LETTER

doi:10.1038/nature22351

Frequency dependence limits divergent evolution by favouring rare immigrants over residents

Daniel I. Bolnick[1] & William E. Stutz[1]†

Two distinct forms of natural selection promote adaptive biological diversity. Divergent selection occurs when different environments favour different phenotypes, leading to increased differences between populations[1]. Negative frequency-dependent selection occurs when rare variants within a population are favoured over common ones[2], increasing diversity within populations[3]. These two diversifying forces promote genetic variation at different spatial scales, and may act in opposition, but their relative effects remain unclear because they are rarely measured concurrently. Here we show that negative frequency-dependent selection within populations can favor rare immigrants over locally adapted residents. We reciprocally transplanted lake and stream ecotypes of threespine stickleback[4] into lake and stream habitats, while manipulating the relative abundance of residents versus immigrants. We found negative frequency-dependence: survival was highest for the locally rare ecotype, rather than natives. Also, individuals with locally rare major histocompatibility complex (MHC) class IIb genotypes were infected by fewer parasites. This negative frequency-dependent selection will tend to favour rare immigrants over common residents, amplifying the effect of migration and undermining the efficacy of divergent natural selection to drive population differences. The only signal of divergent selection was a tendency for foreign fish to have higher parasite loads than residents, after controlling for MHC genotype rarity. Frequency-dependent ecological interactions have long been thought to promote speciation. Our results suggest a more nuanced view in which negative frequency dependence alters the fate of migrants to promote or constrain evolutionary divergence between populations.

Divergent selection (DS) is thought to be a nearly ubiquitous evolutionary force[5]. Geographic variation in biotic and abiotic variables is widespread, imposing selection for different traits in different habitats and driving rapid genetic divergence between populations[6], local adaptation[7] and speciation[8]. DS promotes variation among populations at the expense of within-population polymorphism. DS may ultimately lead to ecological speciation in allopatry or parapatry[8]. However, this evolutionary divergence is often constrained by gene flow among populations that homogenizes allele frequencies[9]. In contrast, negative frequency-dependent selection (NFDS) requires very specific types of ecological interactions[10] and so is widely thought to be less common than DS. When NFDS does occur, it promotes polymorphism within populations and is essential to most models of sympatric speciation[11].

DS and NFDS might act concurrently[12]. For instance, if the rarity of immigrants allows them to exploit locally under-used resources, competitive release may partly compensate for any maladaptation to their new habitat. By favouring rare immigrants, NFDS would amplify the effective migration rate[13]. For populations subject to DS, this enhanced migration will erode between-population divergence[9] and oppose DS (Extended Data Fig. 1). Yet we remain largely ignorant of the relative effects of DS and NFDS because they act at different geographic scales (between and within populations), and so are studied separately using different experimental designs. DS is usually measured by reciprocal transplant experiments that permute genotypes across environments to test for a home-field advantage[7]. Such transplants typically mix genotypes in equal proportions, ignoring the rarity of migrants[14]. In contrast, NFDS is measured by manipulating relative abundance within a single habitat, to test whether rarity confers an advantage[15,16].

Here, we present the results of a field experiment designed to partition the concurrent effects of DS and NFDS acting on adjacent populations of lake and stream threespine stickleback (*Gasterosteus aculeatus*). Numerous studies have reported DS acting between ecologically disparate populations of stickleback[17], including parapatric lake and stream populations whose divergence is constrained by migration[4]. Concurrently, NFDS acts within stickleback populations because rare variants experience weaker resource competition[15,18]. We therefore hypothesized that the rarity of immigrants might compensate for their maladaptation. To test this conjecture, we reciprocally transplanted lake and stream stickleback between adjoining habitats, while factorially manipulating the relative abundance of migrants (the native:migrant ratio per cage was 1:2 or 2:1; Extended Data Fig. 2).

Survival data revealed no evidence of DS (Table 1). By the end of the 6-week experiment, native survival (58%) was not higher than immigrant survival (68%; $\chi^2 = 2.57$, $P = 0.109$). Instead, directional selection favoured stream natives over lake natives in both habitats (80% versus 46% survival for stream versus lake fish; $P < 0.0001$; statistical results in Extended Data Table 1). The advantage of the stream fish may be a legacy of developing in the stream, which appears to be the more productive habitat; both ecotypes survived at a higher rate (77%) in the stream than in the lake (49%; $P < 0.0001$). The origin and destination effects were additive (interaction $P = 0.4033$), providing no evidence for local adaptation and DS. These results are consistent with prior studies of lake–stream stickleback, which typically found directional selection rather than DS, or no selection despite sometimes high mortality[19] (reviewed in Supplementary Information section 1). Survival was independent of pre-experiment mass (controlling for fish origin), and sex (see below). An important caveat is that we measured viability selection only on adult fish in only one year; selection may act differently on juveniles or in other years.

We observed NFDS within cages: minority ecotypes had 1.25-fold higher survival than majority ecotypes (Fig. 1; Extended Data Fig. 3; main effect of minority status $P = 0.013$; statistical details in Extended Data Table 1). The benefit of rarity was strongest for native stream

Table 1 | Survival rates for each combination of stickleback source and destination habitat

From	To lake cages	To stream cages
Lake natives	26.6%	65.0%
Stream natives	71.2%	88.3%

The survival rate is the percentage of 60 fish who survived for each source–destination combination. Within each habitat, fish were randomized among 40 cages, with either 1 native and 2 immigrants, or 2 natives and 1 immigrant ($n = 3$ fish per cage; $n = 240$ total).

[1]Department of Integrative Biology, University of Texas at Austin, Austin, Texas 78712, USA. †Present address: Office of Institutional Research, Western Michigan University, Kalamazoo, Michigan 49008, USA.

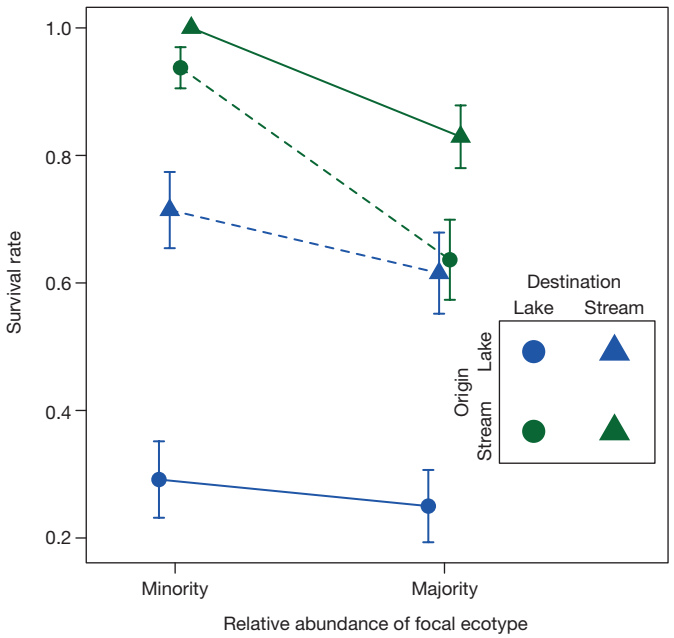


Figure 1 | Survival rate depended on fish origin, transplant destination and relative abundance. Points and bars represent group means ±1 standard error. Solid lines indicate resident ecotypes; dashed lines indicate immigrants. Extended Data Figure 2 diagrams the experimental design. Extended Data Table 1 provides statistics. There is a significant main effect of relative abundance ($P = 0.013$). Rare ecotype survival was on average 1.26-fold higher than common ecotype survival (95% confidence interval: 1.03 to 1.49; fold changes for each source–destination combination are 1.21, 1.47, 1.17 and 1.17 for stream–stream, stream–lake, lake–stream, and lake–lake respectively, the latter two with $P > 0.1$). In Supplementary Information section 1, we discuss the directional selection favouring stream fish. Sample sizes are in Table 1. Symbol colours denote fish origin, and symbol shapes indicate fish transplant destination (inset).

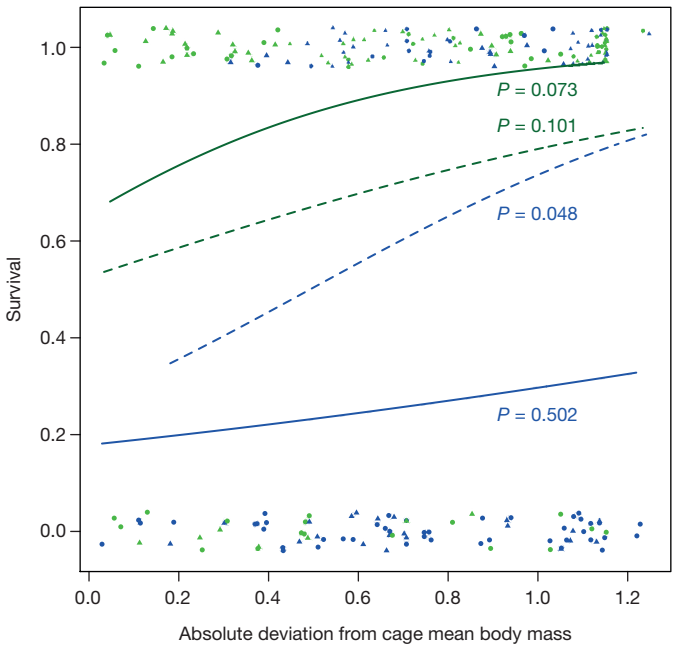


Figure 2 | Survival is higher for atypically sized individuals within cages (a statistically significant main effect of majority/minority status). GLM fitting trends are indicated for stream natives (green lines) and lake natives (blue lines), either in their native habitat (solid lines) or as immigrants into the adjoining habitat (dashed lines). Symbols indicate individuals that survived (1) or died (0). Vertical jitter has been added to separate points. Symbol colours denote fish origin, and symbol shapes indicate fish transplant destination, as in the inset to Fig. 1. Sample sizes are listed in Table 1. Statistical results are summarized in Extended Data Table 1 (model E).

fish placed in either habitat (origin × minority interaction, $P = 0.0367$). This 'neighbour-dependent selection'[20] may result from resource competition among individuals within cages[21]. Intraspecific competition has previously been shown to reduce stickleback fitness[18,21]. This competition is most severe for individuals with common trophic traits (for example, average size), whereas individuals with rare phenotypes use atypical resources and have relatively few competitors[15,18]. Our experiment revealed similar trait-dependent NFDS. Survival was lower for average-sized individuals within a cage than for individuals whose initial mass was much larger or smaller than their cage mean (Fig. 2; $P = 0.0195$). However, greater mass itself was not associated with higher survival ($P = 0.488$ in a general linear model (GLM) including origin and destination effects). Body size deviations were partly confounded with minority status (minority and majority fish average weights were, respectively, 0.91 and 0.66 standard deviations from their cage mean; Student's t-test $= 5.85$, $P < 0.0001$). Multiple regression including both minority status and body size deviations indicated that NFDS can be attributed to the rare-size advantage (Extended Data Table 1). We infer that rare ecotypes experience weaker resource competition because their atypical phenotype (size) confers an atypical diet[22], and are therefore favoured by NFDS.

