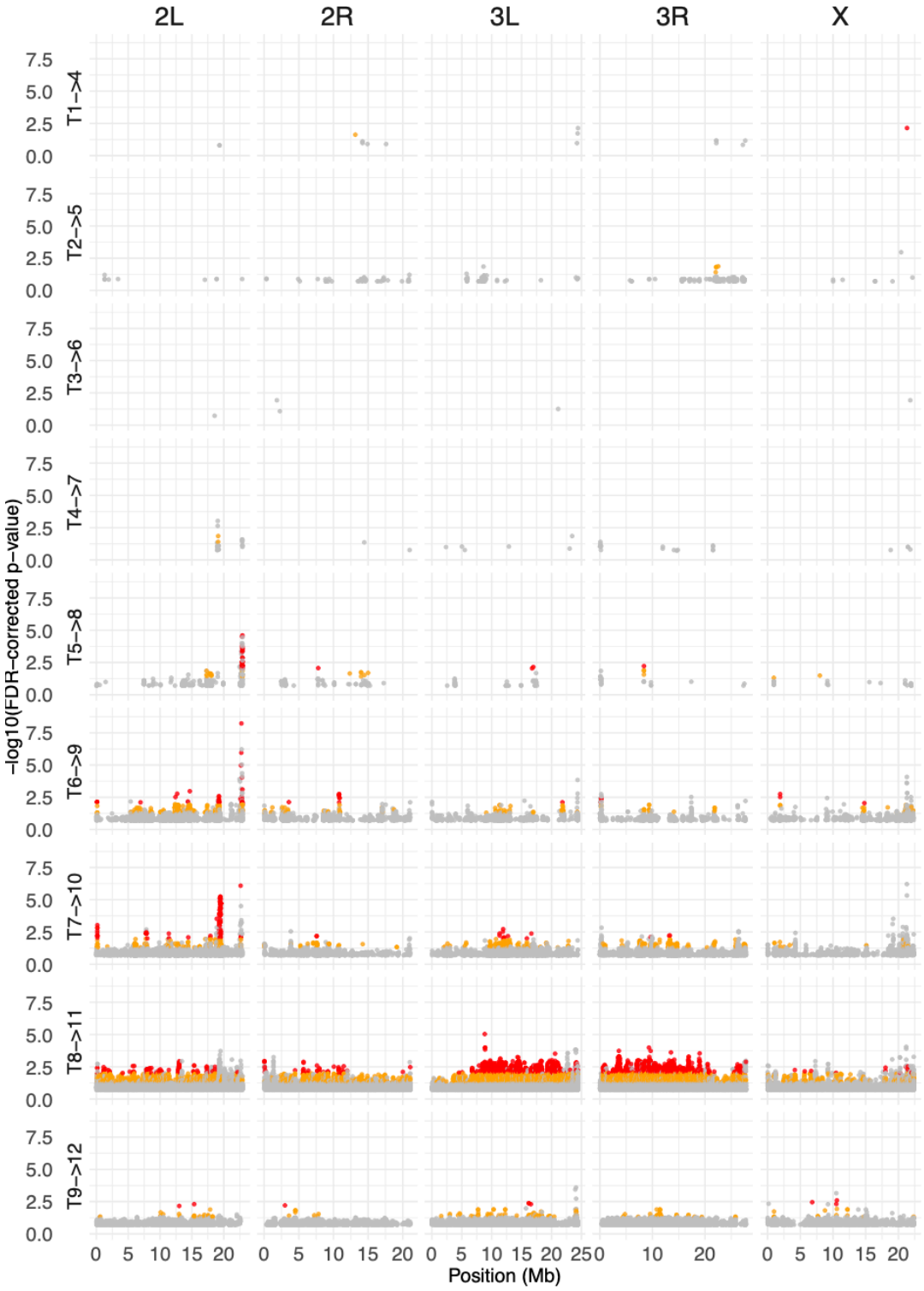


Figure S7 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all three time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

