ARTICLE

Genomic release-recapture experiment in the wild reveals within-generation polygenic selection in stickleback fish

Telma G. Laurentino1✉, Dario Moser1, Marius Roesti2, Matthias Ammann1, Anja Frey1, Fabrizia Ronco1, Benjamin Kueng1 & Daniel Berner1✉

How rapidly natural selection sorts genome-wide standing genetic variation during adaptation remains largely unstudied experimentally. Here, we present a genomic release-recapture experiment using paired threespine stickleback fish populations adapted to selectively different lake and stream habitats. First, we use pooled whole-genome sequence data from the original populations to identify hundreds of candidate genome regions likely under divergent selection between these habitats. Next, we generate F2 hybrids from the same lake-stream population pair in the laboratory and release thousands of juveniles into a natural stream habitat. Comparing the individuals surviving one year of stream selection to a reference sample of F2 hybrids allows us to detect frequency shifts across the candidate regions toward the genetic variants typical of the stream population—an experimental outcome consistent with polygenic directional selection. Our study reveals that adaptation in nature can be detected as a genome-wide signal over just a single generation.

1 Department of Environmental Sciences, Zoology, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland. 2 Institute of Ecology and Evolution, University of Bern, Bern, Switzerland. ✉email: telma.laurentino@unibas.ch; daniel.berner@unibas.ch
adaptation to novel environments can occur rapidly\textsuperscript{1–4}, with evolution in ecologically relevant phenotypes arising within a few generations\textsuperscript{5–8}. Such rapid phenotypic evolution has sometimes been linked to changes in allele frequencies at underlying genetic loci\textsuperscript{9–13}, although this generally concerns a small number of loci harboring genetic variants of large phenotypic effect\textsuperscript{13}. However, adaptation commonly involves a great number of loci spread across the genome\textsuperscript{14,15}, and how rapidly natural selection influences ecologically important loci genome-wide remains largely unexplored empirically outside prokaryotic organisms\textsuperscript{16–18}. To address this gap, we here investigate a rapid genome-wide response to selection under natural experimental conditions in threespine stickleback fish (*Gasterosteus aculeatus*).

In this organism, the postglacial colonization of freshwater by marine ancestors has led to the evolution of distinct ecotypes residing in adjacent, selectively different lake and stream habitats\textsuperscript{19–21}. Such divergent lake-stream adaptation has occurred in stickleback within the Lake Constance basin in Central Europe\textsuperscript{22,24–26}. In this system, the ecotype inhabiting Lake Constance exploits the pelagic (open-water) foraging niche, whereas multiple tributary streams harbor ecotypes with a benthic (bottom-feeding) lifestyle\textsuperscript{22,27}. This ecological diversification is mirrored by divergence between the lake and stream ecotypes in traits, such as foraging and predator defense morphology, and life history\textsuperscript{22,24,27–29}.

The lake and stream ecotypes within the Lake Constance basin are undoubtedly products of adaptive evolution, as experiments in natural streams have revealed that stream individuals consistently outperform lake individuals (and F1 lake-stream hybrids) within a single generation, and that this fitness difference has a strong genetic basis\textsuperscript{30}. At the molecular level, marker-based genomic investigations of natural populations from the Lake Constance basin have found signatures of divergent selection\textsuperscript{25,26}, for instance in the form of exceptionally strong lake-stream difference in the frequency of genetic variants in some genome regions, and indicated that this selection is highly polygenic (that is, involves a great number of genetic loci across the genome).

What is now needed to understand the mode and speed of adaptation at the genomic level is a manipulative experiment connecting rapid ecological adaptation to genome-wide changes in the frequency of genetic variants. We performed such an experiment in nature, involving (i) identifying genome-wide candidate target loci for divergent lake-stream adaptation using whole-genome sequencing in a natural lake-stream population pair; (ii) exposing a laboratory-bred, genetically mixed F2 hybrid population derived from this lake-stream pair to a natural stream habitat for one year; and (iii) assessing variant frequency shifts at the target loci in the survivors. Finding genome-wide evidence of directional polygenic selection in our field experiment, we finally use individual-based simulations to explore the underlying selection.