Host–parasite interactions can also cause DS or NFDS. Lake and stream stickleback harbour distinct parasite communities (Extended Data Fig. 4)[23]. If they experience DS to resist their respective parasites, immigrants should carry heavier parasite loads than locally adapted residents. However, within each habitat, selection favours parasites that better exploit locally common host genotypes[24]. The resulting NFDS on hosts may favour rare immigrants, opposing the effect of DS. We did not expect to find this NFDS within our cages, because this selection results from across-generation parasite evolution to target common hosts. But we can test for habitat-wide NFDS by contrasting infection rates of fish with locally common versus rare genotypes. The proteins produced by genes in MHC class IIb (these are orthologues of the HLA genes in humans) play a key part in recognizing parasite antigens, and are sometimes subject to DS or NFDS[24,25]. Allele frequencies of *MHCIIb* genes differ between lake and stream stickleback (Extended Data Fig. 5), but there is sufficient admixture to sustain lake alleles in stream fish and vice versa. As a result, we can measure DS and NFDS by statistically partitioning how parasite infection load depends on fish ecotype (resident/immigrant, DS) versus its *MHCIIb* genotype (locally rare/common, NFDS). We sequenced the hypervariable exon 2 of all *MHCIIb* gene paralogues in the stickleback genome, and used discriminant function analysis to obtain a quantitative measure of whether fish have locally rare or common *MHCIIb* genotypes in their transplant destination (see Methods). We tested whether total parasite infection load (presumed to reduce fitness) covaried with fish origin or with *MHCIIb* discriminant function analysis score. After controlling for *MHCIIb* genotype, immigrant fish were more heavily infected than native ecotypes (origin effect, $P = 0.024$), revealing *MHCIIb*-independent local adaptation and DS. Simultaneously, foreign (locally rare) *MHCIIb* genotypes were less infected than locally common MHC genotypes ($P = 0.005$; Fig. 3), revealing that NFDS within habitats selects against maladapted resident *MHCIIb* (statistics in Extended Data Table 2). The deleterious effect of native *MHCIIb* was stronger in the lake habitat (destination × *MHCIIb* interaction $P = 0.008$), where parasite diversity is higher overall. The infection rates were lowest (and presumably fitness was highest) in resident fish with foreign *MHCIIb* alleles. We conclude that *MHCIIb*-independent DS is opposed by *MHCIIb*-dependent NFDS. These opposing effects are only detectable in statistical models considering both terms; neither is significant in univariate regression.

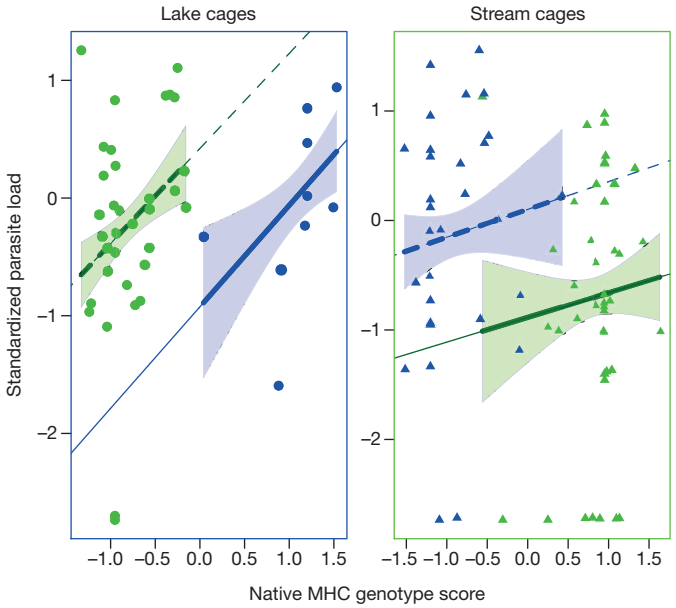


Figure 3 | The effect of native/foreign status and *MHCIIb* genotype on standardized parasite infection load. The native *MHC* habitat score measures the extent to which individuals have a *MHCIIb* genotype that is typical (positive values) or atypical (negative values) in their cage habitat (see Methods). That is, positive numbers denote fish with more distinctively lake *MHCIIb* in the lake cages, and fish with stream *MHCIIb* in stream cages. Parasite load is higher for fish with more native MHC genotypes within each habitat, but also higher for immigrants than residents after controlling for MHC genotype (difference in intercepts). GLM trendlines and ±1 standard error intervals for the regression slopes (shading) are plotted for each combination of origin (colour, see Fig. 1) and destination (panels). Solid (or dashed) lines distinguish resident (or immigrant) fish. Sample sizes are the number of survivors in Table 1. Statistical support is summarized in Extended Data Table 2.

We are able to dismiss some other potential causes of NFDS. Intra- or inter-sexual selection may favour rare variants, for instance if males with locally common nuptial colours are disproportionately attacked by rivals[26]. Within our cages, such NFDS should cause sex-biased mortality. But we observed no such sex bias among survivors, comparing majority versus minority fish ($\chi^2 = 0.55$, $P = 0.455$), or natives versus immigrants ($\chi^2 = 0.56$, $P = 0.453$). Predators can drive NFDS by evolving (or learning) to detect or capture locally common prey[27]. But this NFDS should act at the habitat-wide scale and most predators (piscivorous birds, trout) were excluded by the cages.

Migrants between habitats suffer a variety of disadvantages[9] when it comes to resource acquisition, parasite resistance, predator avoidance or mate attraction[25,28]. Yet migrants are also usually rare, and may thus benefit from NFDS. Our results confirm that NFDS acts within habitats, at multiple spatial scales: neighbour-dependent competition[20] within cages, and parasite-driven NFDS targeting locally common *MHCIIb* genotypes at the habitat-wide scale. Intraspecific competition and parasitism are both important sources of selection in natural populations[29,30] and both can favour rare phenotypes or genotypes. In contrast, the only signal of DS that we detected was higher parasite infection for immigrants than residents. This DS acted concurrently with *MHCIIb*-dependent NFDS, with the result that neither form of selection was detectable without considering the other factor. This highlights the potential for DS and NFDS to act simultaneously and in opposition, and have comparable effect sizes. However, we emphasize that additional unmeasured NFDS or DS could act on juveniles or during courtship and breeding.

By rescuing the fitness of immigrants, NFDS amplifies the rate of gene flow between populations. Our model suggests that NFDS can be almost as effective as migration itself at constraining between-population divergence in response to DS (Extended Data Fig. 1). NFDS constrains divergence most effectively when migration is uncommon; at higher migration rates the immigrants cease to be rare and lose the benefit of NFDS. The model thus confirms an intuitive but often overlooked point: NFDS facilitates migration and thus undermines the response of populations to DS, as we saw with parasite infection rates. However, our survival data are more puzzling at first glance: directional selection should favour stream fish in both habitats, so why are these populations divergent at all? Our simulations suggest an intriguing hypothesis that could resolve this conundrum. Although directional selection will tend to favour the complete fixation of a single genotype in all habitats, NFDS can prevent this by retaining the less-adapted allele at some low equilibrium frequency at which the benefit of rarity compensates for its otherwise lower fitness. Importantly, if this equilibrium differs between habitats, the populations will reach different allele frequencies. Thus, NFDS can promote between-population genetic differences when directional selection and migration would otherwise homogenize the populations (Extended Data Figs 6 and 7; see also extended discussion in Supplementary Information section 1).

Population divergence therefore reflects a three-way balance between DS on one side, opposed by an interaction between migration and NFDS on the other. Ecological interactions causing NFDS may thus inhibit ecological speciation in response to DS, by amplifying constraints from gene flow, especially when gene flow is rare. In other settings, the ecological interactions causing NFDS can promote population divergence when it would not otherwise occur (for example, directional selection, Extended Data Fig. 6). These results present a counterpoint to other work claiming that NFDS promotes speciation via adaptive branching of single populations[11] or suggesting that ecologically driven DS can drive speciation in parapatry[5].

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 December 2016; accepted 7 April 2017.
Published online 31 May 2017.

1. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, 1930).
2. Ayala, F. J. & Campbell, C. A. Frequency-dependent selection. *Annu. Rev. Ecol. Syst.* **5,** 115–138 (1974).
3. Kopp, M. & Hermisson, J. The evolution of genetic architecture under frequency-dependent disruptive selection. *Evolution* **60,** 1537–1550 (2006).
4. Hendry, A. P. & Taylor, E. B. How much of the variation in adaptive divergence can be explained by gene flow? An evaluation using lake-stream stickleback pairs. *Evolution* **58,** 2319–2331 (2004).
5. Shafer, A. B. A. & Wolf, J. B. W. Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology. *Ecol. Lett.* **16,** 940–950 (2013).
6. Savolainen, O., Lascoux, M. & Merilä, J. Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14,** 807–820 (2013).
7. Kawecki, T. J. & Ebert, D. Conceptual issues in local adaptation. *Ecol. Lett.* **7,** 1225–1241 (2004).
8. Rundle, H. D. & Nosil, P. Ecological speciation. *Ecol. Lett.* **8,** 336–352 (2005).
9. Lenormand, T. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* **17,** 183–189 (2002).
10. Doebeli, M. & Dieckmann, U. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *Am. Nat.* **156,** S77–S101 (2000).
11. Bolnick, D. I. & Fitzpatrick, B. M. Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* **38,** 459–487 (2007).
12. Gray, S. M. & McKinnon, J. S. Linking color polymorphism maintenance and speciation. *Trends Ecol. Evol.* **22,** 71–79 (2007).
13. Ebert, D. *et al.* A selective advantage to immigrant genes in a *Daphnia* metapopulation. *Science* **295,** 485–488 (2002).
14. Hereford, J. A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* **173,** 579–588 (2009).
15. Schluter, D. Frequency dependent natural selection during character displacement in sticklebacks. *Evolution* **57,** 1142–1150 (2003).
16. Antonovics, J. & Kareiva, P. Frequency-dependent selection and competition: empirical approaches. *Phil. Trans. R. Soc. Ser. B* **319,** 601–613 (1988).
17. Hendry, A. P., Bolnick, D. I., Berner, D. & Peichel, C. L. Along the speciation continuum in sticklebacks. *J. Fish Biol.* **75,** 2000–2036 (2009).
18. Bolnick, D. I. Can intraspecific competition drive disruptive selection? An experimental test in natural populations of sticklebacks. *Evolution* **58,** 608–618 (2004).

19. Wegner, K. M., Kalbe, M., Milinski, M. & Reusch, T. B. Mortality selection during the 2003 European heat wave in three-spined sticklebacks: effects of parasites and MHC genotype. *BMC Evol. Biol.* **8,** 124 (2008).
20. Vasseur, D. A., Amarasekare, P., Rudolf, V. H. W. & Levine, J. M. Eco-evolutionary dynamics enable coexistence via neighbor-dependent selection. *Am. Nat.* **178,** E96–E109 (2011).
21. Svanbäck, R. & Bolnick, D. I. Intraspecific competition drives increased resource use diversity within a natural population. *Proc. R. Soc. B* **274,** 839–844 (2007).
22. Ingram, T., Stutz, W. E. & Bolnick, D. I. Does intraspecific size variation in a predator affect its diet diversity and top-down control of prey? *PLoS One* **6,** e20782 (2011).
23. Rauch, G., Kalbe, M. & Reusch, T. B. H. Relative importance of MHC and genetic background for parasite load in a field experiment. *Evol. Ecol. Res.* **8,** 373–386 (2006).
24. Spurgin, L. G. & Richardson, D. S. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. R. Soc. B* **277,** 979–988 (2010).
25. Eizaguirre, C., Lenz, T. L., Kalbe, M. & Milinski, M. Divergent selection on locally adapted Major Histocompatibility Complex immune genes experimentally proven in the field. *Ecol. Lett.* **15,** 723–731 (2012).
26. Bolnick, D. I., Hendrix, K. A., Jordan, L. A., Veen, T. & Brock, C. D. Intruder color and light environment jointly determine how nesting male stickleback respond to simulated territorial intrusions. *Biol. Lett.* **12,** 20160467 (2016).
27. Hughes, R. N. & Croy, M. I. An experimental analysis of frequency-dependent predation (switching) in the 15-spined stickleback, *Spinachia spinachia*. *J. Anim. Ecol.* **62,** 341–352 (1993).
28. Nosil, P. Reproductive isolation caused by visual predation on migrants between divergent environments. *Proc. R. Soc. B* **271,** 1521–1528 (2004).
29. Gurevitch, J., Morrow, L. L., Wallace, A. & Walsh, J. S. A meta-analysis of competition in field experiments. *Am. Nat.* **140,** 539–572 (1992).
30. Fumagalli, M. *et al.* Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genet.* **7,** e1002355 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank L. Snowberg, K. Ballare, G. Calabrese, J. Heiling, K. Jiang, C. Brock and M. Bright for field assistance. This project was supported by grants to D.I.B. from the Howard Hughes Medical Institute and the NSF (DEB-1144773 and DEB-1456462) and grants to W.E.S. from the EEB Graduate Program at the University of Texas.

Author Contributions The experiment was jointly designed by W.E.S. and D.I.B. The field research, sequencing, and specimen measurements were conducted by W.E.S. Both authors contributed to data analysis. The manuscript was written by D.I.B. with input from W.E.S.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to D.I.B. (danbolnick@austin.utexas.edu).

Reviewer Information *Nature* thanks R. Calsbeek, J. McKinnon and D. Reznick for their contribution to the peer review of this work.

METHODS

Focal populations. We performed a reciprocal transplant experiment exchanging stickleback between Roberts Lake (Vancouver Island, British Columbia, 50.2265° N and 125.5531° W) and its outlet stream (50.2283° N and 125.5562° W). These parapatric populations differ at genome-wide single nucleotide polymorphisms (SNPs)[31], and exhibit divergent allele frequencies at *MHCIIb* loci[32], and for a variety of morphological traits[33,34]. These genetic and phenotypic differences have been claimed to be adaptive responses to DS arising from distinct prey and parasite communities and flow regimes of the populations' respective habitats[31,34]. Collection and animal handling were approved by the University of Texas Institutional Animal Use and Care Committee (Protocol 07-032201), and a Scientific Fish Collection Permit from the Ministry of the Environment of British Columbia (NA07-32612).

Reciprocal transplant experiment. In April 2011 we constructed 40 cages in Roberts Lake, and 40 cages in the stream. Each cage was constructed by embedding a vertical 1-m diameter cylinder of galvanized 0.25-cm steel mesh into the lake or stream substrate. The mesh cylinder was held upright by rebar hammered into the substrate. Cages were 1.5 m tall, installed along the lake or stream shore in water approximately 1 m deep. Each cylinder was open at both ends. The open bottom of the cage was embedded into the substrate to give stickleback free access to the substrate and any invertebrates on it. A fine seine-net skirt was sewn to the bottom of each cage and buried to prevent stickleback from escaping. The 0.25-cm wire mesh allowed unimpeded water flow and entry of invertebrates upon which stickleback feed. The lake cages were confined to the shallow littoral zone and did not include deep pelagic habitat. However, stickleback in Roberts Lake are typically found in shallow water (0.1–2 m deep) in May–July, because this habitat is where they nest and breed and forage (D.I.B., personal observation). Previous experiments with shallow cages confirmed that stickleback within the cages have diets that closely resemble those of un-enclosed fish captured from the same lake[18,21,35], and immune gene expression profiles are similar between caged and wild stickleback, suggesting there is no undue stress[36].

Stream cages were placed 100 m downstream from the lake along both banks of the stream (2–3 m across) in a pool with low but non-zero flow. Most stream stickleback inhabit pool rather than riffle microenvironments, so this location represents a typical stream habitat. We placed cages 100 m downstream from the lake because this site naturally receives immigrants from the lake (about 10% of captured fish in this section of stream were estimated to be immigrants[31]), so this is the habitat that most immigrants would experience. Furthermore, this location was ethically preferable because any accidental damage to our cages (which did not occur) would have contributed only to natural migration without disrupting the more substantial population divergence observed further downstream. Likewise, our lake cages were placed roughly 100 m away from the outlet, at a site that does receive a few first-generation immigrants from the stream[31].

We used minnow traps to capture adult stickleback from the lake and the stream in early June 2011. We lightly anaesthetized each captured fish with MS-222 and recorded its weight and standard length (measured with digital calipers). We cut off a spine clip from each fish, stored in ethanol for subsequent DNA extraction. Fish were allowed to re-acclimate alone in fresh water before transfer to cages. Any fish that did not quickly and fully recover from spine clipping were not used in the experiment. Each of the 80 cages received three fish by 6 June 2011. Half the cages received a 2:1 ratio of lake:stream fish, and half received a 1:2 ratio, with 20 cages of each ratio within each habitat (Extended Data Fig. 2). For a given source–destination combination, we randomly assigned animals to cage and minority/majority status. Within each trio of fish, each individual received a unique spine clip pattern (first, second, or both dorsal spines) to identify recaptured individuals. Stickleback from these populations exhibit subtle and unreliable sexual dimorphism even in the breeding season. Therefore, the initial sex ratio entering each cage was not known, but should be close to 50:50 as is typical in minnow trap samples of stickleback populations in this area. Likewise, we did not know the age of the wild-caught fish, but all fish were likely to have been born in June or July of 2010 and hence be approximately one year old.

Cage size was constrained to 1 m diameter by the practical necessity of fitting enclosures within the width of the stream and among submerged logs in the lake. To induce high but realistic levels of resource competition within each cage, sample sizes per cage ($n = 3$ fish per square metre of substrate) were set to equal the upper bound of densities used in previous field experiments[18,21], which is roughly the upper bound of our estimates of natural abundances obtained from mark–recapture surveys. The number of cages (40 per habitat) therefore set the total sample size and power. The number of cages was determined by logistical constraints (time and materials to build and install cages). We therefore conducted a power analysis to evaluate our ability to detect biological effects given our design. We conservatively assumed an average 50% survival rate (at which our standard error is highest), which means our 60 fish per source–sink combination would confer a standard error of 6.4% (or less if survival were higher or lower). We therefore expected to be able to detect DS (origin × destination interaction with a 15% lower survival of immigrants), or a 15% survival penalty for majority ecotypes, with at least 65% power for each detection.