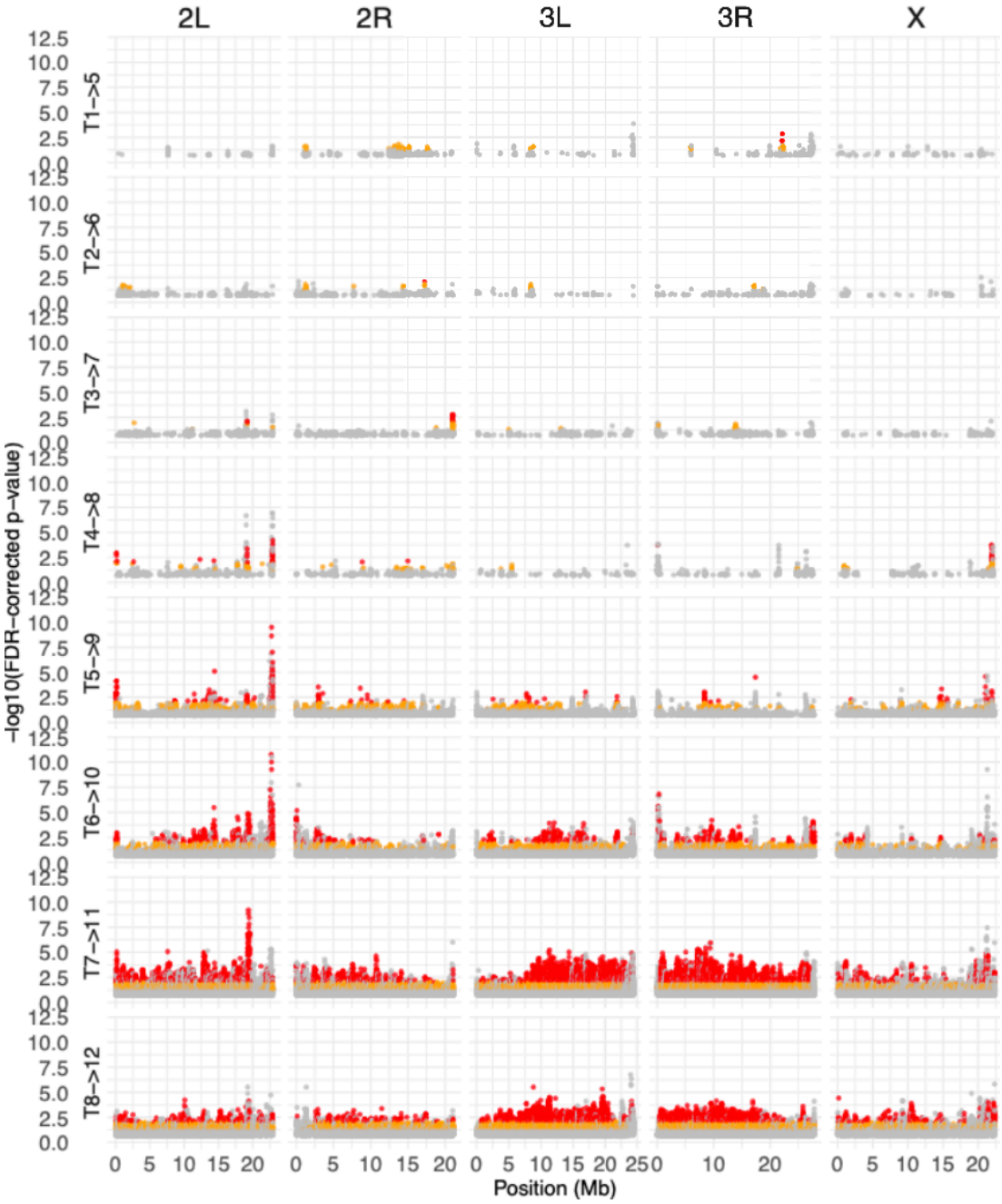
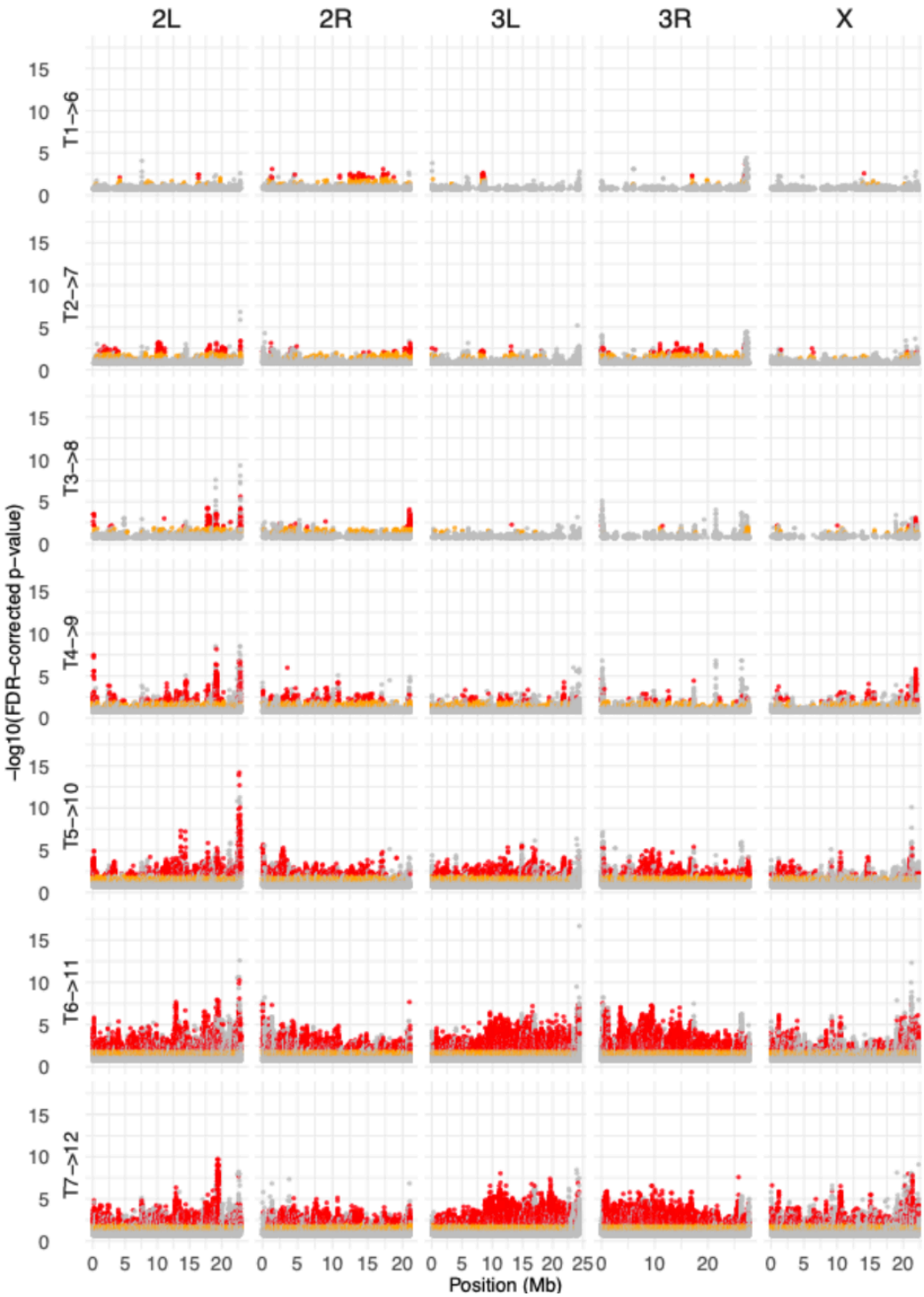


Figure S8 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all four time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

1312



1363
1364 **Figure S9 Genomic distribution of SNPs identified as significant using a generalized linear model and**
1365 **regression of allele frequencies across all five time point intervals** | SNPs (points) are plotted as a function of
1366 chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-
1367 corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%;
1368 red = FDR < 0.01 and effect size > 2%.
1369

