Low concordance of short-term and long-term selection responses in experimental *Drosophila* populations

Anna Maria Langmüller[1,2] and Christian Schlötterer[2,*]

[1]Institut für Populationsgenetik, Vetmeduni Vienna, Veterinärplatz 1, 1210 Wien, Austria

[2]Vienna Graduate School of Population Genetics, Vienna, Austria

[*]corresponding author: christian.schloetterer@vetmeduni.ac.at

September 5, 2019

ABSTRACT

Experimental evolution is becoming a popular approach to study genomic selection responses of evolving populations. Computer simulation studies suggested that the accuracy of the signature increases with the duration of the experiment. Since some assumptions of the computer simulations may be violated, it is important to scrutinize the influence of the experimental duration with real data. Here, we use a highly replicated Evolve and Resequence study in *Drosophila simulans* to compare the selection targets inferred at different time points. At each time point approximately the same number of SNPs deviated from neutral expectations, but only 10 % of the selected haplotype blocks identified from the full data set could be detected in the first 20 generations. Those haplotype blocks that emerged already after 20 generations differ from the others by being strongly selected at the beginning of the experiment and displaying a more parallel selection response. Consistent with previous computer simulations, our results confirm that only Evolve and Resequence experiments with a sufficient number of generations can characterize complex adaptive architectures.

KEYWORDS: experimental evolution, evolve & resequence, *Drosophila simulans*, early adaptation, replicated time series data, concordance of adaptation, window-based analysis

INTRODUCTION

Deciphering the adaptive architecture is a long-term goal in evolutionary biology. In contrast to natural populations, experimental evolution (EE) provides the possibility to replicate experiments under controlled, identical conditions and to study how evolution shapes populations in real time (Kawecki et al. (2012); Schlötterer et al. (2015)). The combination of EE with next-

generation sequencing - Evolve and Resequence (E&R) (Turner et al. (2011); Schlötterer et al. (2015); Long et al. (2015)) - has become a popular approach to study the genomic response to selection and to identify adaptive loci. E&R has been applied to various selection regimes, such as virus infection (Martins et al. (2014)), host-pathogen co-adaptation (Papkou et al. (2019)), thermal adaptation (Orozco-Terwengel et al. (2012); Barghi et al. (2019)), or body weight (Johansson et al. (2010)). A wide range of experimental designs have been used, which vary in census population size, replication level, history of the ancestral populations, selection regime, and number of generations (Garland and Rose (2009); Turner et al. (2011); Kawecki et al. (2012); Lang et al. (2013); Burke et al. (2014); Huang et al. (2014); Hardy et al. (2018); Castro et al. (2019); Michalak et al. (2019); Seabra et al. (2019)). The duration of E&R studies ranged from less than 20 (Kelly and Hughes (2018); Turner and Miller (2012); Rêgo et al. (2019)), over a few dozen (Orozco-Terwengel et al. (2012); Johansson et al. (2010)), up to hundreds of generations (Burke et al. (2010)). Computer simulations showed that the number of generations has a strong influence on the power of E&R studies, and increasing the number of generations typically improved the results (Baldwin-Brown et al. (2014); Kofler and Schlötterer (2014); Vlachos and Kofler (2019)). Since simulations make simplifying assumptions, it is important to scrutinize these conclusions with empirical data. Until recently no suitable data-sets were available, which included multiple time points and replicates. We use an E&R experiment (Barghi et al. (2019)), which reports allele frequency changes in 10 replicates over 60 generations in 10 generation intervals, to investigate the impact of the experimental duration on the observed genomic response. The time resolved genomic data of this experiment allows to contrast putative selection targets inferred at different time points. We show that only a subset of the selection targets are detected at earlier generations, which are not representative of the underlying adaptive architecture.

METHODS

Experimental *Drosophila simulans* populations

A detailed description of the *Drosophila simulans* E&R experiment can be found in Barghi et al. (2019) and Hsu et al. (2019). The pooled individuals (Schlötterer et al. (2014)) from the evolving populations were sequenced every 10^{th} generation starting with the founder population (generation 0) until generation 60. This E&R experiment started from 202 isofemale lines, which were collected in Tallahassee, Florida. 10 replicate populations evolved in the laboratory at a "cycling hot" temperature regime (12 hours light and 28 °C, 12 hours dark and 18 °C). The census size of the replicates was 1,000 individuals with non-overlapping generations (Barghi et al. (2017, 2019); Hsu et al. (2019)).

Genomic analysis hierarchy

We investigated the genomic response of the experimental *Drosophila* populations on three different levels: candidate SNPs, candidate SNPs in a window of fixed length and candidate SNPs shared with reconstructed selected haplotype blocks. A detailed description for each level is given below. Reasoning that the most reliable signal is detected at the most advanced generation (60), we performed the same analysis at earlier time points and determined to what extent the same selection targets were identified as in generation 60.