Cages were monitored periodically and any accumulated debris was brushed off the mesh of stream cages. On 5–7 July 2011, all surviving fish were removed from each cage, and then the cages were removed. Recaptured fish were euthanized with an overdose of MS-222 and preserved in 10% neutral buffered formalin. Standard length and body mass were recorded upon recapture. Fish growth was calculated as the change in mass from the start to end of the experiment (not log-transformed because some fish lost weight). We dissected all recaptured fish to determine their sex. Using a dissecting microscope, we scanned each fish for macroparasite infections as described in ref. 36. MHC class IIb loci were genotyped for each fish using DNA extracted from the spine clips taken before release, following the PCR amplification and 454 pyrosequencing protocol described in ref. 32. Genotype scoring was done using Stepwise Threshold Clustering[32]. Genotyping and phenotypic measurements were carried out by W.E.S., who was blind to individuals' origin or destination while processing a given specimen in the laboratory.

Statistical analysis of survival. We used survival rates as a measure of selection on lake and stream ecotypes in the lake and stream habitats. An essential caveat is that this experiment used only approximately one-year-old adult stickleback. Juveniles were too small to cage, and were not yet present in the area when we initiated the experiment. We therefore do not have measures of lifetime survival rates. Nor do we have measures of fecundity and mating success by adult fish because of the artificial social context of three fish in a 1-m diameter cage. Consequently, we estimate only a portion of selection arising from viability of adult fish. Readers should therefore be aware that selection gradients based on lifetime fitness may be different from our estimates. The relative strength of NFDS and DS measured here must be treated as estimates. Another important caveat is that we are measuring phenotypic selection only. Because we used wild-caught fish, their breeding values are unknown. Consequently, plasticity may have contributed to variation in size or morphology or physiology that influenced individual survival and performance (except *MHCIIb* genotype). We know that the focal populations differ at many loci in the genome, but the heritability of the specific traits examined here remains to be measured.

To test for local adaptation, we used a binomial GLM evaluating whether survival varied as a function of fixed effects of fish 'origin' (lake versus stream ecotypes), 'destination' (lake versus stream cages) and an origin × destination interaction. Cage was included as a random intercept effect, and the full model and its subsets were analysed in the R package *glmm*, using likelihood ratio tests to compare nested models. Local adaptation would be supported had we found an origin × destination interaction in which each ecotype's survival was highest in its native habitat but lower in the foreign habitat. Starting body mass or condition (residuals from a regression of mass on length) had no significant effect on survivorship over the duration of the study, so we omit this from consideration for brevity (the statistical significance of other variables was not altered by doing so). We also tested for local adaptation effects on individual growth rates, using analyses of covariance (ANOVAs) to test for origin and destination (and interaction) effects on change in body mass, change in length or change in condition. We observed no significant signal of DS using growth rates as a fitness measure, so for brevity we do not report details here. All analyses were conducted in the R environment[37]. All tests were two-sided.

As noted above, we did not determine the sex of released individuals, only recaptured ones. However, sex-biased survival should still be detectable in the form of a sex ratio bias among survivors, or sex ratio differences between treatment groups. To test for this bias, we used χ^2 tests to check whether the sex ratios of surviving fish (i) differed from a 50:50 expectation, (ii) differed from the sex ratio of wild-caught fish trapped just outside the enclosures at the end of the experiment, (iii) differed between habitats, or (iv) differed between native versus non-native caged fish. None of these tests yielded significant sex ratio differences so we infer that survival was not sex-dependent.

To test for NFDS, we designated each fish as either a majority or minority ecotype within its cage (2/3 or 1/3 of the fish per cage; $N = 160$ and $N = 80$ fish, respectively). Minority status was factorially crossed with whether fish were resident or immigrant in the cage habitat. We used a binomial GLM (*glmer* function in R) to test whether survival depends on minority versus majority status, origin, and destination, with cage as a random effect. As a post-hoc test (having observed stronger NFDS for stream natives), we also tested for interactions between minority status and origin, or between minority status and destination (Extended Data Table 1).

NFDS could arise from size-dependent competition among neighbours. To evaluate this possibility we tested whether survival depends on individuals' phenotypic deviation from their neighbours within their cage. We calculated phenotypic deviation as the absolute value of the difference between each fish's body mass and

the mean mass of all three fish per cage, rescaled by the standard deviation of body size. We then used a mixed model GLM (using *glmm*) to test for an effect of this size deviation on survival, while accounting for fixed effects of origin, destination, and random cage effects. This analysis does not distinguish between fish that were larger versus smaller than their cage mean mass, so we also tested for a signed effect of mass deviations (and found none). We observed no significant signal of NFDS when using growth rates as a fitness measure, so for brevity we do not report these details.

We used a Student's *t*-test to compare size deviations of minority versus majority fish, to determine whether size deviations and minority status were confounded. Minority fish deviated more from their cage mean size, so we used multiple regression with a mixed-model GLM to partition the effects of minority status or size deviation on survival. We also used path analysis for the same purpose, testing whether minority status affected survival directly, or indirectly via size deviation. Path analysis and regression yielded identical inferences, so we present only the mixed model results in the main text.

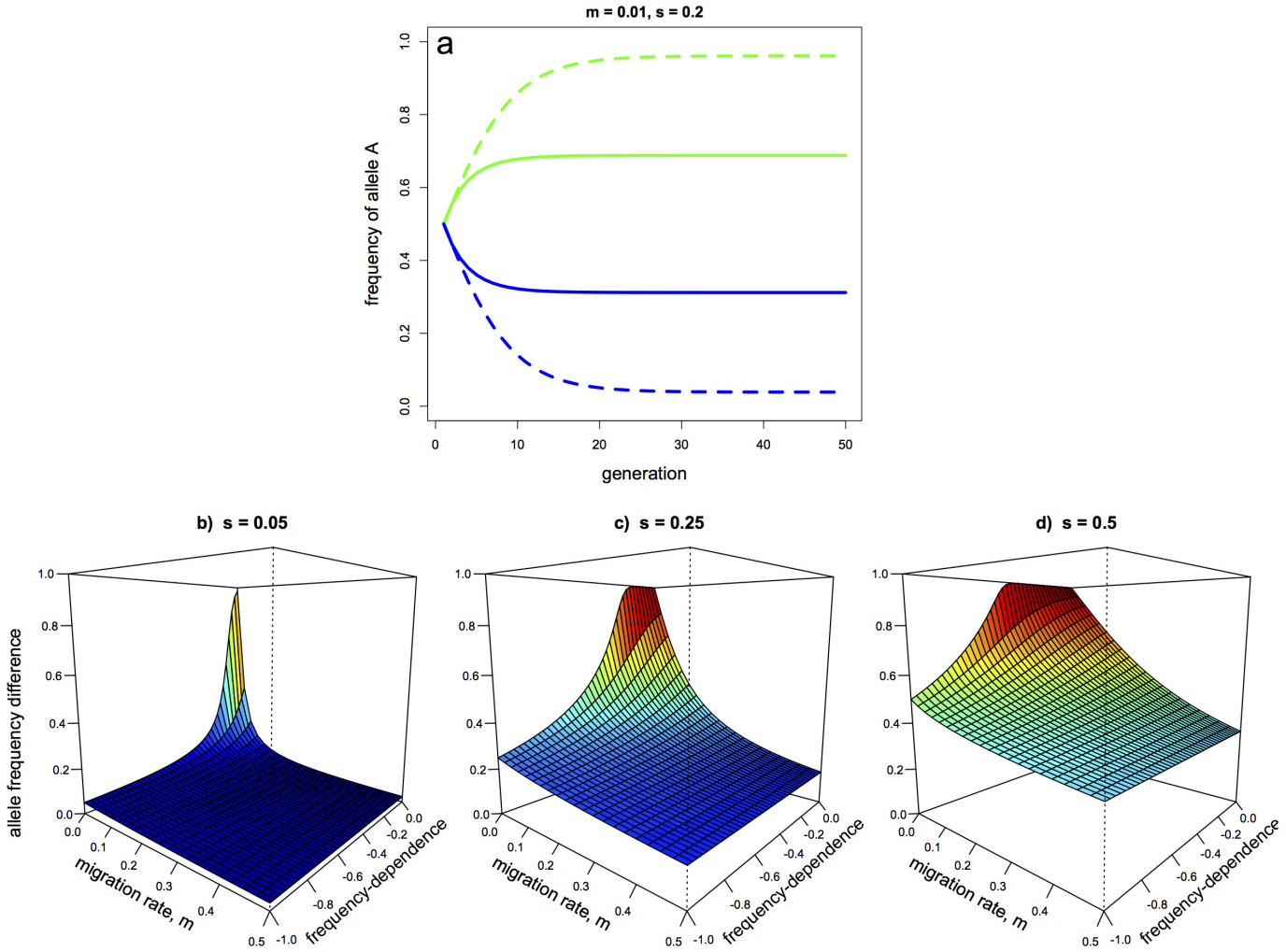
Statistical analysis of parasite load. We calculated a standardized index of total parasite load for each fish, combining data for multiple parasite taxa. We first calculated the maximum within-fish abundance for each parasite taxon. Counts of each parasite taxon, in each fish, were then re-scaled relative to the taxon-specific maximum. The result is that each parasite taxon's standardized load varied from 0 (no infection) to 1 (maximum observed infection load). We then summed these relative loads across all parasite taxa for each individual fish. This total parasite load was then rescaled to fit a standard normal distribution (subtracting the mean and dividing by the standard deviation). Positive numbers represent individuals that are relatively heavily infected by multiple parasite taxa; negative numbers represent individuals with relatively light infection loads by few parasite taxa. Lastly, because infection load increases with body mass in both lake and stream stickleback from these sites, we obtained residuals of these standardized total infection loads in a regression against ln(mass) of the fish.