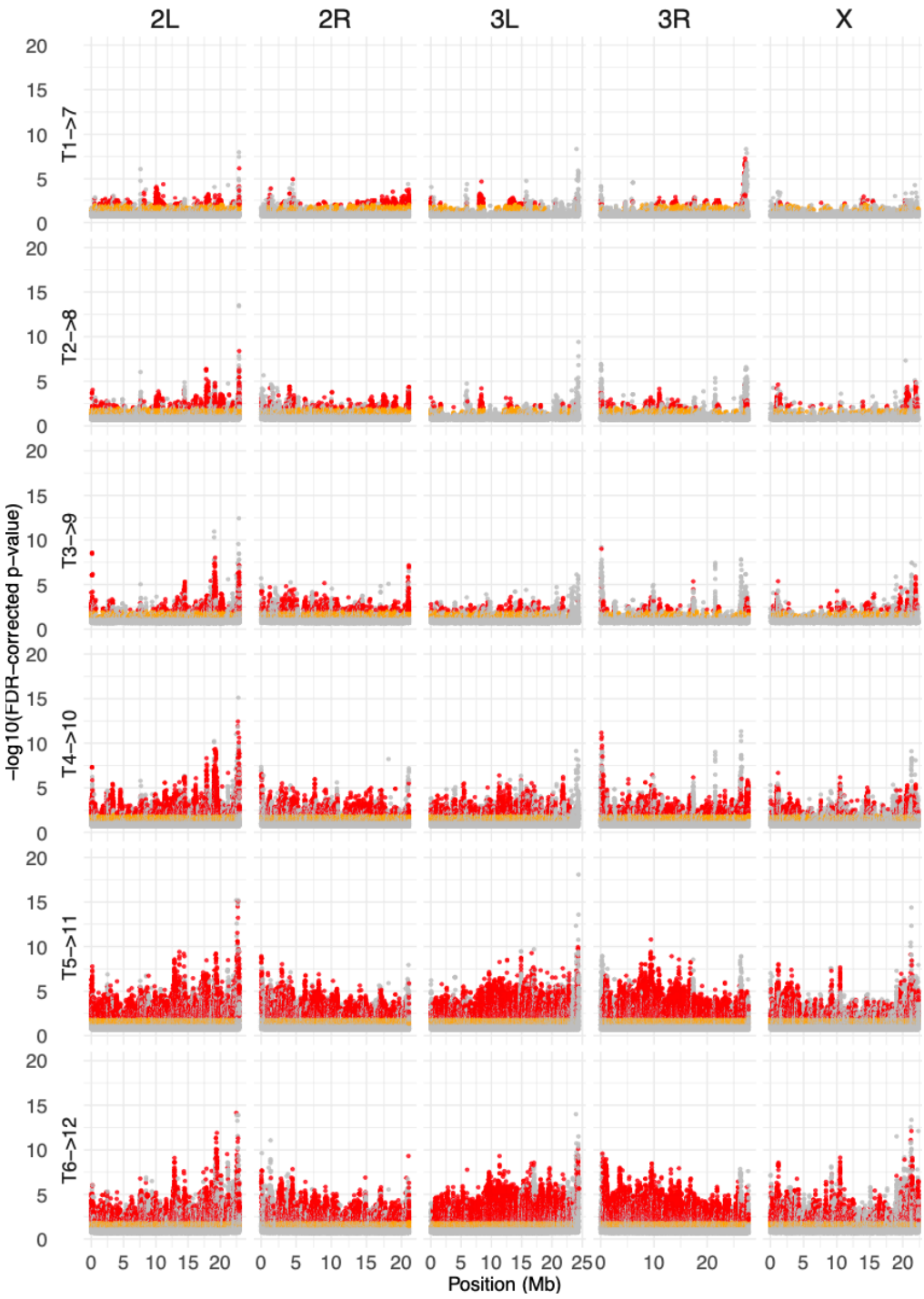


Figure S10 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all six time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

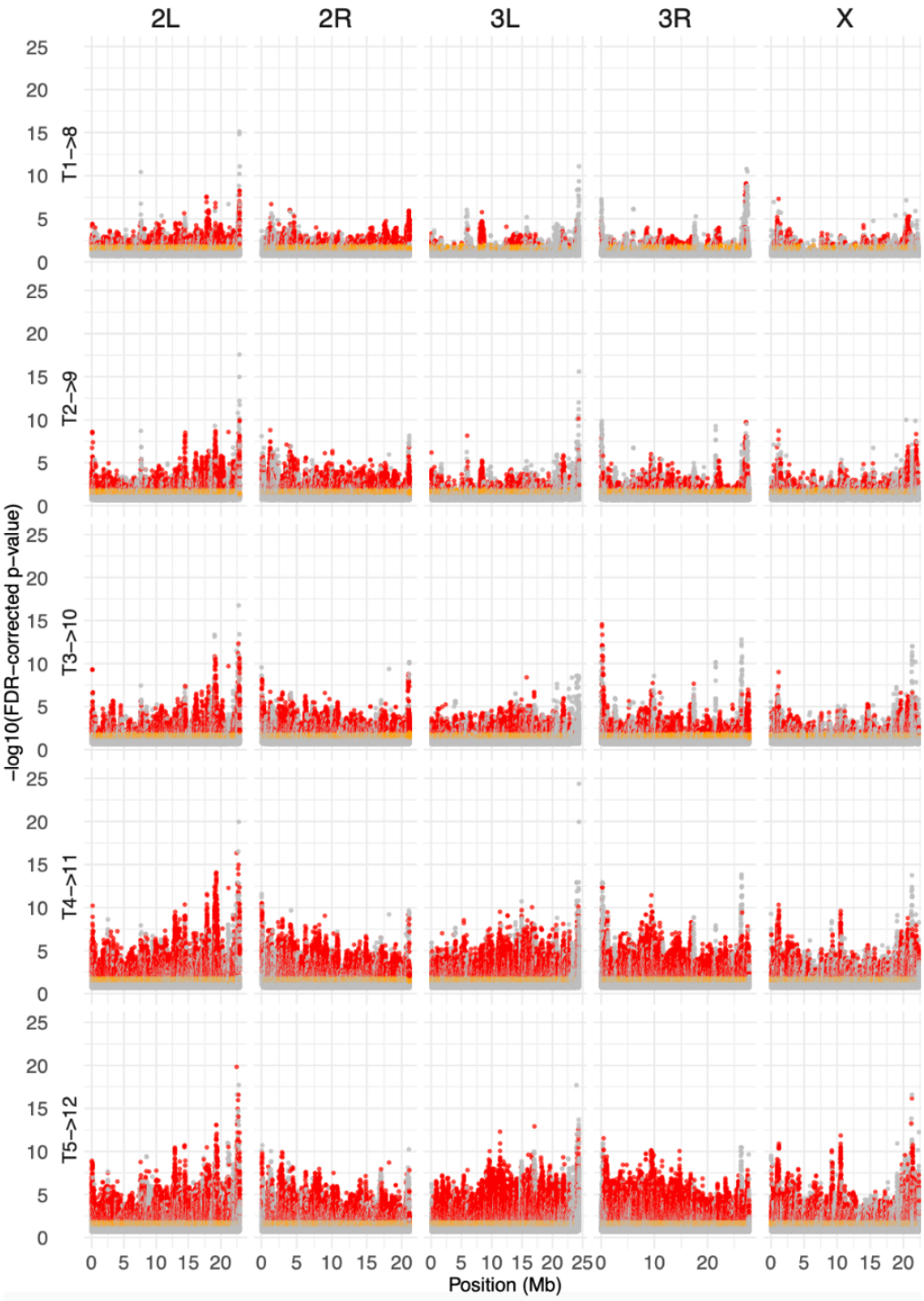


Figure S11 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all seven time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