Identification of candidate SNPs

Barghi et al. (2019) applied various filtering steps to obtain a robust SNP set from the ancestral population. In short, SNPs were called applying the following criteria: base quality of 40 in at least one replicate, a coverage between the 2^{nd} and 98^{th} percentile, and the minor allele is sup-

ported by at least ten reads. Repeats, transposable elements, SNPs specific to Y-translocated genes (Tobler et al. (2017)) and 5-bp regions around indels were excluded from the analysis to increase the robustness of the SNP set (for further details, see Barghi et al. (2019)), resulting in 5,096,200 SNPs on chromosome X, 2, 3, and 4.

To study the selection response at different time points, we identified "candidate SNPs" based on the frequency difference between the ancestral and evolved populations for each of the six time points. Following Barghi et al. (2019), replicates were tested separately (Fisher's exact test) and jointly (Cochran-Mantel-Haenszel test, CMH) to identify SNPs with pronounced allele frequency change (AFC) using `PoPoolation2` (Schlötterer et al. (2011)). Minimum and maximum coverage restrictions were not imposed because outlier SNPs with extreme coverage had already been removed. Neither the CMH test nor the Fisher's exact test account for AFC due to genetic drift. To detect SNPs that show more AFC than expected under drift, we performed neutral simulations with `Nest` (Jónás et al. (2016)) using estimates of the effective population size (N_e) between generation 0 and the focal time point (Table S1- S3). The simulations further used the empirical starting allele frequencies and sequencing coverages. For the CMH test, N_e estimates were averaged across replicates for autosomes and the X chromosome separately. For Fisher's exact test, we used replicate-specific N_e estimates of the autosomes. Based on these neutral simulations we determined candidate SNPs with a false discovery rate smaller than 5 % (Barghi et al. (2019)).

We identified 56,166 candidate SNPs in generation 60, compared to 55,199 in Barghi et al. (2019). This small discrepancy can be explained by stochastic differences arising from the neutral simulations used to determine the significance threshold. We excluded six haplotype blocks (3.17, 2.27, 3.21, 3.48, 3.49 and 3.54) (Barghi et al. (2019)) with less than 90% of the previously reported candidate SNPs.

Identification of candidate windows

The number of candidate SNPs is inflated as a result of linkage disequilibrium in the experimental populations (Nuzhdin and Turner (2013); Tobler et al. (2014)). To account for non−independence of candidate SNPs we used a window based approach. We split the main chromosomes (X, 2 and 3) into non-overlapping windows of 5,000 SNPs that are segregating in all generations and replicates. We chose SNPs instead of base pairs as window size measure to allow for variation in SNP density along the genome. To determine if a given window contains more candidate SNPs than expected, we sampled the same number of random SNPs as candidate SNPs in this window (1,000 iterations). "Candidate windows" contained at least as many candidate SNPs as the 99^{th} percentile of randomly sampled SNPs. Applying the procedure independently to candidate SNPs from all time points provides time point specific candidate windows (Figure S1). We evaluated the similarity of two time points with the Jaccard index (for both candidate SNPs, and candidate windows).

The number of candidate SNPs in a window is a summary statistic which ignores the significance of the candidate SNPs. If a signal is robust between two time points, we expect the same p-value based ranking of candidate SNPs. Thus, we also evaluated whether candidate SNPs in a given window had a similar relative significance. For each candidate window we created a ROC-like curve (similar to Jakšić and Schlötterer (2016)) by ranking the candidate SNPs by their p-values - the most significant SNP was assigned rank 1 - and calculating the overlap in top-ranked SNPs between two time points.

HAplotype block Discovery Rate (HADR)

Barghi et al. (2019) clustered candidate SNPs from F60 into selected haplotype blocks based on similar allele frequency trajectories over time and replicates

(Franssen et al. (2016)). The reconstructed haplotype blocks were further validated with experimentally phased haplotypes from ancestral and evolved populations (Barghi et al. (2019)) and 96 % of the reconstructed haplotype blocks could be confirmed. This suggests that reconstructed haplotype blocks provide a reliable set of linked candidate SNPs.

Taking advantage of this additional confirmation of the candidate SNPs in a selected haplotype block we developed a third measure of similarity between time points. We determined the fraction of candidate SNPs comprising a haplotype block that were also discovered at a given time point (haplotype block discovery rate, HADR) using the `poolSeq` R-package (Taus et al. (2017)). We note that inference of selected haplotype blocks at each generation does not provide a good alternative to HADR, as the ability to cluster SNPs into haplotype blocks is dependent on the number of time points (Franssen et al. (2016)), resulting in less power at early time points compared to later ones.