**Results**

**Lake-stream stickleback under polygenic divergent selection.** A key assumption underlying our study was that if natural selection drives allele frequency shifts in an experimental population within a single generation, these shifts are likely subtle and hence difficult to detect by just comparing the experimental population before and after selection. Our strategy was therefore to define genomic regions likely to be targeted by selection during the experiment a priori. To discover such regions, we focused on a single lake-stream stickleback pair\textsuperscript{22,24,25,30} residing within the Lake Constance basin (Fig. 1). From each population, we collected a large sample of individuals (\(N = 240\) and 229) in the wild.

These natural population samples were then subjected to pooled whole-genome sequencing at high read depth (210x), and the sequences were aligned to the 447 megabase (Mb) threespine stickleback genome and screened for single-nucleotide polymorphisms (SNPs). For each of the 977,723 autosomal SNPs discovered, we then quantified the magnitude of differentiation between the lake and stream population by the absolute allele frequency difference (AFD)\textsuperscript{31}.

This revealed a modest magnitude of differentiation between the natural populations (median AFD = 0.139, mean = 0.165). Numerous genomic regions, however, stood out clearly from this background level of differentiation, reaching maximal values up to 0.934 (Fig. 2; Supplementary Fig. 1). Nevertheless, no single SNP with fixed differences between the habitats was observed, which may reflect dispersal and gene flow between lake and stream stickleback within the Lake Constance basin\textsuperscript{25,26}, or that adaptation does not require the complete fixation of locally favorable alleles\textsuperscript{14,32,33}. Patterns of differentiation along chromosomes were qualitatively similar to those recovered in a previous lower-resolution genome scan for the same population pair based on reduced-representation (RAD) sequencing (compare Supplementary Fig. 1 to the ‘Lake vs. NID’ panel in Supplementary Fig. 7 from ref. 25). For instance, the SNP with the highest differentiation in the latter analysis (Fig. 4a in ref. 25) also showed extreme differentiation in the present lake-stream comparison (AFD = 0.6), and an inversion on chromosome 1 emerged as highly differentiated in both studies (Supplementary Fig. 1; Fig. 6b in ref. 25). Our comparison of the natural populations performed with whole-genome resolution clearly confirms the view that adaptive divergence between lake and stream stickleback involves differentiation in hundreds of genomic regions\textsuperscript{23}, and hence qualifies as polygenic.

From the most strongly differentiated of these regions—considered most likely to respond to selection during the release-recapture experiment, we then selected a single representative SNP (Fig. 2). These 126 total target SNPs displayed AFD values ranging from 0.477 to 0.934 (median = 0.589, mean = 0.602).

**Predicted targets of selection evolve in a single generation.** To obtain an experimental population for studying selection in action, we derived a large F2 hybrid population from our focal natural lake and stream stickleback population pair in the laboratory. Owing to random assortment and recombination, these F2 hybrids represented a genomic mixture of lake and stream ancestry (Fig. 1). From the F2 hybrid population, 3000 juvenile individuals were released into the wild at a natural stream site suitable to, but not currently inhabited by, stickleback (Fig. 1). At the same time, we took a reference sample of 510 individuals from the laboratory hybrid population to obtain a baseline of the genomic composition of the F2 hybrid population before the release. One year after the release, the F2 hybrids were recaptured in the field, recovering 37 total fish hereafter called survivors. To study evolution during the exposure to natural field conditions, both the reference sample and the survivors were subjected to whole-genome sequencing at high read depth (127 and 115x). After stringent quality filtering, these data allowed us to assess through a resampling approach whether the target SNPs predicted to be under divergent lake-stream selection showed elevated allele frequency shifts from the reference sample to the survivors compared to genome-wide neutral SNPs.