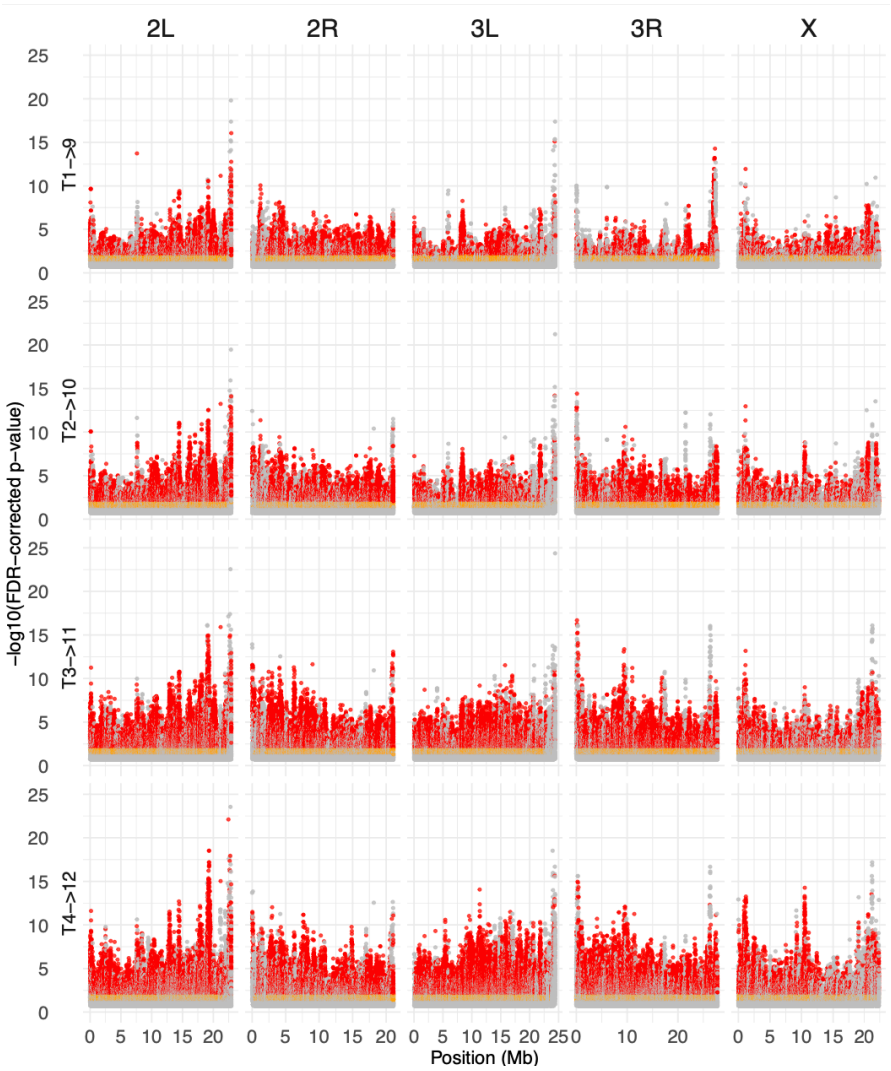


Figure S12 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all eight time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

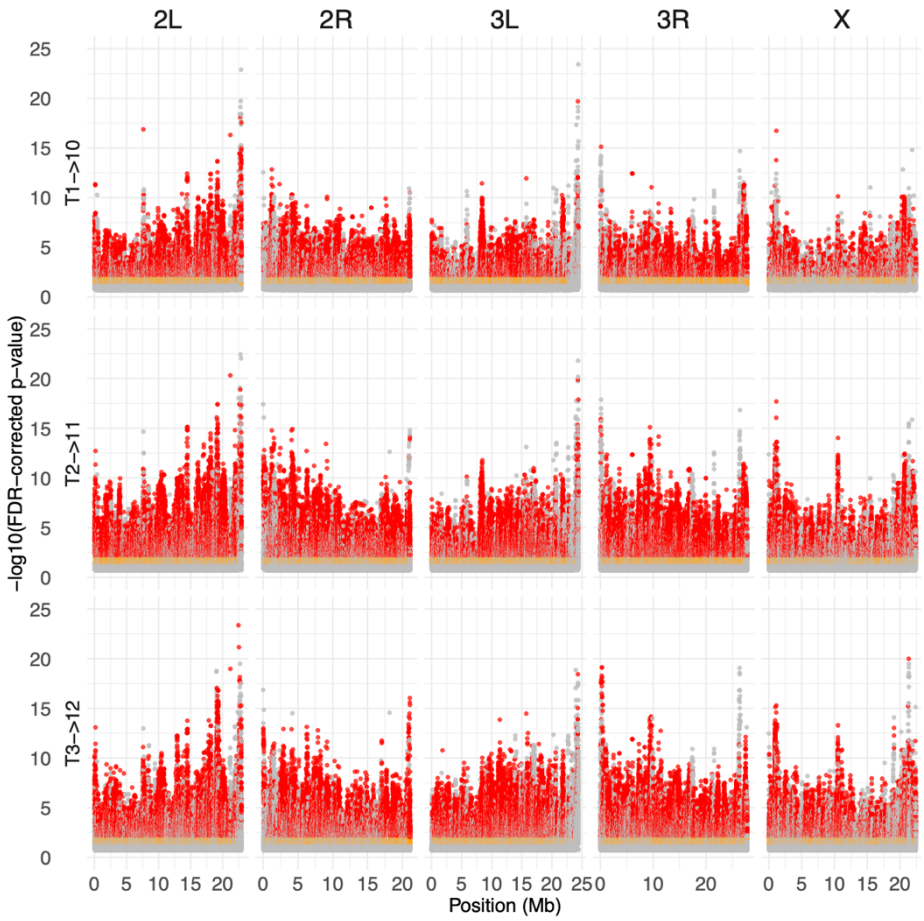


Figure S13 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all nine time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

