Supplemental Materials

Mitochondrial lineage assignment

To ensure the selection experiment did not confound selection between mitochondrial lineages, which are potentially cryptic species, and selection within a lineage, we determined the major and minor mitochondrial haplotypes present in our samples. This analysis follows previous COI work conducted in *A. tonsa* (1, 2), particularly Figueroa et al. (3). We used a reference COI from Figueroa et al. to align and pull out the COI region in our raw FASTQ files. Variants were called using Varscan2 (4) in the same manner as our full dataset. Because these are pooled data, we cannot determine haplotypes of individuals. Further, if two lineages were present at equal frequencies, it would be difficult to determine which two haplotypes were present; the variants of each would be mixed in these pooled data. We assessed the distribution of allele frequencies to predict the number and frequency of mitochondrial haplotypes present in these data. Given the bimodal distribution of variants, where SNPs are either nearly fixed or absent (Fig. S9), there are likely two mitochondrial lineages present in these samples; one at very high frequency, and the other at low frequency. Looking at the shift in variant frequency from F0 to F25, it is apparent that the low frequency variant drops out over time while the high frequency variant (which is nearly fixed) remains. Because of the low number of variant sites, the high frequency variant also appears to match the reference sequence, which is from an individual from clade X. All results below hold when a reference from a different clade is used instead.

We reconstruct the major and minor haplotypes present in our data. Given the bimodal distribution of variants, we can reconstruct the consensus sequence of the major haplotypes as variants that are nearly fixed, while the minor haplotype consists of the low frequency variants. We only construct the minor haplotypes for F0 samples, as these haplotypes are no longer present by F25. Variants were filtered in R and VCF files containing either the major or minor variants were created. *bcftools* was used to generate consensus sequences for each sample and haplotype and sequences were concatenated with fasta sequences from Figueroa et al. to allow for direct comparisons downstream. This fasta file was aligned with MUSCLE (5), converted to NEXUS format in R using APE (6), and a phylogenetic tree was built using MrBayes (7) using the substitution model HKY + I + G, following Figueroa et al. The plot was generated with ggtree in R (8). The resulting tree clearly shows that the major haplotype is clade X, with a low frequency of clade S present at F0 (Fig. S10). The S clade drops out by F3 and is not present in any samples at F25, including the ambient selection line. Given the low

frequency of clade S and its absence from all lines by F25, the divergence detected in our experiment is not simply due to selection for a mitochondrial clade, but due to selection within individuals of clade X.

Analysis of covariance in allele frequency change

The replicated nature of our experiment coupled with the multiple selection regimes allows us to disentangle the relative contributions of drift and selection on the genome-wide changes in allele frequencies. We use and expand upon the approach developed by Buffalo and Coop (9). This method quantifies genome-wide covariance in the change in allele frequencies between replicates of a single treatment and between treatments to determine the relative contributions of drift and selection as well as to assess the degree of shared selection between selection regimes. Finally, we can leverage the presence of the control ambient line to estimate and remove the effects of the aforementioned estimates of lab adaptation.

First, it is possible to partition the changes in allele frequencies within a treatment into selection and drift components for replicate A as:

$$
\Delta P_{t,A} = \Delta_D P_t + \Delta_S P_t
$$

where $\Delta_S P_t$ is the change due to selection and $\Delta_D P_t$ is drift. The proportion of change due to selection can be further defined as the allele frequency change due to lab selection common to all replicates within a treatment group, $\Delta_L P_t$, as well as the change due to selection within a specific selection regime, $\Delta_R P_t$. The change in allele frequency in replicate A can be partitioned into,

$$
\Delta P_{t,A} = \Delta_D P_t + \Delta_L P + \Delta_R P_t
$$

Because the terms above are uncorrelated, the variance is,

$$
Var(\Delta p_t) = Var(\Delta_D P_t) + Var(\Delta_L P_t) + Var(\Delta_R P_t)
$$