Early Detected HAplotype blocks (EDHAs)

We applied hierarchical clustering (Pollard and Laan (2005)), PCA and kmeans (Hartigan and Wong (1979)) to group haplotype blocks based on their HADR patterns. The hyper-parameter k, which determines the number of clusters, was set to 5 based on the gap statistic approach (Tibshirani et al. (2001)). The k-means clustering resulted into a group of 10 haplotype blocks with elevated HADR in generation 20 (Figure S2). This group of 10 haplotype blocks can also be separated from other haplotype blocks by the first principal component of a PCA applied to HADR from F10 to F50 (Figure S3). We refer to the haplotype blocks in this cluster as early detected haplotype blocks (EDHAs). We evaluated whether EDHAs have distinct characteristics compared to all other haplotype blocks using the following features: haplotype block length, median starting allele frequency, average recombination rate (*D. simulans* recombination map

from Howie et al. (2019)), selection coefficient (s, estimated with `poolSeq` (Taus et al. (2017))) in generation 20, s in generation 60, selection coefficient ratio $r_s = \frac{s_{20}}{s_{60}}$, and number of rising replicates in generation 20. Following Barghi et al. (2019) , we classified a haplotype block as replicate specific, if the allele frequency of candidate SNPs from a haplotype block increases on average by at least 10%. We also used AFC thresholds of 5, 15, and 20% to determine whether these haplotype blocks were rising in a given replicate. Selection coefficients were averaged (mean) over replicates that passed the AFC threshold.

Results & Discussion

Subsequent time points are more similar for advanced generations

We studied the similarity of selection signatures for different time points using 10 replicates of a *D. simulans* population, which evolved for 60 generations to a novel hot environment (Barghi et al. (2019)). With Pool-Seq data from every 10^{th} generation, we evaluated the selection signature on three different levels: candidate SNPs, candidate SNPs in a window of fixed length and candidate SNPs shared with reconstructed selected haplotype blocks. The similarity of two time points was determined by the Jaccard index, a dimensionless parameter ranging from 0 (no overlap between two sets) to 1 (sets are identical). We found that all candidate SNP sets are more similar than expected by chance. The Jaccard index ranged from 0.08 (generation 10 vs generation 60) to 0.40 (generation 50 vs generation 60), where subsequent time points are more similar than those separated for more than 10 generations (e.g. J=0.15 (generation 10 vs generation 20); J=0.08 (generation 10 vs generation 60)). Furthermore, the similarity of candidate SNP sets from subsequent time points increases with time until it ultimately more than doubles for the last two generations (J=0.15 (generation 10 vs generation 20); J=0.34 (generation 50 vs

generation 60), Figure 1A). The monotonic increase in similarity with time shows that selection patterns are more reliably detected at later generations.

Since the analysis of single SNPs suffers from considerable stochasticity, and neighboring SNPs are not independent (Tobler et al. (2014); Howie et al. (2019)), we also repeated the analysis of different time points using non-overlapping windows of 5,000 SNPs. Reasoning that windows containing a target of selection will harbor multiple candidate SNPs, we defined selected windows as those, which harbor more candidate SNPs than expected by chance. Consistent with higher stochasticity on the SNP level, a higher similarity was observed for candidate windows (from J=0.26 (generation 10 vs generation 60) to J=0.62 (generation 40 vs generation 50)). Again, adjacent time points have a higher Jaccard index than time points farther apart (J=0.26 (generation 10 vs generation 60) J=0.39 (generation 10 vs generation 20)). The similarity of subsequent time points also increases with the duration of the experiment (J=0.39 (generation 10 vs generation 20) ; J=0.59 (generation 50 vs generation 60), Figure 1A). In contrast to the SNP level, the set of selected windows after 10 generations is only significantly similar to generation 20, but not to any other generation. Thus, the pattern of less reliable selection targets in the early generations is confirmed on the window level, albeit with different significance levels.

For an alternative measure of similarity we used the ranking of candidate SNPs in a specific window based on their p-values and compared it between different time points. If a signal is robust between two time points, we expect the same SNP ranking in a selected window. Consistent with the other tests, we found that the congruence in candidate SNP ranking increases with time (Figure 2). To rule out that rare SNPs are responsible for the dissimilarity between early and late time points, we calculated similarity measures based on SNPs that are segregating at all generations and time points. Nevertheless, including SNPs which were lost in at least one replicate during the experiment did not result in a pronounced decrease in similarity (Figure

S4).

The analysis of selected haplotype blocks provides another possibility to control for non-independence of single candidate SNPs. We calculated the haplotype block discovery rate (HADR) - the fraction of candidate SNPs in a haplotype block that are rediscovered at a given time point. Similar to the other analyses, we observe higher similarity between later time points (Figure 1B), with a pronounced increase of median HADR between generation 30 (< 25 %) and 40 (> 50 %).