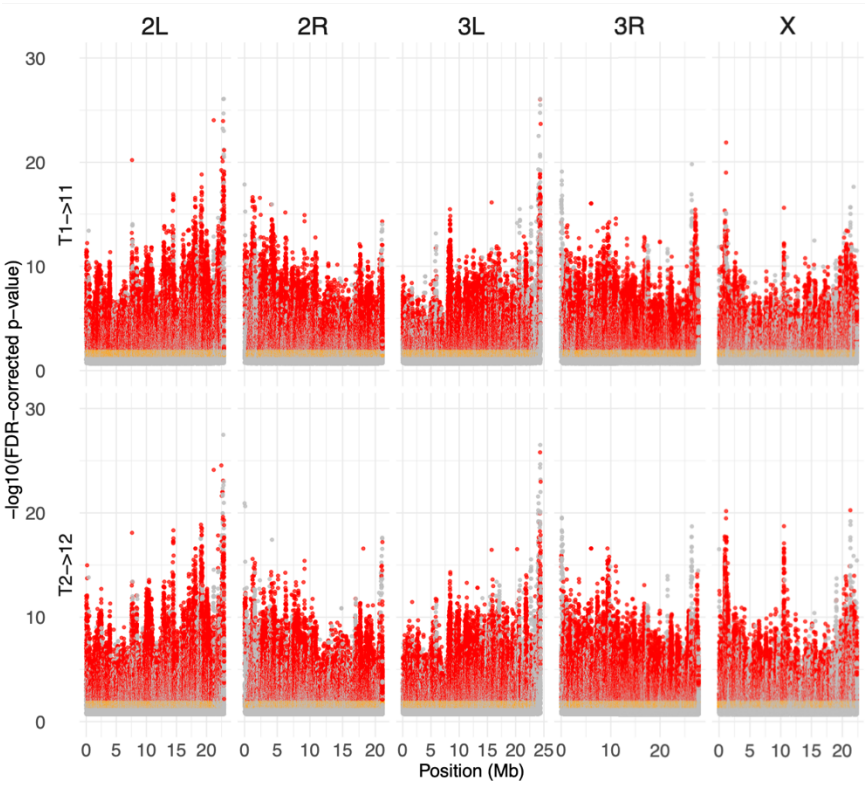


Figure S14 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all ten time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

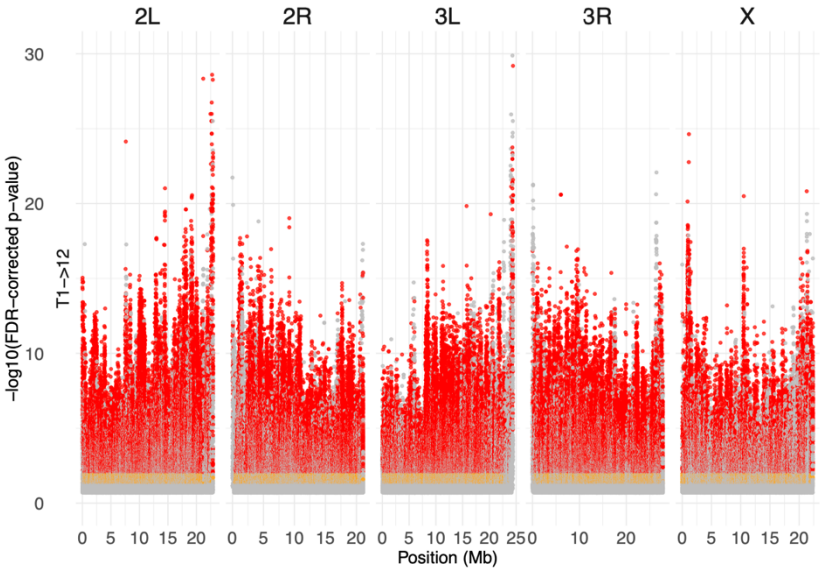


Figure S15 Genomic distribution of SNPs identified as significant using a generalized linear model and regression of allele frequencies across all ten time point intervals | SNPs (points) are plotted as a function of chromosomal location and Benjamini-Hochberg FDR-corrected P value. SNPs are colored according to FDR-corrected P value and effect size: grey = FDR < 0.2, effect size > 0.05%; orange = FDR < 0.05 and effect size > 2%; red = FDR < 0.01 and effect size > 2%.