We estimate the shared effects of selection regime from the allele frequency changes as the covariance of allele frequency change between any two replicates,

$$
Var(\Delta_R P_t) = Cov(\Delta P_{t,A}, \Delta P_{t,B})
$$

Where A and B indicate different replicates within a selection regime. Thus, to estimate the total shared response to selection within a selection regime, we estimate the covariance between all pairwise replicates and take their mean. Further, because of the presence of an ambient control line in our study, we can estimate the contribution of adaptation to the lab environment to the overall changes in allele frequency. The shared variance between a selection regime and the control represents the contribution of lab adaptation within a treatment group, again estimated as the mean of all pairwise comparisons between the two groups, giving $\text{Var}\Delta_L P_t$. We can subtract this value from the shared response within a selection regime, $\text{Var}(\Delta_R P_t)$, to get an estimate of the response to selection that is independent of estimated average lab selection effect.

After determining the contributions from selective regime and lab adaptation, the remaining variance can be attributed to the drift component. Finally, these values can be divided by the total variance to find the proportion each contributes to the overall variance in change in allele frequency.

Next, we can use similar principles to determine the shared response to selection between each selective regime. Here, the shared response to selection is again the covariance in allele frequency change between any two replicates, as above, but now from two different treatments,

$$
Var(\Delta_R P_t) = Cov(\Delta_{pt,A}, \Delta_{pt,B})
$$

We take the mean of all possible pairwise comparisons between treatments and scale this by the total variance to determine the proportion of total variance of allele frequency change that is shared. Finally, lab adaptation can be estimated and accounted for as described above.

Accounting for shared variance due to limited F0 replication

In these data, covariances are calculated from the change in allele frequency from the same set of F0 samples for all treatments. While these F0 samples likely represent the pre-selective state of the population, replicating these samples in this way leads to a spurious increase in the covariance estimates between samples in different treatments due to shared sampling variance. This is made clear in the covariance heatmap (Fig. S3A) where along the diagonal where the same replicate is compared between treatments the covariance is increased; Fig. S3B shows these covariance calculations in a different format where the same pattern is evident. Similarly, we see the same pattern when the convergence correlation is calculated (Fig. S4). To avoid this inflation, we dropped any covariances between samples with the same F0 reference when calculating shared response to selection. For example, the covariance in allele frequency change between OWA replicate 1 and Acidification replicate 1 would not be included when quantifying the shared selection response between these treatments. While this reduces our replication to an extent, the estimates are likely much more accurate estimates of the true impact of selection on allele frequency changes as all covariances are independent as a result.

Supplemental table 1: Number of loci assigned to each functional category for each SNP set. The first value is the total number of loci, the 2nd is the proportion of the total. P-values from chi-squared tests relative to the genome-wide distribution. Underlined values are significant (Bonferroni correction: P < 0.05/15 = 0.0033). The "No Annotation" category refers to loci that were not on a scaffold with any annotated genes, indicating they are likely in fragmented regions of the genome assembly.

Supplemental table 2: find at the end of this document.

Supplemental Figures

Figure S1: Empirical p-values from drift simulations. Points are individual loci with Tukey's boxplots. Values above each boxplot indicate the proportion of total loci in that bin.

Figure S2. Number of significant go terms per category. Conducted with the same snps as the similar plot in the main text (Fig. 2).

Figure S3: Pairwise covariance in allele frequency change between samples from F0 to F25. A) Each square and color indicate the covariance between two samples. Above the diagonal is a mirror of below. B) The same covariance estimates as in A, but with 95% bootstrap confidence intervals. Color indicates the group comparison. Note the increase in covariance between samples sharing the same replicate number due to an artifact from calculating allele frequency change from the same F0 sample. Samples with shared F0 samples (i.e., the same replicate number) were dropped from further calculations to avoid this bias.

Figure S4: Convergent correlations of allele frequency change from F0 to F25 between samples. Black lines for each point show the 95% bootstrap confidence interval. Note the increase in convergent correlation between samples sharing the same replicate number due to an artifact from calculating allele frequency change from the same F0 sample. Samples with shared F0 samples (i.e., the same replicate number) were dropped from further calculations.