Independent of the measure of similarity between time points, we consistently find that selection signatures at early time points are less reliable than those from later time points. With the limitation that the true targets of selection are not known, this observation highlights that a more reliable identification of selection targets strongly benefits from additional generations of selection. Given that fewer "real" selection targets were identified in the first time points, it is remarkable that a similar number of candidate SNPs was detected at each time point (Table S4). This may imply that earlier time points harbor more false positives, but it is also possible that these targets were only selected during the first generations. To distinguish between these two alternative explanations, the analysis of haplotypes will be required to separate linked hitchhikers from selected sites. Furthermore, experimental validation of selection targets in secondary E&R studies (Burny et al. (2019)) may be another route to confirm selection signatures beyond statistical testing.

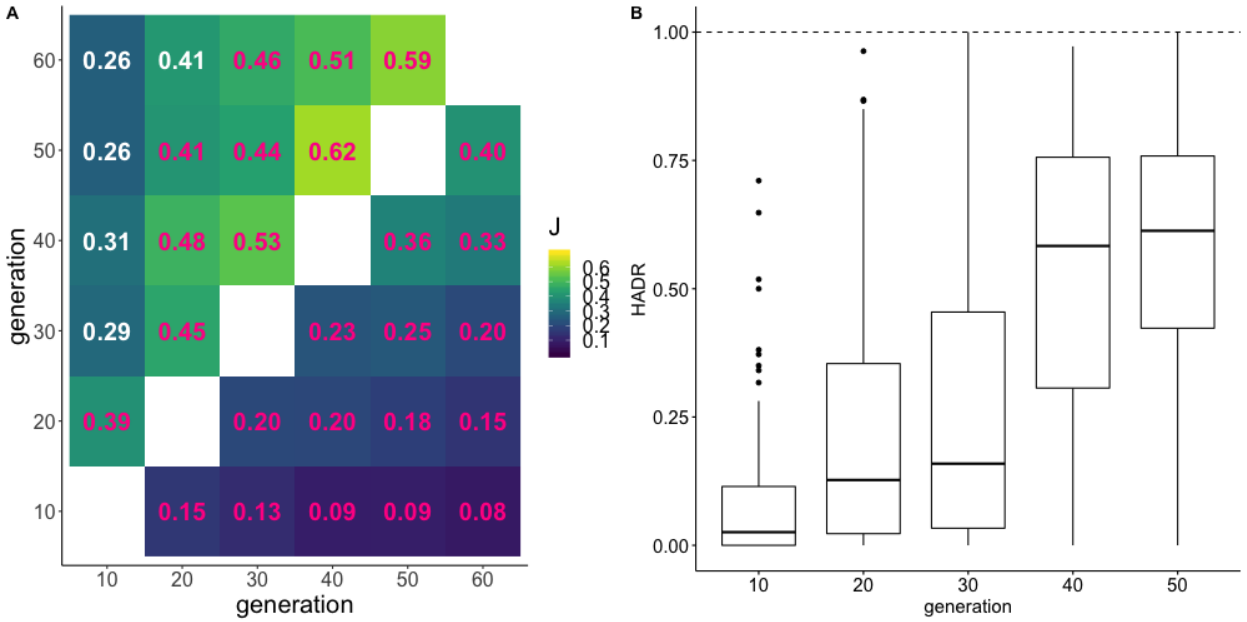

Figure 1: Similarity measures for candidate SNPs, candidate SNPs in a window of fixed length and candidate SNPs shared with reconstructed selected haplotype blocks
(**A**): Jaccard index for pairwise comparisons of candidate sets. The top triangle shows candidate window sets, the bottom triangle candidate SNP sets. Significant similarities (p-value <0.05 after multiple testing correction, 10,000 bootstraps) are marked in pink.
B:The rate at which selected SNPs of 93 haplotype blocks from generation 60 were already discovered at earlier generations (HADR).