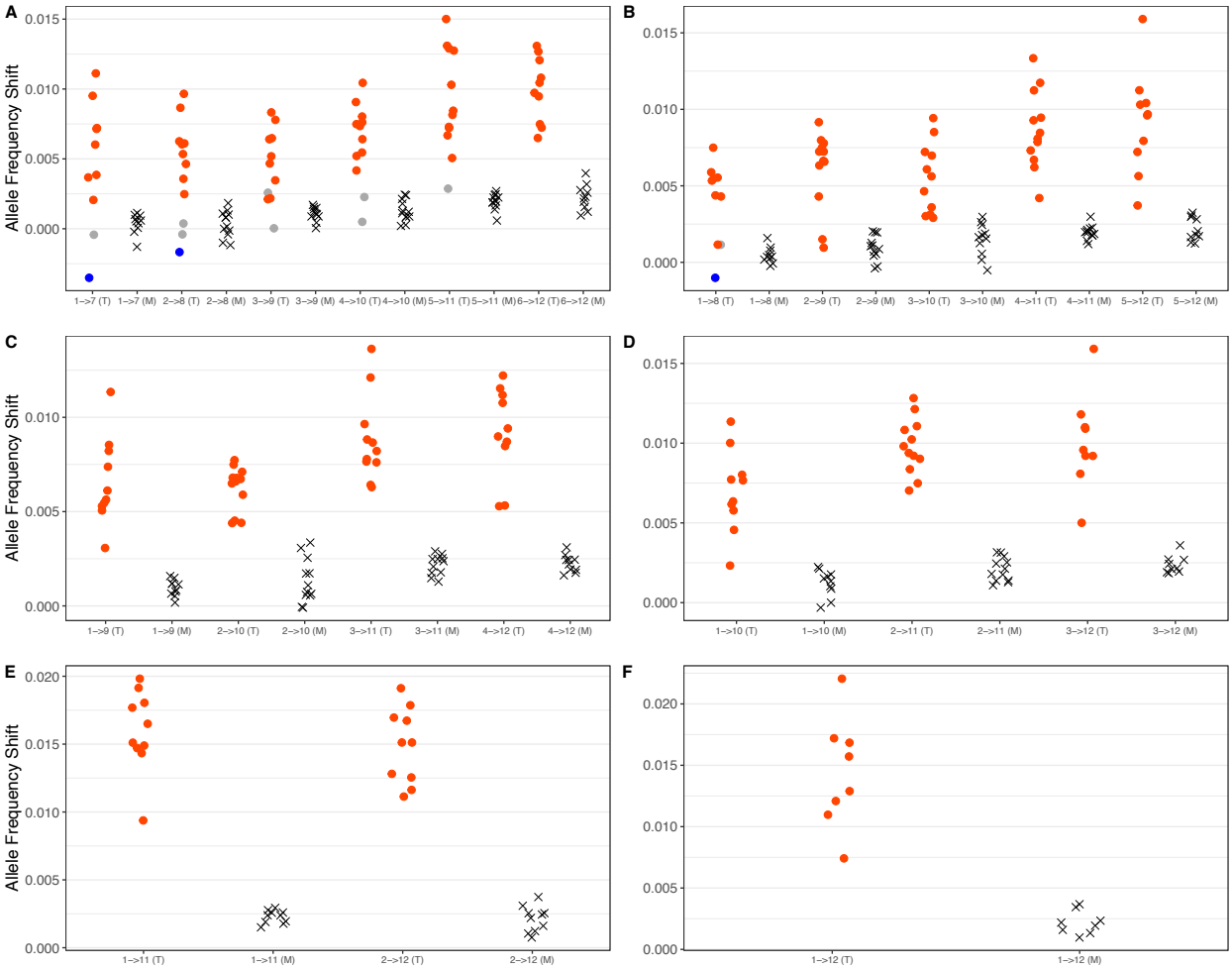


Figure S15 | Median allele frequency shifts for each replicate at target (T) SNPs (circles) and matched (M) controls (M, X's) identified using a GLM and leave-one-out approach for all six (A), seven (B), eight (C), nine (D), ten (E), and eleven (F) time point intervals. Allele frequency shifts were measured, and are plotted, across the interval over which sets of test SNPs were identified. Median shifts at target sites are colored red or blue if the distribution of phased allele frequency shifts was significantly greater or less than that for matched control sites, respectively.

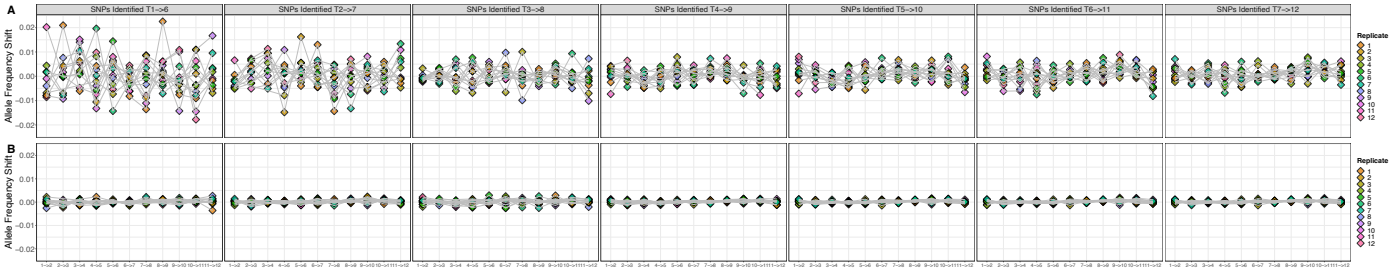


Figure S16 | Median shifts at target SNPs (A) and matched control SNPs (B) identified in each possible five time point interval and measured across each pair of single time points. Point color corresponds to replicate cage ID's and grey lines connect the median shifts for each cage between consecutive time point intervals.

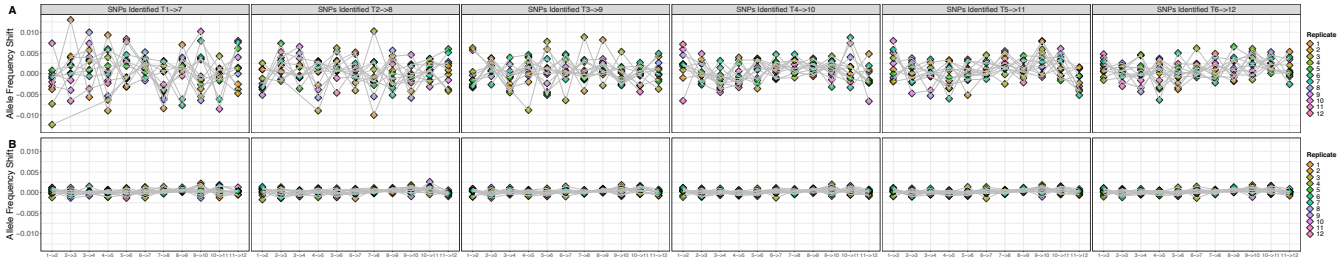


Figure S17 | Median shifts at target SNPs (A) and matched control SNPs (B) identified in each possible six time point interval and measured across each pair of single time points. Point color corresponds to replicate cage ID's and grey lines connect the median shifts for each cage between consecutive time point intervals.

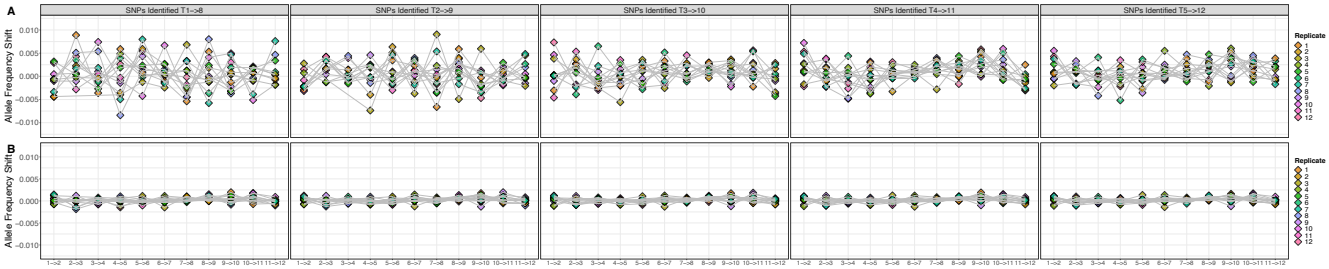


Figure S18 | Median shifts at target SNPs (A) and matched control SNPs (B) identified in each possible seven time point interval and measured across each pair of single time points. Point color corresponds to replicate cage ID's and grey lines connect the median shifts for each cage between consecutive time point intervals.