Our sequence data showed that in the reference sample characterizing the F2 hybrid population before the field release, the frequency of the stream allele (i.e., the allele showing a higher
relative frequency in the natural stream than the lake sample) at the 126 target SNPs was almost perfectly intermediate between the frequencies of the natural lake and stream population samples (Fig. 3a). Our laboratory breeding protocol thus mixed lake and stream genomes in the F2 hybrids reliably. Note, however, that our target SNPs were generally relatively far from the fixation for alternative alleles in the lake and stream populations, and that the F2 hybrids were derived from ten independent F1 hybrid families. Hence, most of the haplotype-level diversity exposed to selection in the F2 hybrids was not generated by recombination when intercrossing the F1 hybrid generation, but pre-existed in the natural populations. During the experimental period, the majority of the target SNPs (77 out of 126; 61%) exhibited an allele frequency shift in favor of the stream allele (Fig. 3a, b), a numerical imbalance unlikely to arise by chance (two-tailed binomial probability: 0.016). The median allele frequency shift across the target SNPs was 2.5% (mean 2.3%). Resampling 126 genome-wide neutral SNPs at random 9999 times and recalculating the median shift for each iteration indicated that observing an overall shift of 2.5% or greater in any direction was unlikely (two-tailed probability: 0.0173; based on the mean: 0.006) (Fig. 4). All these findings remained robust to changing analytical detail (robustness checks described in the Methods and summarized in Supplementary Fig. 2).

In genome regions inferred to be under divergent lake-stream selection based on the natural population samples, our genetically mixed lake-stream F2 hybrid fish exposed to natural stream conditions thus exhibited exceptionally large allele frequency shifts in the expected direction. This pattern is consistent with a slight response to polygenic directional selection within a single generation. Conversely, our experiment also confirms that the regions of high differentiation between the natural lake and stream populations detected in our genome scan are indeed under divergent selection.
Allele frequency changes observed in our stickleback release-recapture experiment suggest a polygenic response to directional selection within a single generation. Strong independent support for this interpretation derives from a previous experiment transplanting juvenile Lake Constance and tributary stream stickleback and their F1 hybrids into multiple, ecologically different streams, consistently and unambiguously demonstrating directional viability selection within a single generation. The experimental fish in that study were derived from laboratory lines; hence ecotype-dependent survival was largely genetically determined. In this light, there can be little doubt that the survivors recaptured in the present experiment represent a genetically non-random subset of the F2 hybrid population initially released. Also, selective shifts in the order of magnitude observed are not only plausible, but actually required to explain allele frequency shifts at candidate adaptation loci arising during phenotypic evolution from standing genetic variation over dozens of generations in wild stickleback populations.
exposed to novel habitats\cite{4,5,34,35}. Indeed, our simulations of polygenic selection over multiple generations confirm that selection coefficients appearing plausible in our single-generation experiment are compatible with the rapid allele frequency shifts observed in naturally evolving stickleback. Our genomic experiment thus suggests that adaptive allele frequency shifts can be detected over a single generation when focusing on a collective signal across many loci predicted a priori to be targets of natural selection.

Given the absence of downstream dispersal barriers in our experimental stream, a possibility worth considering is that the observed allele frequency shifts may to some extent reflect genotype-dependent dispersal. That is, experimental individuals in poor phenotypic condition due to relatively unfavorable combinations of alleles across the ecologically relevant loci may have dispersed, thus altering the genomic composition of the remaining population\cite{36-40}. This mechanism represents a (particularly effective) form of, rather than an alternative to, natural selection, because genetically based fitness differences among genotypes are a prerequisite; habitat preference mechanisms unrelated to individual fitness, whether learned or genetically determined, are not expected to systematically shift allele frequencies in genetically mixed F2 hybrids generated under laboratory conditions. Phenotype-related habitat preference has indeed been suggested in lake-stream stickleback\cite{41}, although further evidence, including on a potential genetic basis, is needed. Such information would help understand whether genome-wide responses to selection are facilitated by dispersal behavior.