Figure S5: Linkage disequilibrium estimates. Decay curves were fit by regressing the log of the physical distance with LD between base pairs. LD estimates increased with the strength of selection relative to the F0 founding population (P < 0.001). Founding population: intercept = 0.212 ± 0.001 , slope = -0.0497 ± 0.0003; Ambient: 0.245 ± 0.002, -0.0562 ± 0.0003; Acidification: 0.235 ± 0.001, -0.0549 ± 0.0003; Warming: 0.242 ± 0.001, -0.0556 ± 0.0003; OWA: 0.251 ± 0.002, -0.058 ± 0.0003

Fig. S6: Genetic diversity estimates. All treatment groups lost genetic diversity relative to the founding F0 population (Wilcoxon signed-rank test, P < 0.0001). Estimates for each group: F0 founding population: 0.0148 ± 0.0111; ambient: 0.0133 ± 0.0111; acidification: 0.0127 ± 0.0111; warming: 0.0133 ± 0.0110; OWA: 0.0138 ± 0.0112. Between the F25 treatments, only acidification lines were significantly lower than other lines (P < 0.001).

Fig. S7: Starting minor allele frequencies at F0 for loci significant for different treatments. Panel A includes lab adaptation loci in each group while these are removed in panel B. Error bars around the all loci group are 95% confidence intervals from 1000 random samples from all loci where the number of loci sampled was equal to the number of significant Acidification loci (A: 1,713, B: 506). Because the acidification group had the least number of significant loci, these confidence intervals are broader than if sampling was based on warming or OWA values.

Figure S8: Representative candidate genes underlying rapid adaptation. Points represent the average allele frequency change among replicates from the starting frequencies at F0. Loci have been filtered for minor allele frequency > 0.1 at F0.

Figure S9: Frequency of mitochondrial haplotypes. Histograms show the frequency of haplotypes from variants in the pooled data. At F0, there were low frequency variants present in the data, indicating the presence of a low frequency haplotype. These minor alleles drop out by F3 and nearly absent by F25 across all samples. This indicates that the vast majority of samples matched the reference genome after 3 generations of selection. Given the low starting frequency of minor mitochondrial haplotypes, this suggests the pooled samples likely consisted of a dominant mitochondrial haplotype and selection across the rest of the genome was likely not due to shifting frequencies of mitochondrial clades.

Figure S10: Phylogenetic tree from COI sequence data. We determined the clade of samples from this study using samples from Figueroa et al. (2020). Black and blue bars are the tree tips indicate major and minor haplotypes in the pooled samples, respectively. Samples in our study were predominantly clade X.

Supplemental References:

- 1. G. Chen, M. P. Hare, Cryptic ecological diversification of a planktonic estuarine copepod, Acartia tonsa. *Mol. Ecol.* **17**, 1451–1468 (2008).
- 2. G. Chen, M. P. Hare, Cryptic diversity and comparative phylogeography of the estuarine copepod Acartia tonsa on the US Atlantic coast. *Mol. Ecol.* **20**, 2425–2441 (2011).
- 3. N. J. Figueroa, D. F. Figueroa, D. Hicks, Phylogeography of Acartia tonsa Dana, 1849 (Calanoida: Copepoda) and phylogenetic reconstruction of the genus Acartia Dana, 1846. *Mar. Biodivers.* **50**, 23 (2020).
- 4. D. C. Koboldt, *et al.*, VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
- 5. R. C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).
- 6. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
- 7. F. Ronquist, *et al.*, MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
- 8. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T. Lam, Ggtree : An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
- 9. V. Buffalo, G. Coop, Estimating the genome-wide contribution of selection to temporal allele frequency change. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 20672–20680 (2020).

Supplemental table 2: Go enrichment of candidate SNPs from topGO. The group column indicates the set of SNPs in the enrichment test. "_all" means all significant loci for that treatment were tested. "_unique" means only those SNPs unique to that treatment. When multiple treatments are listed this indicates SNP sets that are significant under multiple treatments.