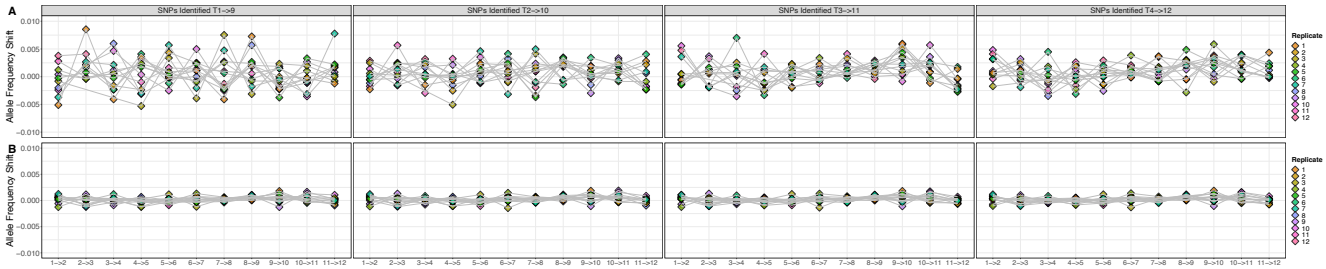


Figure S19 | Median shifts at target SNPs (A) and matched control SNPs (B) identified in each possible eight time point interval and measured across each pair of single time points. Point color corresponds to replicate cage ID's and grey lines connect the median shifts for each cage between consecutive time point intervals.

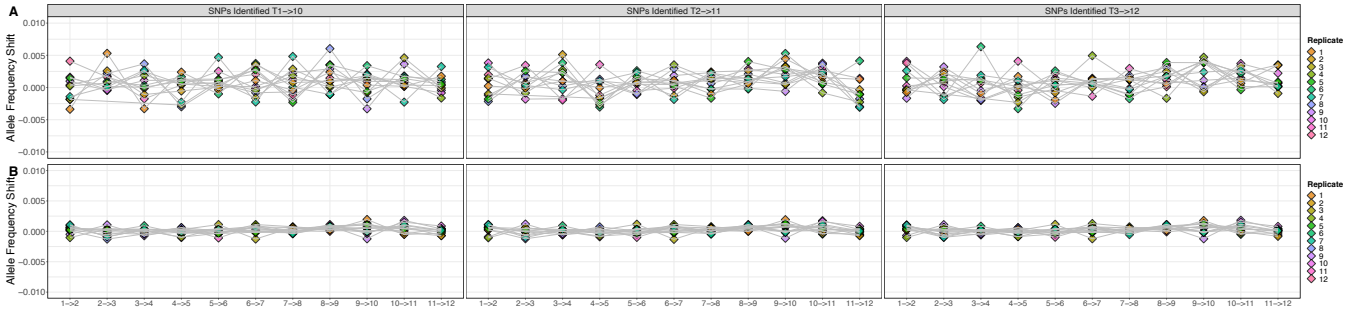


Figure S20 | Median shifts at target SNPs (A) and matched control SNPs (B) identified in each possible nine time point interval and measured across each pair of single time points. Point color corresponds to replicate cage ID's and grey lines connect the median shifts for each cage between consecutive time point intervals.

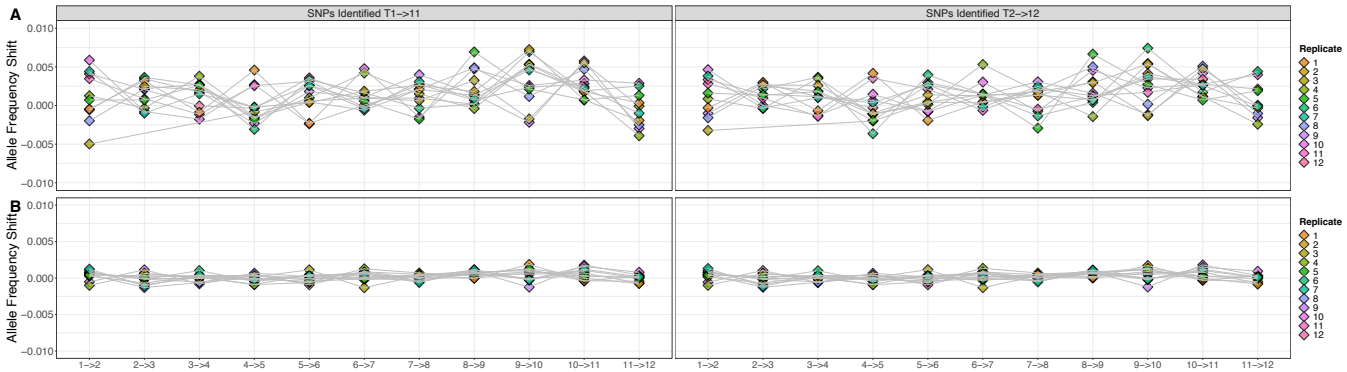


Figure S21 | Median shifts at target SNPs (A) and matched control SNPs (B) identified in each possible ten time point interval and measured across each pair of single time points. Point color corresponds to replicate cage ID's and grey lines connect the median shifts for each cage between consecutive time point intervals.