A further insight, emerging from our simulations, is that a substantial within-generation polygenic response to directional selection may be plausible despite weak selection at the level of individual loci. This holds in particular when assuming that the loci affect fitness additively, as suggested by another stickleback experiment\cite{42}. In this case, the fate of a given allele is highly contingent on the allelic makeup at other loci within an individual. Specifically, selection can here become very effective when unfavorable alleles across all loci together drive the whole population toward an absolute mean fitness near zero. In this domain, many individuals actually do have zero fitness and selection is truncational. Individuals by chance carrying particularly favorable multilocus genotypes will then display an exceptionally high fitness relative to the population mean.

Our simulation finding of a substantial response to selection despite weak per-locus selection under additive fitness supports the notion that polygenic truncation selection—including departures from strict truncation in which individuals are simply ranked by multilocus genotype, allows for strong responses to selection by eliminating unfavorable alleles jointly\cite{43,44}. While we believe that truncation selection is plausible in our stickleback system, because of very high juvenile mortality measured under natural conditions\cite{30}, we emphasize the urgent need for more refined experimental information on the connection between multilocus genotype and fitness in this and other organisms. As long as this

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Observed allele frequency shift at the target SNPs compared to neutral SNPs. Distribution of the median reference-survivors frequency shift of the stream allele across 126 neutral SNPs, based on 9999 replicate resampling iterations. The neutral SNPs were required to lie within an AFD range of 0–0.1 in the comparison of the lake and stream population, and were further constrained to a minor allele frequency (MAF) spectrum matching the one observed across the target SNPs by applying a MAF threshold of 0.35 in the reference sample. The dashed blue line indicates the grand median random shift. The dashed red line gives the median frequency shift of the stream alleles observed across the 126 target SNPs during the field experiment.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Allele frequency shifts in simulations of polygenic directional selection. \textbf{a} Response to viability selection on 100 loci over a single generation for a range of per-locus selection coefficients and two different fitness schemes. For each of the 10,000 replicate simulations, the median shift of the selectively favored allele across the loci was recorded. The blue curves represent the median of these values across the replicates, and the blue bands give the associated 95 percentiles. The dashed red line indicates the median frequency shift of the stream alleles observed across the 126 target SNPs in the field experiment. \textbf{b} Response to selection on 100 loci over multiple generations, for the same two fitness schemes as in \textbf{a}. For each fitness scheme, a single selection coefficient compatible with the experimentally observed median allele frequency shift was chosen (see \textbf{a}, 0.1 and 0.01 for multiplicative and additive fitness). The blue bands display the full range across the 40 replicate simulations of the median frequency of the favorable allele across all selected loci. The black lines indicate the grand median allele frequency of the focal allele across 20 neutral loci and replicate simulations (the lines are not labeled because they largely overlap between the two fitness schemes). In both \textbf{a} and \textbf{b}, the initial frequencies of the favorable alleles at the selected loci were drawn at random from the distribution observed empirically at the target SNPs.}
\end{figure}
relationship is not better understood, it remains possible that estimates of the strength of selection at single ecologically important loci based on allele frequency changes observed over generations are inflated when numerous loci across the genome are under selection simultaneously.

Methods

Study system and experimental field site. Our investigation focuses on a single lake-stream stickleback pair (the ROM lake and the NID stream populations22,24,25,30) residing within the Lake Constance basin (Fig. 1). For the field experiment, we required a natural stream site suitable to, but not currently inhabited by, stickleback. Such a site was identified in the headwater of the stream inhabited by the NID population. Our experimental stream was formerly piped, but opened and restored one year before the experiment. To increase water volume of this small stream, and thus captive capacity, we constructed two successive shallow lagoons (Supplementary Fig. 4c). These hindered upstream dispersal and produced a total stream section of c. 50 m suitable to stickleback. Rapids downstream of the experimental site made the natural colonization of this headwater stream by stickleback highly unlikely. Accordingly, extensive minnow trapping before dam construction (April 2015) and in maximum of 6 clutches) combined haphazardly across the F1 families. The hybrid females (crossed for a maximum of three times) and 58 males (siring a in the F2 hybrid population28,30. Laboratory mortality was negligible. All pedigrees, our crossing protocol ensured that a reasonably high proportion of 6 generations are in important loci based on allele frequency changes observed over estimates of the strength of selection at single ecologically

ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-020-15657-3

To initiate the field experiment, all F2 hybrids were pooled in a single large oxygenated tank on 16 September 2015. From this pool, we randomly sampled were transported without mortality to the experimental site, and 1500 individuals individuals composition of the F2 hybrid population before the release

To evaluate whether the experimental allele frequency shifts across the target SNPs were exceptional as a whole, we generated a baseline distribution for shifts at neutral SNPs were exceptional as a whole, we generated a baseline distribution for shifts at every genomic position by applying the pileup function. Next, we determined the magnitude of genetic differentiation (quantified as absolute allele frequency difference AFD39) between the lake and stream natural populations across all genome-wide SNPs. These SNPs were required to exhibit a read depth within 100–360× in each population (thus excluding poorly sequenced and repeated regions), and a minor allele frequency (MAF) of at least 0.25 across the pool of the two populations (to ensure adequate information content39). This strategy yielded 1,009,247 SNPs across the 447 megabase (Mb) stickleback genome, thus resulting in 1,009,247 SNPs per 440 bp on average.

To define candidate regions under selection, we then identified all high-differentiation SNPs in the top 0.1 percentile of the AFD distribution. This was done chromosome-specifically (autosomes only), thus taking into account variation in baseline differentiation among chromosomes due to differences in crossover rate and hence total selection effect. Using the 0.1 percentile threshold, we performed a genomic comparison of the natural populations, considering excep-

tional strongly differentiated genome regions as putative targets of relatively long-term divergent natural selection between the lake and stream ecotype. Then we assessed if the reference-survivors differentiation in these regions was greater than expected by chance, which would offer evidence of selection on the released F2 hybrids. We started by parsing all sequence output according to barcodes, followed by alignment to the third generation assembly66 of the stickleback reference genome12 with Novoalign 3.05.00 (Novocrafts Sdn Bhd) (options: -F STDFFQ -s 1 -i 0). Using this tool, the alignments were converted to BAM format, and nucleotide counts were performed for every genomic position by applying the pileup function. Next, we determined

We thus considered it crucial to identify genomic regions a priori in which allele frequency differen-
tion SNPs thus identified as belonging to the same genomic region. From each independent
to allow separating potential signatures of selection from background stochasticity. We thus considered it crucial to identify genomic regions a priori in which allele frequency differen-
tion SNPs thus identified as belonging to the same genomic region. From each independent
to allow separating potential signatures of selection from background stochasticity. We thus considered it crucial to identify genomic regions a priori in which allele frequency differen-
tion SNPs thus identified as belonging to the same genomic region. From each independent
to allow separating potential signatures of selection from background stochasticity. We thus considered it crucial to identify genomic regions a priori in which allele frequency differen-
tion SNPs thus identified as belonging to the same genomic region. From each independent
subset was then further restricted by retaining only those SNPs exhibiting a MAF of at least 0.35 in the reference sample. The rationale for this highly stringent MAF filtering was that the magnitude of genetic differentiation between samples is contingent on the MAF across their pool; markers displaying strong differentiation necessarily also show a high MAF, whereas low differentiation is possible across a broader MAF range. Accordingly, our target SNPs—representing high-differentiation markers—showed relatively high MAFs in the reference sample (median: 0.426; Supplementary Fig. 5). With a threshold of 0.35, the MAF spectrum of our neutral SNPs closely approximated the MAF distribution observed across the target SNPs (median: 0.426; Supplementary Fig. 6). Stringent MAF filtering thus precluded that a difference in the magnitude of experimental shifts at the target versus neutral SNPs was an artifact caused by different levels of genetic diversity between these SNP classes.