Using these standardized infection load scores, we first tested whether infection load depended on host source or destination or a source × destination interaction (with a random effect of cage). We then tested whether infection load depends on whether a host was the native ecotype, and/or had a typical native *MHCIIb* genotype. To do so we quantified the extent to which the *MHCIIb* genotype of a given fish was typical of its cage habitat. Our *MHCIIb* sequence data (described above) provided a matrix of allele presence/absence for every fish placed into the cages. We used discriminant function analysis of this matrix of *MHCIIb* allele presence/absence, to obtain a quantitative axis measuring how lake-like versus stream-like each individual's *MHCIIb* genotype was. Initially, individuals with positive scores along this axis had more typical lake *MHCIIb* alleles, whereas negative axis values indicate stream-like *MHCIIb.* To convert this axis into a measure of how 'native' the *MHCIIb* of each fish was in its cage habitat, for all stream-caged fish we reversed the sign of this discriminant axis. The result is a numerical score that, in both habitats, is positive for fish with locally typical *MHCIIb* genotypes (lake-like *MHCIIb* in the lake, stream-like in the stream) and negative for fish with typically foreign genotypes (stream-like *MHCIIb* in the lake, lake-like *MHCIIb* in the stream). Note that we are testing whether individuals' *MHCIIb* genotype is rare for its destination habitat as a whole, not whether it is rare relative to the two other fish in its particular cage. We then used a linear mixed model to test whether standardized total parasite load depended on whether fish were native/non-native, or had a native versus foreign *MHCIIb* genotype score. We included destination habitat and all two-way interactions, plus a random effect of cage. Model statistical significance was determined by sequential deletion and likelihood ratio tests, while also using Akaike Information Criterion (AIC) model selection procedures to choose among alternative statistical models (these approaches yielded the same final inference). Note that we obtained the same inferences if we used normalized parasite load without size correction, and instead used fish mass as a covariate in our statistical model. We present the size-standardized parasite load measure because it is easier to visualize without the added dimension of fish size. We assume that high standardized parasite load is bad for fish and can thus be used as an indirect proxy for selection. However, we recognize that not all parasites affect host fitness equally (or at all). Likewise, as with viability selection we wish to emphasize that parasite load of adult fish is only one component of lifetime fitness; for example, infections can affect juvenile survival and adult reproductive success as well.

Data availability. All data necessary for the results presented in this paper are archived at Datadryad.org doi:10.5061/dryad.61kr4. Source Data to recreate specific figures are provided with this publication.

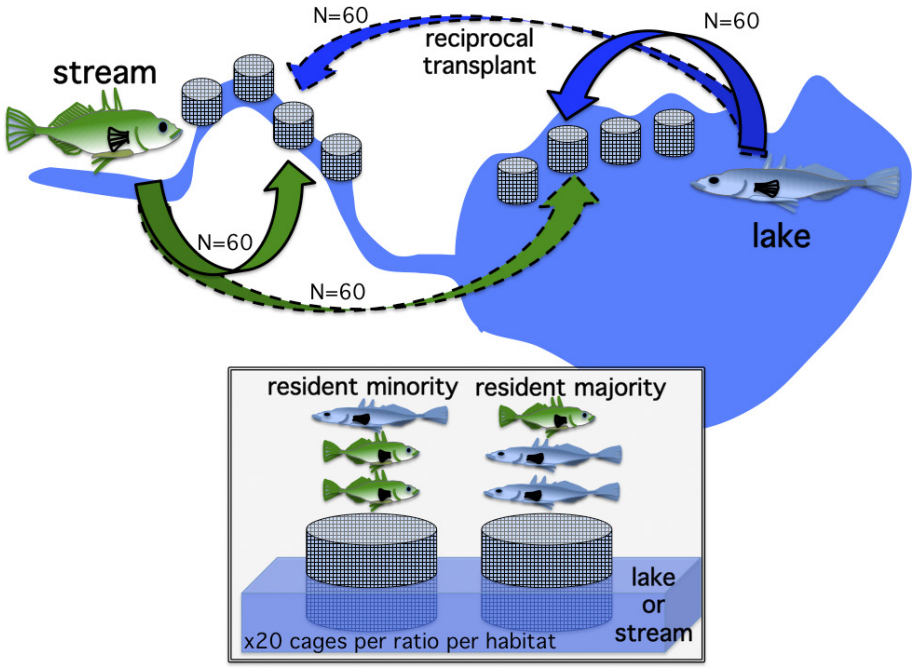
Code availability. The code required to recreate simulations is provided in the Supplementary Information.

31. Weber, J., Bradburd, G. S., Stuart, Y. E., Stutz, W. E. & Bolnick, D. I. The relative contributions of distance, landscape resistance, and habitat, to genomic divergence between parapatric lake and stream stickleback. *Evolution* **71,** 342–356 (2017).
32. Stutz, W. E. & Bolnick, D. I. A Stepwise Threshold Clustering (STC) method to infer genotypes from error-prone next-generation sequencing of multi-allele genes such as the Major Histocompatibility Complex (MHC). *PLoS One* **9,** e100587 (2014).
33. Oke, K. *et al.* Plasticity enhances phenotypic parallelism: evidence from lake-stream stickleback. *J. Evol. Biol.* **29,** 126–143 (2016).
34. Berner, D., Grandchamp, A.-C. & Hendry, A. P. Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution* **63,** 1740–1753 (2009).
35. Bolnick, D. I. *et al.* Ecological release from interspecific competition leads to decoupled changes in population and individual niche width. *Proc. R. Soc. B* **277,** 1789–1797 (2010).
36. Stutz, W. E., Coates, J., Schmerer, M. & Bolnick, D. Among-population divergence in stickleback immune gene expression is predominantly environmentally-induced, rather than heritable. *Mol. Ecol.* **24,** 4629–4646 (2015).
37. R Development Core Team *R: a Language and Environment for Statistical Computing.* http://www.R-project.org (R Foundation for Statistical Computing, 2008).
38. Bolnick, D. & Otto, S. The magnitude of local adaptation under genotype-dependent dispersal. *Ecol. Evol.* **3,** 4733–4735 (2013).

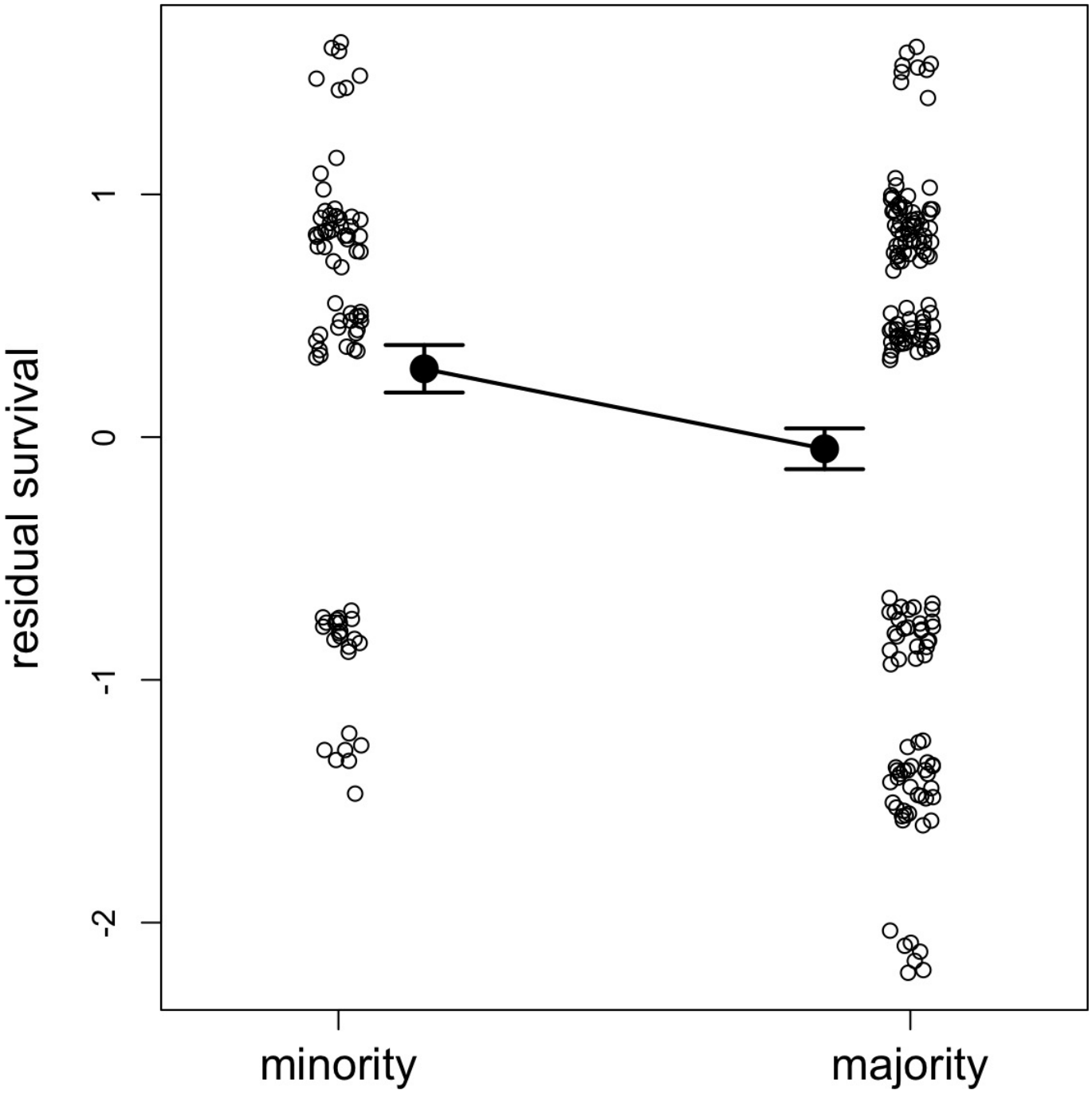


Extended Data Figure 1 | See next page for caption.