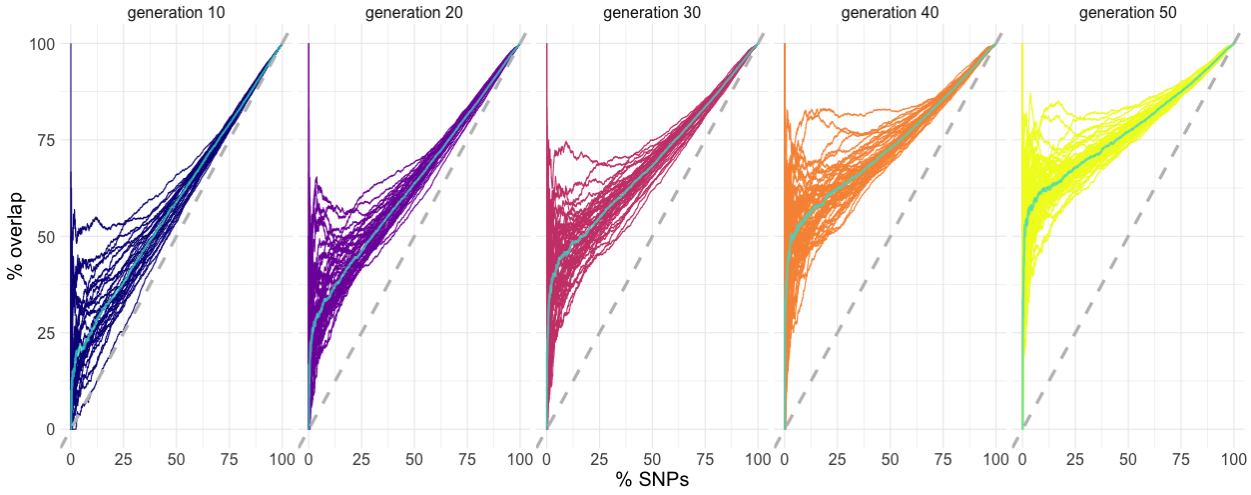

Figure 2: The rank of candidate SNPs becomes more congruent with time. In this ROC-like graph the ranking of all candidate SNPs in candidate windows is compared. Each panel shows one intermediate time point compared to generation 60. The percent overlap for each candidate window is indicated by a separate line. The median overlap (turquoise line) monotonically increases with experimental duration, demonstrating that the ranking of candidate SNPs is more robust for advanced generations. The expected overlap is shown as dashed, grey line.

Only few selection targets are shared across all generations

More than 27,000 candidate SNPs can be identified at each time point (Table S4), but only a small (4.8 %) subset is consistently detected at every generation (Figure 3; including rare SNPs see Figure S5). Independent of the importance of more reliable selection signatures with an increasing number of generations, this analysis raises an important concern about the usefulness of meta-analyses on the SNP level. With less than 5% of the SNPs being shared in the same selection experiment, it will be extremely difficult to compare studies that started from different founder populations and were selected for a different number of generations.

We repeated the analysis for windows and determined the number of selected windows that are shared across all generations. With 18 out of 74 candidate windows in generation 60 (24.3 %, Figure 3) being detected at all generations, the window analysis shows more consistency

across time points than a SNP-based analysis. This observation is independent of window size (Table S5- S6) and the inclusion of rare SNPs into the analysis (Figure S5). We propose that meta-analyses of E&R data should be performed on the level of windows, or probably based on selected haplotype blocks to avoid false negatives due to the high stochasticity of SNP-based analyses.

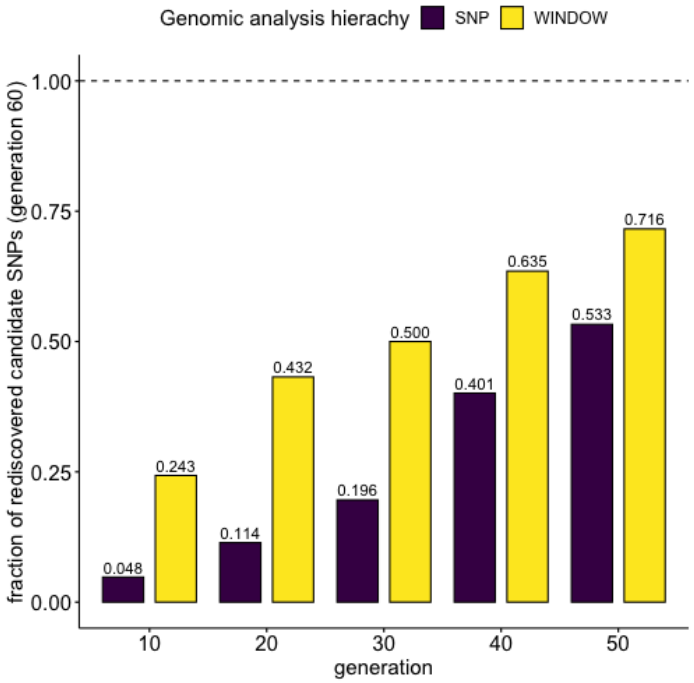

Figure 3: Less than 5% of candidate SNPs in generation 60 are detected consistently at every generation. The bars depict the fraction candidate SNPs (purple) and candidate windows (yellow) at generation 60, which are candidates in all subsequent generations (e.g. 40.1% of generation 60 candidate SNPs are candidates in generation 50 and 40). Candidate windows are more consistent than candidate SNPs. Figure S5 depicts the ratios for candidate sets that are not restricted to SNPs segregating in all generations and time points.