From the MAF-filtered neutral SNPs, we then drew (with replacement) 9999 random samples of SNPs equal in size to the number of target SNPs (126). We here applied exactly the same standards as for the target SNPs: a physical spacing of at least 50 kb between SNPs, a minimum read depth of 70× in both the reference and survivor sample, and nucleotide counts from at least 35 survivors. Characterizing the SNP-specific reference-survivor allele frequency shifts for each of these samples finally allowed us to evaluate if the median shift observed across the target SNPs was uncommon relative to the distribution of median shifts across the neutral SNPs (throughout our study, we consider the median the most appropriate statistic of location, but additionally report the mean for key results). We emphasize that this strategy investigated a global signature of selection across the genome only; given the low expected signal-to-noise ratio, we made no attempt to infer selection on individual SNPs or genome regions.

Robustness checks. To assess the validity of the above statistical protocol to investigate selection during our field experiment, we implemented several alternative parameterizations of the base model (Supplementary Fig. 2a–c). First, while a MAF threshold of 0.35 in the reference sample was assumed to be uniform across all loci for a chosen value. The second individual frequencies of the stream allele observed at the 126 target SNPs in the reference site in the beginning of the experiment (as there was no downstream dispersal possibility that a substantial fraction of individuals may have left the experimental population). From the MAF-filtered neutral SNPs, we then drew (with replacement) 9999 random samples of SNPs equal in size to the number of target SNPs (126). We here applied exactly the same standards as for the target SNPs: a physical spacing of at least 50 kb between SNPs, a minimum read depth of 70× in both the reference and survivor sample, and nucleotide counts from at least 35 survivors. Characterizing the SNP-specific reference-survivor allele frequency shifts for each of these samples finally allowed us to evaluate if the median shift observed across the target SNPs was uncommon relative to the distribution of median shifts across the neutral SNPs (throughout our study, we consider the median the most appropriate statistic of location, but additionally report the mean for key results). We emphasize that this strategy investigated a global signature of selection across the genome only; given the low expected signal-to-noise ratio, we made no attempt to infer selection on individual SNPs or genome regions.

Simulations—single generation. To develop a sense for the selection strength at individual loci required to produce a frequency shift in the order of the one observed here across the target SNPs, we performed forward simulations of polygenetic directional viability selection over a single generation. Our base model involved a population of 1000 diploid individuals under selection at 100 independent (unlinked), biallelic, codominant loci. We chose a population size lower than the number of individuals actually released, thus taking into account the possibility that a substantial fraction of individuals may have left the experimental site in the beginning of the experiment (as there was no downstream dispersal barrier). Initial frequencies of the favorable alleles were drawn at random from the frequencies of the stream allele observed at the 126 target SNPs in the reference sample, and individual diploid multilocus genotypes were constructed according to these frequencies. Selection was modeled in analogy to our empirical experiment by drawing 40 individuals as survivors, the survival probability being a stochastic function of an individual’s relative multilocus genotype fitness. Because the true link between multilocus genotype and fitness is not known for this stickleback system (or any other organism), we explored two distinct fitness functions. The first fitness was standard multiplicative fitness, defined as \( 1 - s^n \), where \( s \) is the selection coefficient and \( n \) is the total number of unfavorable alleles across all loci within an individual (e.g. refs. 54,55). Key features of this fitness function are that the effect of a given allele on an individual’s fitness is independent from its multilocus genetic background, and that fitness always remains positive as long as \( s < 1 \). The second fitness function used was additive, \( 1 - s \cdot n \), with each unfavorable allele reducing an individual’s fitness by \( s \) (54,56). With this latter fitness function, the effect of a given allele is contingent on the genetic background, thus allowing for interactions among different fitness values, which in turn was a direct function of the number of unfavorable alleles across the genome. Individuals were hermaphrodites and were allowed to be drawn mating multiple times, thus allowing variation among individuals in reproductive success. The transmission of alleles from parents to offspring occurred in a standard Mendelian way. With additive fitness \( (s = 0.01) \), a small proportion (~5%) of the population initially displayed negative fitness (which was set to zero), hence selection was truncational in the beginning.