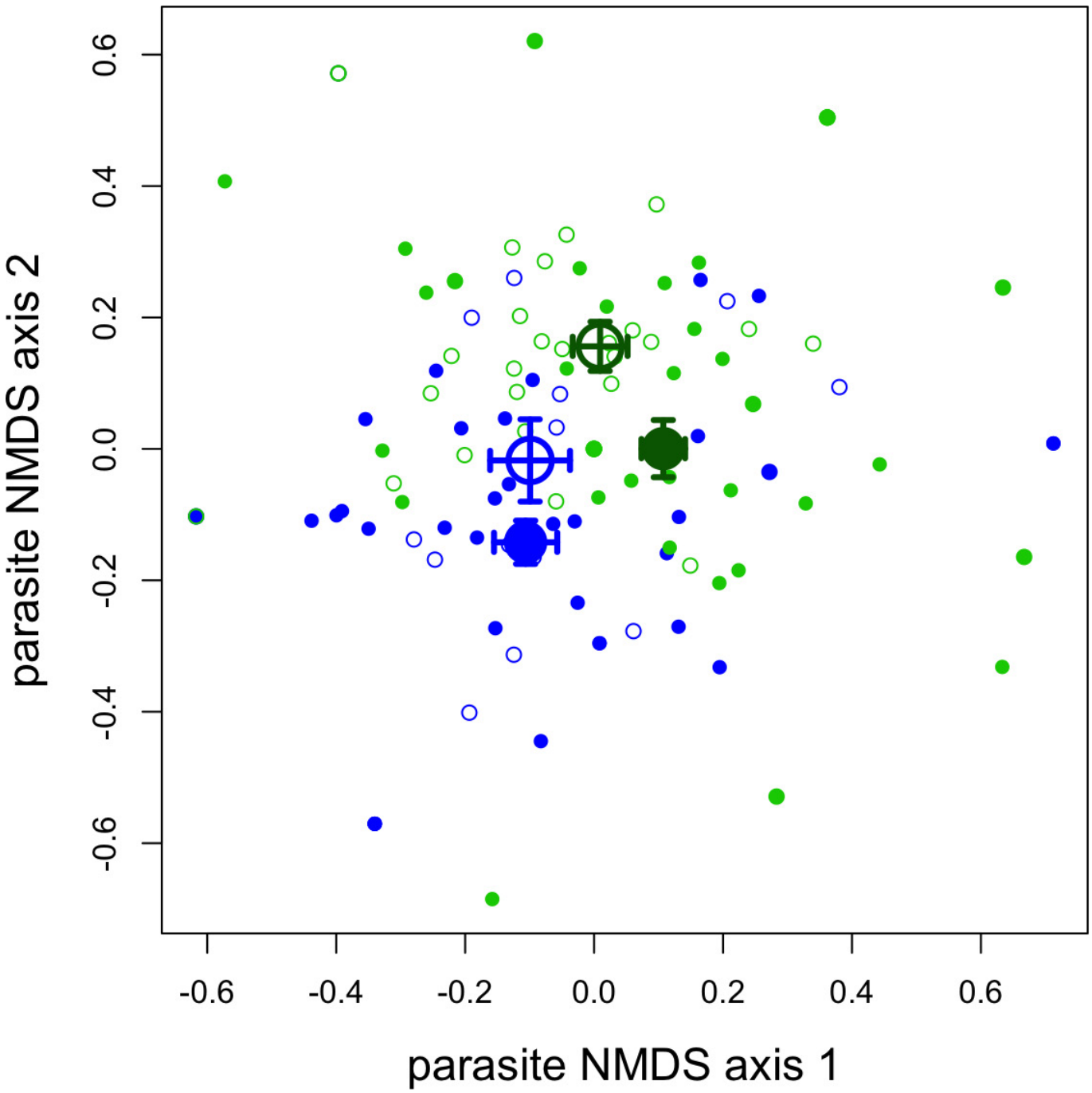
Extended Data Figure 1 | Simulations of how migration, NFDS and DS jointly affect the equilibrium evolutionary divergence between populations. a, To illustrate the effect of NFDS on between-population divergence, here we plot a single instance of a simulation in which we began with two non-diverged populations (allele frequencies $p_A = p_B = 0.5$). Simulation details are provided below. In **a** we plot the process of population divergence to an equilibrium, with or without NFDS (for example, $\gamma = 0$ or -0.5; dashed and solid lines respectively), using blue and green lines to distinguish allele frequencies in habitats A and B, respectively. When NFDS is acting (solid lines), population divergence is much less pronounced than without NFDS, for a given strength of selection and migration. **b**–**d** show the equilibrium allele frequency difference (Δp, vertical axis) as a function of the migration rate, m, and the strength of NFDS, γ, for three strengths of DS, $s = 0.001$, 0.1 and 0.5 (**b**, **c** and **d**, respectively). This figure shows us the well known tendency for migration to constrain divergence (Δp declines as function of m). However, our addition of NFDS reveals a comparable constraint on divergence as a result of NFDS (Δp is smaller for more negative values of γ, indicating that NFDS constrains divergence). There is also an interaction between m and γ, which reflects the fact that the constraining effect of NFDS is most pronounced when migration rates are low, because the migrants are present but rare enough to benefit strongly from NFDS. At higher immigration rates, immigrants become increasingly common and their frequency-dependent advantage is reduced. We used multiple regression to measure the relative effects of m, s, and γ, and the $m \times \gamma$ interaction on Δp. Within the parameter space that we examined, selection has the strongest effect on final allele frequency divergence (55.9% of variance explained), migration is next strongest (20.8% of variance), and NFDS explains 5.8% of variance with an additional 5.7% of variance attributed to a $\gamma \times m$ interaction. Linear regression coefficients indicate that Δp declines by 0.050 per 0.1 unit of migration (a linear approximation of a nonlinear trend), whereas Δp declines by 0.022 for each 0.1 of γ below 0. We infer that the effect of γ (NFDS) is roughly 1/3 as strong as the effect of migration, but that these variables strongly interact. Thus, it is clear from our simulations that NFDS constrains population divergence by amplifying the effective migration rate. Simulation methods: we used a simple population genetic model to evaluate the relative effects of, and synergistic interaction between, migration, frequency-dependent selection, and DS. We used discrete-time deterministic numerical simulations to model the migration–selection balance between two discrete habitats A and B. Each habitat contained a haploid population with a polymorphic locus (alleles a and b) with allele frequencies $p(a) + p(b) = 1$. Each population was kept at a constant and effectively infinite population size so that genetic drift had no effect, and there were no demographic source–sink dynamics. In habitat A, the fitnesses of the two alleles were $w_{a,A} = 1.0$ and $w_{b,A} = 1 - s$, where s is the strength of selection against immigrants. Fitnesses were reversed in habitat B ($w_{a,B} = 1 - s$ and $w_{b,B} = 1$). This symmetric DS was frequency-independent. Divergence in response to this selection was undermined by migration. Every generation, a fraction m of individuals in each habitat migrated to the other habitat as juveniles (before selection acts). Within each habitat, a fraction 1 − m of the residents did not disperse. To incorporate frequency-dependent selection, we adjusted each genotype's fitness (after migration) to account for its relative abundance. Specifically, the frequency-dependent fitness of each allele i in habitat j was: $w'_{i,j} = w_{i,j} + \gamma(p(i) - 0.5)$ where γ dictates the strength and direction of frequency dependence ($\gamma = 0$ imparts no frequency dependence, $\gamma < 0$ imparts increasingly strong negative frequency dependence, and positive frequency dependence occurs when $\gamma > 0$; this is subject to the usual constraint that $0 \leq p(i) \leq 1.0$). We focused exclusively on $\gamma < 0$ for our simulations, as we were interested specifically in negative frequency dependence. Note that this is merely a heuristic model to generate negative frequency-dependence to illustrate our point, and is not tailored to reflect a specific biological process. We then allowed selection to act on this population using the frequency-dependent fitness within each habitat. The frequency of allele a in the next generation in habitat j was: $p_j(a)_{t+1} = \frac{p_j(a)_t w'_{a,i}}{\left(p_j(a)_t w'_{a,i} + p_j(b)_t w'_{b,i}\right)}$ To examine the dynamics of this model, we initiated both populations with allele frequencies of 0.5, and iterated through multiple generations, with each generation containing a bout of migration, then frequency-dependent adjustment to fitness, then selection. We ran each simulation until the allele frequency reached an equilibrium in each habitat. For each simulation run we recorded the ending allele frequency difference between the habitats, $\Delta P = p_A(a) - p_B(a)$. Larger values of Δp denote more substantial genetic divergence between the populations. We repeated this simulation for a fully factorial combination of values of s (0.001, 0.005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5), m (0.001 to 0.5 in increments of 0.001), and γ (0 to 1 in increments of 0.001).