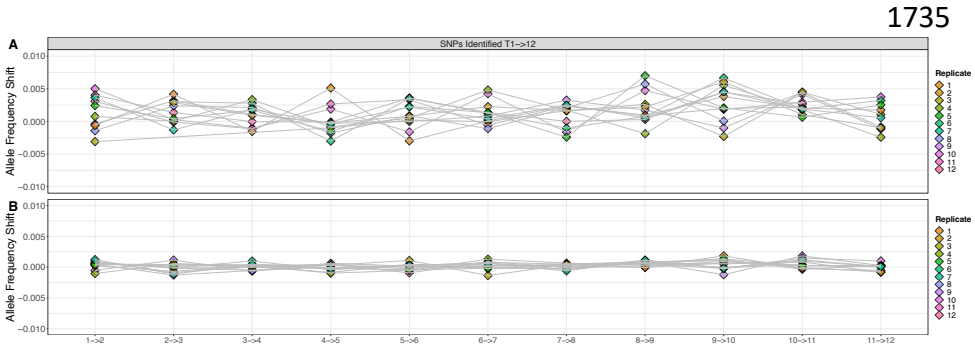


Figure S22 | Median shifts at target SNPs (A) and matched control SNPs (B) identified in each possible eleven time point interval and measured across each pair of single time points. Point color corresponds to replicate cage ID's and grey lines connect the median shifts for each cage between consecutive time point intervals.

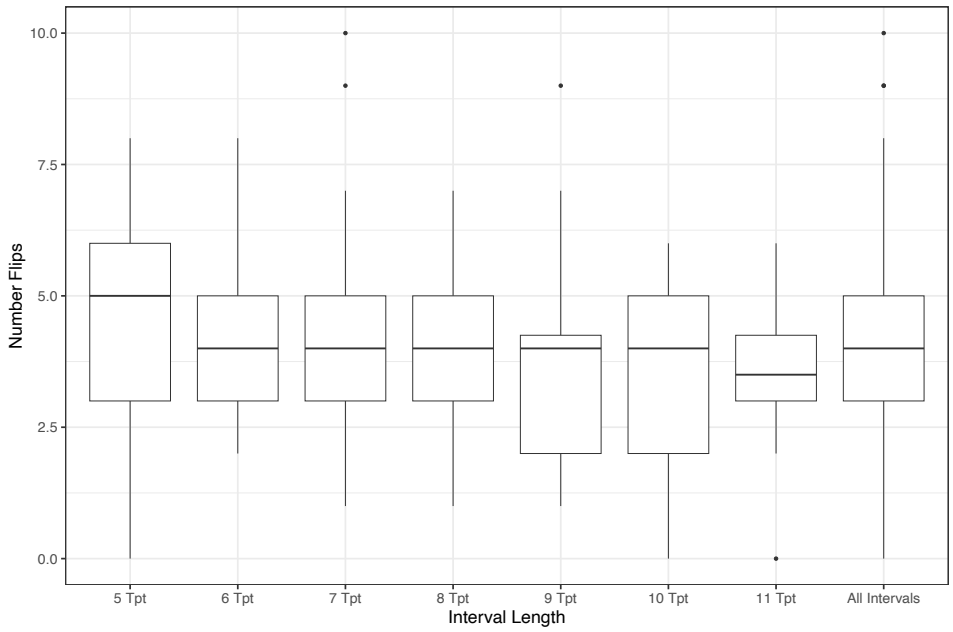


Figure S23 Distributions of the number of flips in the dominant sign of selection between consecutive time point intervals on sets of target SNPs identified across all interval lengths and comparisons and tested in left-out replicate cages. A flip in the dominant sign of selection was inferred if, for an individual cage, the distribution of shifts at SNPs changed from significantly greater to significantly less (or vice versa) than that of the matched control shifts across consecutive time point intervals. All sets of test SNPs were generated using a GLM and leave-one-out approach.

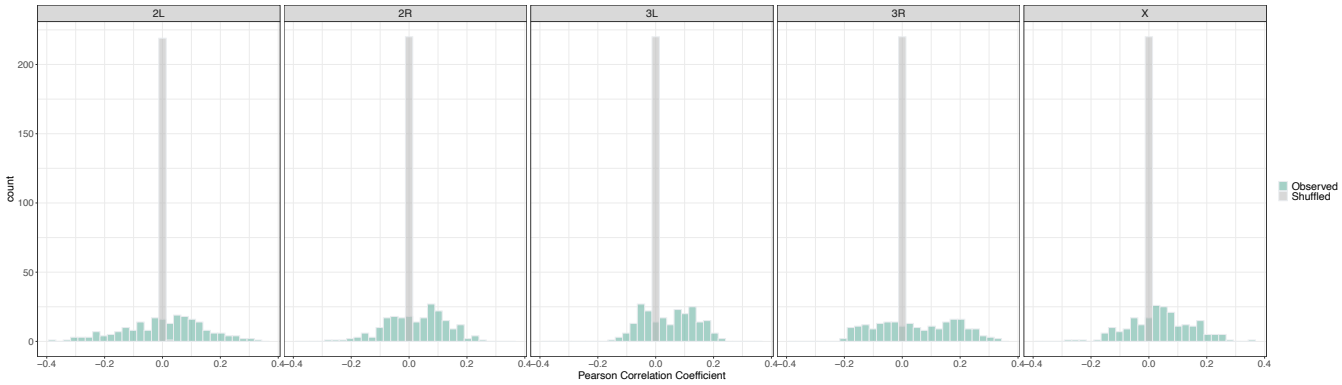


Figure S24 Histogram distributions of Pearson correlations computed between allele frequency shifts for each monthly interval of the 2014 experiment (N = 4), and each interval of the 2021 experiment (N = 56), segregated by chromosomal arm. Bar color correspond to coefficients generated via site information permutations (grey) and observed values (green).

Supplementary Tables

Table S1 | Enriched phenotypes identified using top unlinked cluster genes (see Methods), non-overlapping between the 2014 study conducted by Rudman et al. (2022) and the 2021 study reported here. Enriched phenotypes identified using ModPHEA (Weng and Lio 2017). Phenotypes putatively associated with environmental sensing are bolded.

2014 Enriched Phenotypes	2021 Enriched Phenotypes
adult antennal lobe glomerulus	**nervous system**
behavior defective	**head**
embryonic/larval tracheal system	**embryonic/larval nervous system**
size defective	embryonic/larval tracheal system
circadian rhythm defective	**neuron**
embryonic/larval imaginal precursor	**peripheral nervous system**
larval segment	cuticle
adult sense organ	**adult sensillum**
tracheal system	tracheal system
embryonic/larval digestive system	**central nervous system**
larval thorax	organism
circadian behavior defective	sclerite
stage 7 embryo	**adult head**
stage 8 embryo	ganglion
fused midgut primordium	**adult nervous system**
mitotic domain	muscle system
stage 16 embryo	appendage
imaginal disc	cuticular specialization
embryonic/larval central nervous system	developing embryonic structure
larval thoracic segment	organ system