Evolution was allowed for 500 generations. In each generation, we recorded the median frequency of the locally favored allele across all selected loci (for simplicity, we made no effort to additionally characterize statistics of dispersion in allele frequencies across loci). Allele frequency changes at the neutral loci were tracked analogously by initially defining one of the two alleles as the focal one. For each fitness scheme (and locus number), we performed 40 replicate simulations. All analyses, simulations and data visualization were performed in R version 3.6.0 (57).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
All raw whole-genome sequence data are available from the NCBI sequence read archive (Supplementary Data 1). The high-diploid individuals (K = 20,000 produced very similar results, not presented), and individuals were recruited for mating with probabilities dependent on their relative fitness (which was set to zero), hence selection was truncational in the beginning. Evolution was allowed for 500 generations. In each generation, we recorded the median frequency of the locally favored allele across all selected loci (for simplicity, we made no effort to additionally characterize statistics of dispersion in allele frequencies across loci). Allele frequency changes at the neutral loci were tracked analogously by initially defining one of the two alleles as the focal one. For each fitness scheme (and locus number), we performed 40 replicate simulations. All analyses, simulations and data visualization were performed in R version 3.6.0 (57).

Code availability
All analytical code underlying this work is provided as Supplementary Software.

Received: 18 September 2019; Accepted: 19 March 2020; Published online: 21 April 2020

References
1. Reznick, D. N. & Ghalambor, C. K. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. Genetica 112–113, 183–198 (2001).
2. Hendry, A. P. & Kimmins, M. T. Perspective: the pace of modern life: measuring rates of contemporary microevolution. *Evolution* 53, 1637 (1999).

3. Deo, R. N. & Schuster, S. A. How fast is fast? Eco-evolutionary dynamics and rates of change in populations and phenotypes. *Evol. Ecol.* 6, 573–581 (2016).

4. Lescak, E. A. et al. Evolution of stickleback in 50 years on earthquake-uplifted islands. *Proc. Natl Acad. Sci. USA* 112, E7204–E7212 (2015).

5. Bell, M. A., Aguirre, W. E. & Buck, N. J. Twelve years of contemporary armament evolution in a three-spined stickleback population. *Evolution* 58, 814–824 (2004).

6. Cook, I. M., Grant, B. S., Saccheri, I. J. & Mallet, J. Selective bird predation on the peppered moth: the last experiment of Major Majerus. *Biol. Lett.* 8, 609–612 (2012).

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

8. Grant, P. R. & Grant, B. R. Unpredictable evolution a 30-year study Darwin Finches. *Science* 296, 707–712 (2002).

9. Manceau, M., Domingues, V. S., Linnen, C. R., Rosenblum, E. B. & Hoekstra, H. E. Convergence in pigmentation at multiple levels: mutations, genes and function. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 2439–2450 (2010).

10. Reid, N. M. et al. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* 334, 1305–1309 (2016).

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

12. Jones, F. C. et al. The genomic basis of adaptive evolution in threespine stickleback. *PLoS Genet.* 10, e1004696 (2014).

13. Marques, D. A., Jones, F. C., Di Palma, F., Kingsley, D. M. & Reimchen, T. E. Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. *Nat. Ecol. Evol.* 2, 1130–1140 (2018).

14. de Meulis, T., Michalakis, Y., Renaud, F. & Olivieri, I. Polymorphism in heterogeneous environments: evolution of habitat selection and sympatric speciation: soft and hard selection models. *Evol. Ecol. J.* 7, 175–198 (1993).

15. Armsworth, P. R. & Roughgarden, J. E. The structure of clines with fitness-dependent dispersal. *Am. Nat.* 172, 648–657 (2008).

16. Edeleah, P., Serpielski, A. M. & Clobert, J. Matching habitat choice causes directed gene flow: a neglected dimension in evolution and ecology. *Evol.* 62, 2462–2472 (2008).