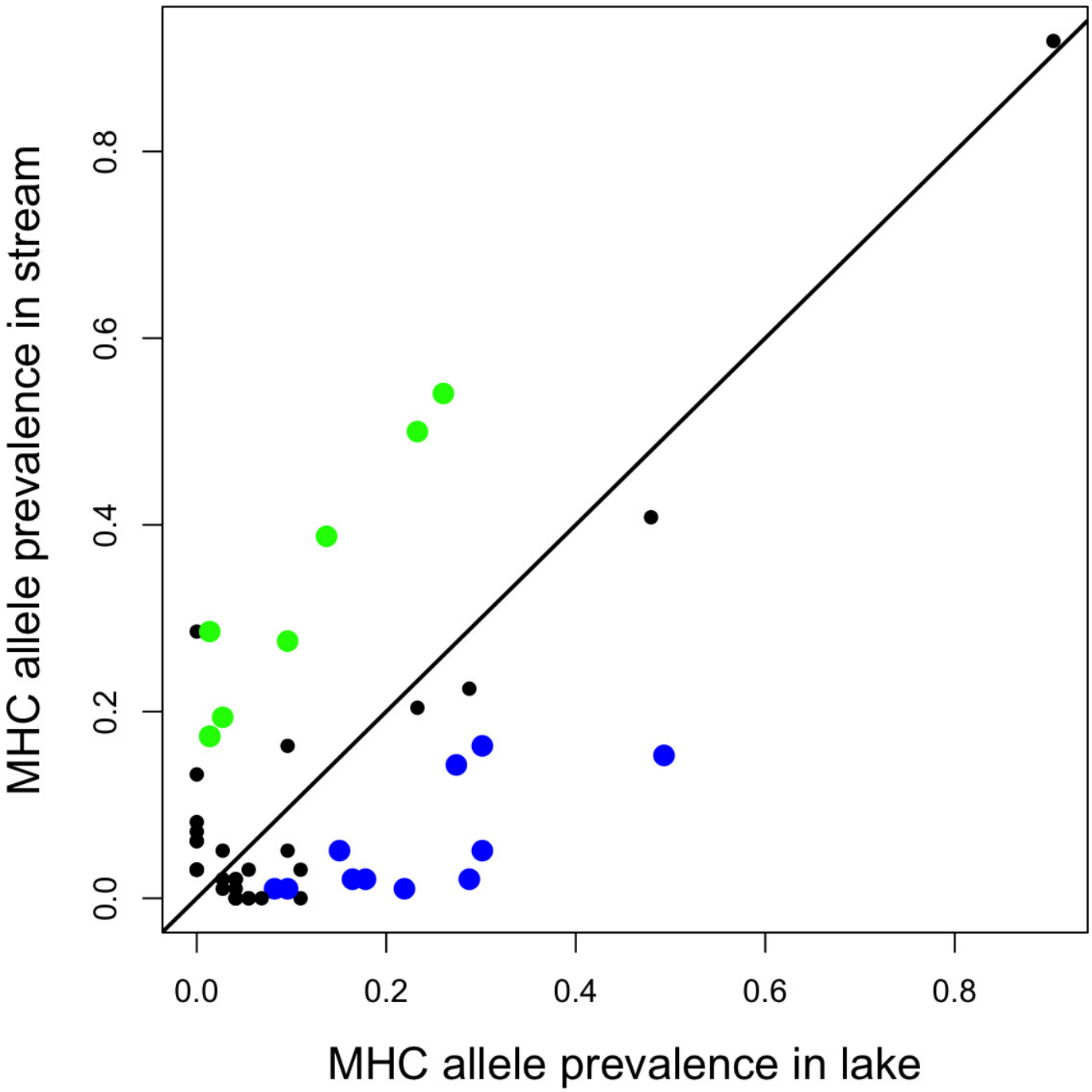
Extended Data Figure 2 | A schematic diagram of our experimental design. We constructed 40 cages in Roberts Lake, Vancouver Island, British Columbia, and 40 cages in the adjoining stream that drains out of the lake. We captured 120 wild-caught lake stickleback (indicated by blue shading), and 120 wild-caught stream stickleback (green shading). These were split evenly between lake and stream cages (60 per source and destination combination). Arrows with dashed perimeters indicate immigrants, solid perimeter arrows indicate residents. Within each habitat, we factorially manipulated the relative abundance of resident and immigrant fish (1:2 or 2:1 ratio; 20 cages per habitat per ratio).



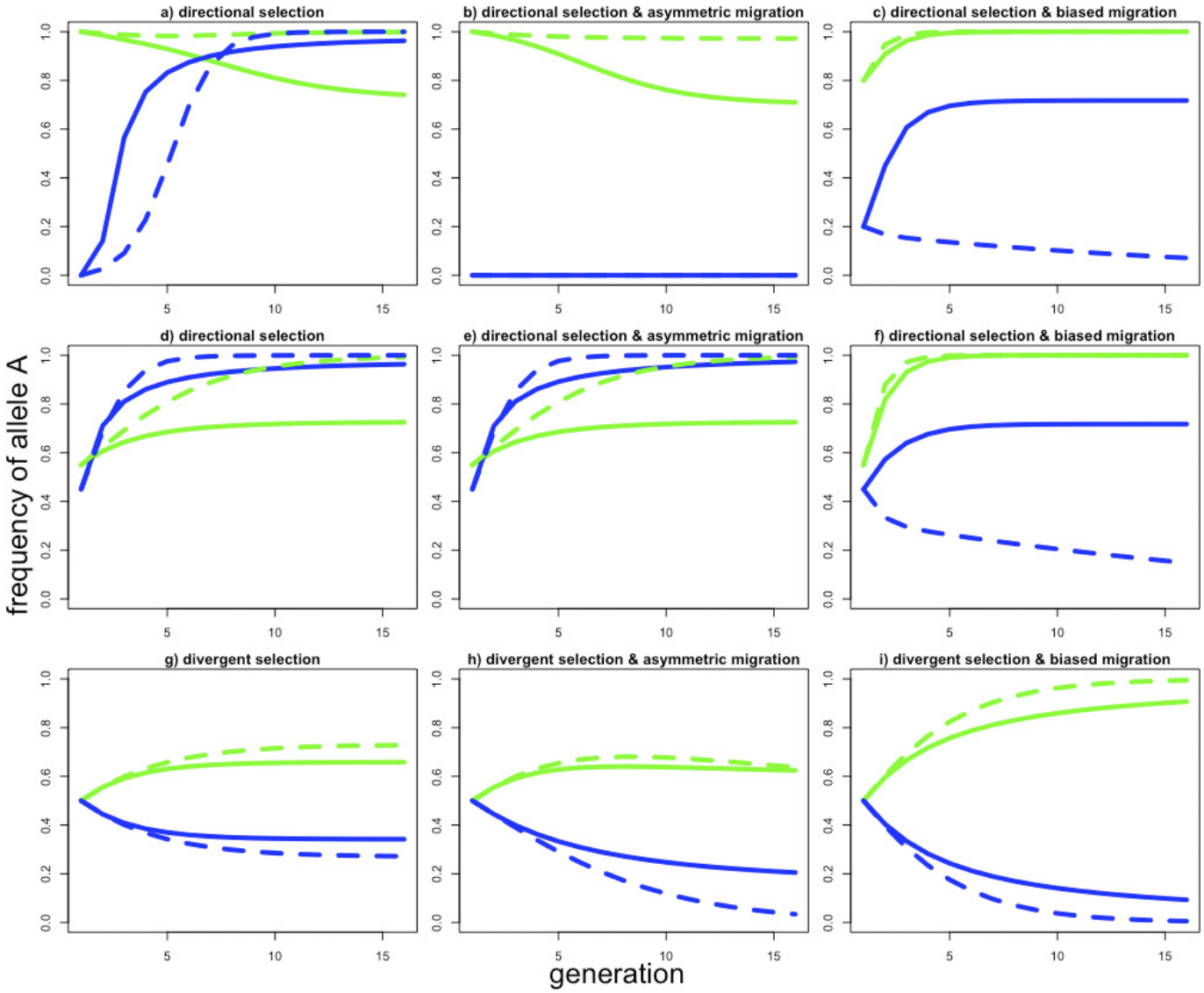
Extended Data Figure 3 | The main effect of minority versus majority status on stickleback survival (replotted from Fig. 1), after accounting for origin and destination. We calculated the residuals from a binomial general linear mixed model regressing survival on fish origin and destination fixed effects with a random effect of cage. Here, we plot these residuals as a function of fish minority/majority status, showing residuals for all fish (open circles) and minority/majority means (filled circles) with ±1 standard error bars. A slight vertical and horizontal jitter has been added to distinguish otherwise overlapping points. Residual survival is significantly higher for minority ecotype fish (Wilcox rank sum test $W = 7656$, $P = 0.0132$).



Extended Data Figure 4 | Both fish origin and transplant destination affect the identity and abundance of parasites infecting surviving fish. Fish destination explains relatively more variance in parasite community composition. We used counts of each parasite in each fish to calculate Euclidean NMDS scores (using the *metaMDS* package in R). Here we plot individuals' first and second NMDS axis scores. Individual points are colour-coded by origin (green for stream; blue for lake) and symbols denote destination (circle for lake; triangle for stream) as described in Fig. 1. Larger symbols indicate group means with one standard error bars. We subjected the two leading NMDS axes to a multivariate analysis of covariance (MANOVA) to test for effects of fish origin (Pillai's trace = 0.086, $P = 0.0015$), destination (Pillai's trace = 0.163, $P = 0.0000029$), and their interaction (Pillai's trace = 0.008, $P = 0.555$). Fish with positive values of NMDS axis 1 carried more *Neoechinorhynchus* parasites but had fewer blackspot infections. Fish with positive values of NMDS axis 2 carried more *Diplostomum* but fewer *Eustrongylides* and *Thersitina*. NMDS axis 2 is primarily responsible for both the origin and destination effects of our transplant experiment: stream natives in stream cages had the highest NMDS axis 2 score, whereas lake fish in lake cages had the lowest NMDS axis 2 score. Non-native fish (transferred from lake to stream or vice versa) were intermediate between these extremes and not significantly different from each other. Thus, fish transferred between habitats converged partially on the parasite community of the native fish of their new habitat. Sample sizes are the number of survivors listed in Table 1.