Selection signatures detected early in the experiment are not representative of the underlying adaptive architecture

This study focused on the comparison of selection targets detected at early and late time points. Since analyses based on single SNP are very stochastic, we investigated the fraction of candidate SNPs comprising a haplotype block that were also discovered at earlier time points (HADR). We detected 10 haplotype blocks with elevated HADR in generation 20 (EDHAs, Figure S3). 10 EDHAs were detected based on kmeans clustering (see Material & Methods and Figures S2-S3). We found that EDHAs do not differ in their starting allele frequency, haplotype block length, average recombination rate or absolute selection coefficients from other haplotype blocks (Figure S6). EDHAs are, however, more strongly selected at the beginning of the experiment, but do not differ from the remaining haplotype blocks at later generations. The comparison of the relative selection intensity of early and late time points identified significant differences of EDHAs from the other haplotype blocks (Figure 4A). Consistent with stronger selection at earlier time points, the selection signature of EDHAs is significantly more parallel across replicates after 20 generations of adaptation. (Figure 4B). All statistical tests, which are evaluating a parallel selection signature across replicates, are more likely to detect selection signatures, which are shared across replicates, even with only moderate allele frequency changes. This could result in a biased picture of the underlying genetic architecture. The analysis of selection signatures in replicated experiments running for only a moderate number of generations is more likely to detect parallel than replicate specific selection signatures. This bias is not restricted to our study, but also an experimental study of *D. simulans* populations adapting 10 to 20 generations to a new temperature regime (Kelly and Hughes (2018)) found more parallel selection responses. We propose that additional analyses contrasting selection signatures of early and late time points are needed to confirm the enrichment of parallel selection signatures in short-term experiments.

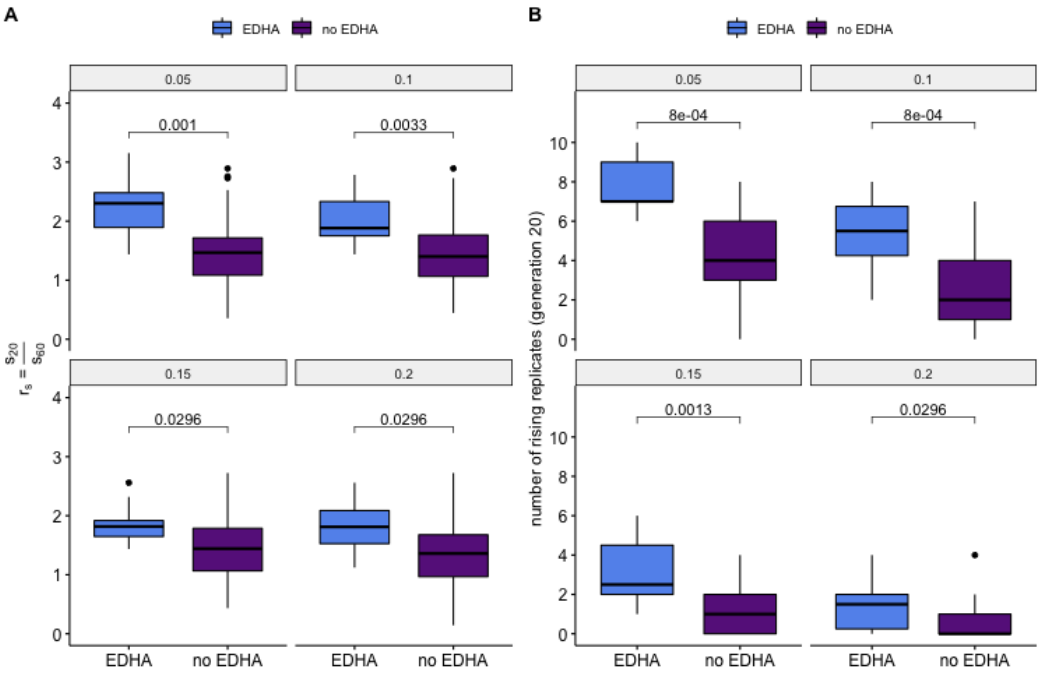

Figure 4: Early Detectable HAplotype blocks (EDHAs) differ from the other selected haplotype blocks. The ratio of selection coefficients determined for early generations (generation 20) and late generations (generation 60) is significantly higher for EDHAs (**A**). EDHAs rise in more replicates (**B**) than other haplotype blocks. Both observations are robust to different AFC thresholds. Values above the boxplots represent the two-tailed Mann-Whitney test *p*-values corrected for multiple testing with the Benjamini-Hochberg procedure.

ACKNOWLEDGMENTS

We thank the members of the Institut für Populationsgenetik for fruitful discussion and support. Special thanks to Neda Barghi, Sheng-Kai Hsu and Claire Burny for helpful comments on earlier versions of the manuscript. This work was supported by the Austrian Science Fund (FWF, grant W1225) and an Europen Research Council (ERC) grant (ArchAdapt).