17. Clobert, J., Le Galliard, J. F., Cote, J., Meylan, S. & Massot, M. Inferred dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Evol. Lett.* 12, 197–209 (2009).

18. Berner, D. & Thübert-Plante, X. How mechanisms of habitat preference evolve and promote divergence with gene flow. *Evol. Biol.* 28, 1641–1655 (2015).

19. Bolnick, D. I. et al. Phenotype-dependent native habitat preference facilitates divergence between parapatric lake and stream stickleback. *Evolution* 63, 2004–2016 (2009).

20. Attezrad, M. E. et al. Genetics of ecological divergence during speciation. *Nature* 511, 307–311 (2014).

21. Kimura, M. & Crow, J. F. Effect of overall phenotypic selection on genetic change at individual loci. *Proc. Natl Acad. Sci. USA* 75, 6188–6171 (1978).

22. Crow, J. F. & Kimura, M. Efficiency of truncation selection. *Proc. Natl Acad. Sci. USA* 76, 396–399 (1979).

23. Berner, D. et al. Sexual isolation promotes divergence between parapatric lake and stream stickleback. *Evol. Biol.* 38, 401–411 (2017).

24. Deagle, B. E. et al. Population genomics of parallel phenotypic evolution in stickleback across stream–lake ecological transitions. *Proc. R. Soc. B Biol. Sci.* 279, 1277–1286 (2012).

25. Moser, D., Roesti, M. & Berner, D. Repeated lake–stream divergence in stickleback life history within a Central European Lake Basin. *PloS ONE* 7, e50620 (2012).

26. Ravinet, M., Prodohl, P. A. & Harrod, C. Parallel and nonparallel ecological, morphological and genetic divergence in lake–stream stickleback from a single catchment. *J. Evol. Biol.* 26, 186–201 (2013).

27. Berner, D., Roesti, M., Hendry, A. P. & Salzburger, W. Constraints on speciation suggested by comparing lake–stream stickleback divergence across two continents. *Mol. Ecol.* 19, 4963–4978 (2010).

28. Roesti, M., Kuehn, B., Moser, D. & Berner, D. The genomics of ecological varicance in threespine stickleback fish. *Nat. Commun.* 6, 1–14 (2015).

29. Marques, D. A. et al. Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genet.* 12, 1–34 (2016).

30. Lucek, K., Sivasundar, A. & Seehausen, O. Evidence of adaptive evolutionary divergence during biological invasion. *PloS ONE* 7, e93777 (2012).

31. Moser, D., Kuehn, B. & Berner, D. Lake–stream divergence in stickleback life history: a plastic response to trophic niche differentiation? *Evol. Biol.* 42, 328–338 (2015).

32. Lucek, K., Sivasundar, A. & Seehausen, O. Disentangling the role of phenotypic plasticity and genetic divergence in contemporary ecotype formation during a biological invasion. *Evolution* 68, 2619–2632 (2014).

33. Moser, D., Frey, A. & Berner, D. Fitness differences between parapatric lake and stream stickleback revealed by a field transplant. *J. Evol. Biol.* 29, 711–719 (2016).

34. Berner, D. Allele frequency AFD—a intuitive alternative to FST for quantifying genetic population differentiation. *Genes* 10, 1–13 (2019).

35. Kremer, A. & Le Corre, V. Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity* 108, 376–385 (2012).

36. Latta, R. G. Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *Am. Nat.* 151, 283–292 (1998).
Author contributions
Project supervision: D.B.; funding acquisition: D.B. and T.G.L.; experimental design: D.B., D.M. and T.G.L.; fish husbandry: D.M., M.A., A.F. and D.B.; field work: D.M., M.R., B.K., F.R., D.B., T.G.L. and M.A.; wet lab: T.G.L.; data analysis: T.G.L. and D.B.; visualization: T.G.L. and D.B.; coding: D.B. and T.G.L.; writing: T.G.L. and D.B., with feedback from: M.R. and F.R.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-15657-3.

Correspondence and requests for materials should be addressed to T.G.L. or D.B.

Peer review information Nature Communications thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020