DATA AVAILABILITY

All scripts will be uploaded and available on GitHub upon publication.

Bibliography

Baldwin-Brown, J. G., A. D. Long, and K. R. Thornton

2014. The power to detect quantitative trait loci using resequenced, experimentally evolved populations of diploid, sexual organisms. *Molecular Biology and Evolution*, 31(4):1040–1055.

Barghi, N., R. Tobler, V. Nolte, A. M. Jakšić, F. Mallard, K. A. Otte, M. Dolezal, T. Taus, R. Kofler, and C. Schlötterer

2019. Genetic redundancy fuels polygenic adaptation in *Drosophila*. *PLOS Biology*, 17(2):1–31.

Barghi, N., R. Tobler, V. Nolte, and C. Schlötterer

2017. *Drosophila simulans*: A species with improved resolution in evolve and resequence studies. *G3: Genes, Genomes, Genetics*, 7(7):2337–2343.

Burke, M. K., J. P. Dunham, P. Shahrestani, K. R. Thornton, M. R. Rose, and A. D. Long

2010. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature*, 467(7315):587–590.

Burke, M. K., G. Liti, and A. D. Long

2014. Standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. *Molecular Biology and Evolution*, 31(12):3228–3239.

Burny, C., V. Nolte, P. Nouhaud, M. Dolezal, and C. Schlötterer

2019. Secondary evolve and re-sequencing: an experimental confirmation of putative selection targets without phenotyping. *bioRxiv*.

Castro, J. P., M. N. Yancoskie, M. Marchini, S. Belohlavy, L. Hiramatsu, M. Kučka, W. H. Beluch, R. Naumann, I. Skuplik, J. Cobb, N. H. Barton, C. Rolian, and Y. F. Chan

2019. An integrative genomic analysis of the longshanks selection experiment for longer limbs in mice. *eLife*, 8:e42014.

Franssen, S. U., C. Schlötterer, and N. H. Barton

2016. Reconstruction of Haplotype-Blocks Selected during Experimental Evolution. *Molecular Biology and Evolution*, 34(1):174–184.

Garland, T. and M. Rose

2009. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*.

Hardy, C. M., M. K. Burke, L. J. Everett, M. V. Han, K. M. Lantz, and A. G. Gibbs

2018. Genome-Wide Analysis of Starvation-Selected *Drosophila melanogaster*-A Genetic Model of Obesity. *Molecular Biology and Evolution*, 35(1):50–65.

Hartigan, J. A. and M. A. Wong

1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108.

Howie, J. M., R. Mazzucco, T. Taus, V. Nolte, and C. Schlötterer

2019. DNA Motifs Are Not General Predictors of Recombination in Two *Drosophila* Sister Species. *Genome Biology and Evolution*, 11(4):1345–1357.

Hsu, S.-K., A. M. Jaksic, V. Nolte, N. Barghi, F. Mallard, K. A. Otte, and S. Christian

2019. A 24 h Age Difference Causes Twice as Much Gene Expression Divergence as 100 Generations of Adaptation to a Novel Environment. *Genes*, 10(2).

Huang, Y., S. I. Wright, and A. F. Agrawal

2014. Genome-wide patterns of genetic variation within and among alternative selective regimes. *PLOS Genetics*, 10(8):1–15.

Jakšić, A. M. and C. Schlötterer

2016. The interplay of temperature and genotype on patterns of alternative splicing in *Drosophila melanogaster*. *Genetics*, 204(1):315–325.

Johansson, A. M., M. E. Pettersson, P. B. Siegel, and Ö. Carlborg

2010. Genome-wide effects of long-term divergent selection. *PLOS Genetics*, 6(11):1–12.

Jónás, Á., T. Taus, C. Kosiol, C. Schlötterer, and A. Futschik

2016. Estimating the effective population size from temporal allele frequency changes in experimental evolution. *Genetics*, 204(2):723–735.

Kawecki, T. J., R. E. Lenski, D. Ebert, B. Hollis, I. Olivieri, and M. C. Whitlock

2012. Experimental evolution. *Trends in Ecology & Evolution*, 27(10):547 – 560.

Kelly, J. K. and K. A. Hughes

2018. Pervasive linked selection and intermediate-frequency alleles are implicated in an evolve-and-resequencing experiment of *Drosophila simulans*. *Genetics*, 211(March):genetics.301824.2018.

306 Kofler, R. and C. Schlötterer

307 2014. A guide for the design of evolve and resequencing studies. *Molecular Biology and*
308 *Evolution*, 31(2):474–483.

309 Lang, G. I., D. P. Rice, M. J. Hickman, E. Sodergren, G. M. Weinstock, D. Botstein, and M. M.
310 Desai

311 2013. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast popula-
312 tions. *Nature*, 500(7464):571–574.

313 Long, A., G. Liti, A. Luptak, and O. Tenaillon

314 2015. Elucidating the molecular architecture of adaptation via evolve and resequence exper-
315 iments. *Nature Reviews Genetics*, 16(10):567–582.

316 Martins, N. E., V. G. Faria, V. Nolte, C. Schlötterer, L. Teixeira, É. Sucena, and S. Magalhães

317 2014. Host adaptation to viruses relies on few genes with different cross-resistance proper-
318 ties. *Proceedings of the National Academy of Sciences*, 111(16):5938–5943.

319 Michalak, P., L. Kang, M. F. Schou, H. R. Garner, and V. Loeschcke

320 2019. Genomic signatures of experimental adaptive radiation in *Drosophila*. *Molecular Ecol-*
321 *ogy*, 28(3):600–614.

322 Nuzhdin, S. V. and T. L. Turner

323 2013. Promises and limitations of hitchhiking mapping. *Current Opinion in Genetics & De-*
324 *velopment*, 23(6):694–699.

325 Orozco-Terwengel, P., M. Kapun, V. Nolte, R. Kofler, T. Flatt, and C. Schlõtterer

326 2012. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally hetero-
327 geneous trajectories of selected alleles. *Molecular Ecology*, 21(20):4931–4941.

328 Papkou, A., T. Guzella, W. Yang, S. Koepper, B. Pees, R. Schalkowski, M.-C. Barg, P. C. Rosenstiel,
329 H. Teotónio, and H. Schulenburg
330 2019. The genomic basis of Red Queen dynamics during rapid reciprocal host–pathogen
331 coevolution. *Proceedings of the National Academy of Sciences*, 116(3):923–928.

332 Pollard, K. S. and M. J. V. D. Laan
333 2005. Analysis of Genomic Data with Applications in R. *U.C. Berkeley Division of Biostatistics*
334 *Working Paper Series*, Working Paper 167.

335 Rêgo, A., F. J. Messina, and Z. Gompert
336 2019. Dynamics of genomic change during evolutionary rescue in the seed beetle *Calloso-*
337 *bruchus maculatus*. *Molecular Ecology*, 28(9):2136–2154.

338 Schlötterer, C., R. Kofler, E. Versace, R. Tobler, and S. U. Franssen
339 2015. Combining experimental evolution with next-generation sequencing: a powerful tool
340 to study adaptation from standing genetic variation. *Heredity*, 114(5):431–440.

341 Schlötterer, C., R. Tobler, R. Kofler, and V. Nolte
342 2014. Sequencing pools of individuals — mining genome-wide polymorphism data without
343 big funding. *Nature Reviews Genetics*, 15(11):749–763.

344 Schlötterer, C., R. V. Pandey, and R. Kofler
345 2011. PoPoolation2: identifying differentiation between populations using sequencing of
346 pooled DNA samples (Pool-Seq). *Bioinformatics*, 27(24):3435–3436.

347 Seabra, S. G., I. Fragata, M. A. Antunes, G. S. Faria, M. A. Santos, V. C. Sousa, P. Simões, and

M. Matos

2019. Different Genomic Changes Underlie Adaptive Evolution in Populations of Contrasting History. *Molecular Biology and Evolution*, 36(6):1358–1358.

Taus, T., A. Futschik, and C. Schlötterer

2017. Quantifying Selection with Pool-Seq Time Series Data. *Molecular Biology and Evolution*, 34(11):3023–3034.

Tibshirani, R., G. Walther, and T. Hastie

2001. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society Series B*, 63:411–423.

Tobler, R., S. U. Franssen, V. Nolte, and C. Schlötterer

2014. Patterns of Linkage Disequilibrium and Long Range Hitchhiking in Evolving Experimental *Drosophila melanogaster* Populations. *Molecular Biology and Evolution*, 32(2):495–509.

Tobler, R., V. Nolte, and C. Schlötterer

2017. High rate of translocation-based gene birth on the *Drosophila* Y chromosome. *Proceedings of the National Academy of Sciences*, 114(44):11721–11726.

Turner, T. L. and P. M. Miller

2012. Investigating natural variation in *Drosophila* courtship song by the evolve and resequence approach. *Genetics*, 191(2):633–642.

Turner, T. L., A. D. Stewart, A. T. Fields, W. R. Rice, and A. M. Tarone

2011. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLOS Genetics*, 7(3):1–10.

370 Vlachos, C. and R. Kofler

371 2019. Optimizing the power to identify the genetic basis of complex traits with Evolve and

372 Resequence studies. *Molecular Biology and Evolution*.