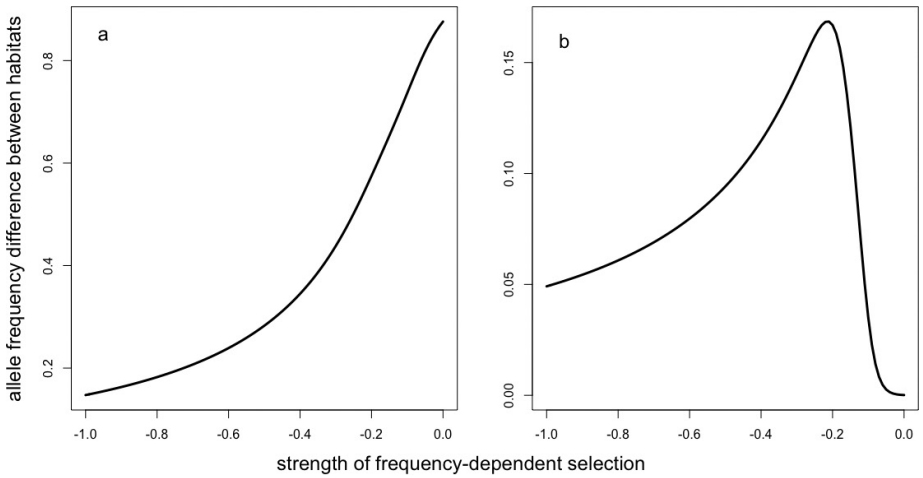


Extended Data Figure 5 | Lake versus stream divergence in MHC IIb genotypes. Each unique exon 2 amino acid sequence ('allele') is plotted as a point, showing its prevalence in the lake-origin fish used in our transplant experiment (*x* axis) and stream-origin fish in the experiment (*y* axis). Because alleles can be distributed across MHC paralog copies, we calculate prevalence (the fraction of individuals carrying a given allele) rather than allele frequencies. Therefore the prevalences of all alleles sum to more than one. Alleles showing a significant lake versus stream difference in prevalence (binomial GLM, $P < 0.05$) are enlarged and coloured to indicate the habitat in which the allele is more common (blue for lake; green for stream). The diagonal line shows where allele frequencies are identical between the populations. Sample sizes of fish are 120 lake and 120 stream natives (Table 1).



Extended Data Figure 6 | See next page for caption.

Extended Data Figure 6 | Simulations to evaluate the outcome of population divergence given both NFDS and directional selection (for example, higher fitness of stream fish everywhere). As in Extended Data Fig. 1, here we plot the temporal dynamics of allele frequency in habitat A (green lines) and B (blue lines), with and without negative frequency dependence (solid and dashed lines respectively). Unlike in Extended Data Fig. 1, here we consider the evolutionary effects of symmetric migration (**a, d** and **g**), asymmetric migration (**b**, **e** and **h**) and biased migration (**c**, **f** and **i**) when there is directional selection (**a**–**f**) or DS (**g**–**i**). **a**–**c** illustrate the potential for initially divergent populations to remain at least partly differentiated despite directional selection. **d**–**f** illustrate the potential for initially similar populations to become genetically divergent despite directional selection. **g**–**i** illustrate the evolutionary dynamics for the same phenomena but when there is DS (as in Extended Data Fig. 1). These simulations are not meant to be an exhaustive analysis, but rather provide examples of possible evolutionary outcomes. In particular, the simulations are intended to demonstrate that lake–stream genetic divergence is plausible despite the directional selection favouring stream fish in both habitats. We repeated the simulations outlined in Extended Data Fig. 1 (R code is provided in the Supplementary Information), with the following modifications. (i) We compared the effects of DS (as in Extended Data Fig. 1) versus directional selection. For the directional selection we used fitness values directly drawn from Table 1, standardized to a maximum fitness of 1 within each habitat. That is, assuming the stream represents habitat A, the fitness of allele i in habitat j, w_{ij}, is $w_{a,A} = 1.0$, $w_{a,B} = 1.0$, $w_{b,A} = 0.74$ and $w_{b,B} = 0.37$. (ii) We introduced asymmetric migration (for instance, caused by water flow). We assumed that the emigration rate of the fitter ecotype $m_{A\to B}$ is some fraction f of the reverse migration rate $m_{B\to A}$, so that the reciprocal migration rates can be expressed as $(f \times m, m)$. When $f = 0$, the universally fittest ecotype (for example, stream fish) cannot invade the habitat occupied by the less fit ecotype (for example, the lake), unless it is already present at non-zero frequency. (iii) We considered the case of biased movement. We assumed that allele a prefers to remain in habitat A, and allele b prefers to remain in habitat B. We therefore set their migration rates to zero. In contrast, the mismatched allele (b in habitat A; a in habitat B) emigrates at a rate m which we assume to be symmetric. For instance, in habitat A migration changes the frequency of allele a from p_A to: $p'_A = \frac{p_A + p_B m_{B\to A}}{p_A + (1+p_A)(1-m_{A\to B}) + p_B m_{B\to A}}$ The parameter value sets used to illustrate these possible outcomes are illustrated in Extended Data Fig. 6. In panel **a**, $\gamma = -0.6$, $m = 0.01$, $p_A = 1.0$, $p_B = 0.0$, and selection coefficients are from Table 1. Panel **b** is the same as **a** except that migration is zero in the upstream (habitat A to habitat B) direction. In panel **c**, $p_A = 0.8$ and $p_B = 0.2$, the better to visualize complete fixation in the frequency-independent case, and $m = 0.65$, indicating that mismatched genotypes actively rather than randomly ($m = 0.5$) switch habitats. Panels **d**–**f** are the same as panels **a**–**c** but with different starting allele frequencies ($p_A = 0.55$ and $p_B = 0.45$) to illustrate that the equilibrium results in panels **a** and **c** are independent of initial conditions. However, evolution with asymmetric migration is highly sensitive to initial conditions because allele a will sweep to fixation within habitat B as long as it is present initially, whether or not there is migration. Panels **g**–**i** are the same as panels **a**–**c** but assuming symmetric DS (as in Extended Data Fig. 1) to examine the effects of asymmetric migration (**h**) and biased migration (**i**). For all three scenarios with DS, NFDS acts to undermine genetic divergence between the habitats (as in Extended Data Fig. 1). **h**, However, asymmetric migration means that the sink population is more polymorphic than the source population. **i**, Genotype-dependent dispersal enhances overall divergence between populations, as previously described[38]. As one would expect, the directional selection documented in Table 1 leads to rapid fixation of the 'stream' allele a in both habitats (**a**, dashed lines). However, the addition of NFDS (solid lines) leads to a counter-intuitive result: population divergence. Populations that would collapse to a single genotype (all $p(a) = 1.0$, dashed lines), instead maintain modest genetic differences when NFDS is added (the solid blue and green lines diverge; **a**). This result is easily explained: NFDS tends to maintain polymorphism within populations. For the particular parameter values chosen to illustrate this point ($\gamma = -0.6$, $m = 0.01$), NFDS maintains both alleles a and b in both habitats. The equilibrium for this stable polymorphism is higher for the blue line (habitat B) than the green line (habitat A) because NFDS must overcome stronger selection against allele b in habitat B. The equilibrium is insensitive to initial conditions (such as starting allele frequencies; **d**). This simulation thus points out that NFDS promotes polymorphism within populations (as is well known), but when these equilibrium allele frequencies differ between populations NFDS can sustain between-population genetic differences when they would not otherwise occur. This is an important point for our empirical system, because it suggests that strong NFDS could explain the persistence of allele frequency differences between lake and stream fish despite directional selection favouring one ecotype over the other. NFDS might thus resolve the conundrum posed in Supplementary Information section 1, subsection 3: why does lake–stream divergence persist despite widespread directional selection rather than the expected DS? When we add both directional selection and asymmetric migration (**b**), several results can occur. If we start with completely fixed differences between populations, then population differences persist despite directional selection. Even though allele a is favoured in both habitats in this simulation, it is initially absent in habitat B (blue, 'lake') and never arrives because migration is strictly directional from habitat B to A. Allele a fixes in habitat A (where it is favoured), and never makes it to habitat B (where a is also favoured). However, if we relax the initial conditions even slightly (non-fixed differences at the start, or weak upstream migration), then population divergence rapidly collapses as allele a fixes in both habitats either by a selective sweep within habitat A and B separately, or by immigration into habitat B followed by a sweep. Adding NFDS to the strict initial conditions (solid lines in **b**), allele a remains absent in habitat B because it is initially absent and there is no immigration. However, emigration from habitat B into A introduces allele b into the latter population, where it is maintained by NFDS despite frequency-independent directional selection (green line, whose equilibrium outcome matches the result in **a**). In the more relaxed initial conditions, the combination of directional selection, NFDS, and moderately asymmetric migration results in an outcome that looks very much like that in **a**. Next, we consider the case of biased dispersal in which allele a exhibits philopatry for habitat A, and allele b prefers habitat B. As described elsewhere[38], this genotype-dependent dispersal facilitates population divergence (for example, divergence is greater in panels **c**, **f** and **i** than in panels **a**, **d** and **g**. Notably, this is possible even when allele a is favoured everywhere by selection (**c**). This is plausible if, for example, stream fish evolved to prefer the stream where they are fittest (Table 1), and lake fish prefer the lake as a refuge from stream fish. However, incorporating NFDS can prevent fixation of the philopatric types in one habitat (as shown in **c**) or in both habitats (not shown), depending on parameter values. In conclusion, these simulations confirm that there are multiple mechanisms that can explain the persistent divergence between lake and stream stickleback, even if the populations experience persistent directional rather than DS. The top row of figures here provides examples of allele-frequency differences arising from NFDS (**a**), asymmetric migration (**b**), and biased migration of genotypes (**c**). Of these, the most intriguing is the role of NFDS in generating stable allele frequency differences between populations (albeit not fixed differences).



Extended Data Figure 7 | NFDS can suppress or enhance population divergence. a, When the populations in habitats A and B are subject to DS (as in Extended Data Fig. 1) and NFDS is absent ($\gamma = 0$), equilibrium allele frequency divergence is substantial and reflects the migration–selection balance. Stronger NFDS (more negative γ) reduces this equilibrium between-habitat divergence. **b**, When there is directional selection and no NFDS ($\gamma = 0$) the populations will fail to diverge because a single allele fixes in both habitats (see Extended Data Figs 1 and 6). Introducing NFDS ($\gamma < 0$) facilitates allele frequency difference between populations as long as selection is not identical in both habitats (for example, here $w_{a,A} = 1.0$, $w_{a,B} = 1.0$, $w_{b,A} = 0.8$ and $w_{b,B} = 0.9$). This is because NFDS favours different polymorphic equilibria in the two habitats, generating allele frequency differences (see Fig. 1a and Extended Data Fig. 6). However as NFDS strengthens ($\gamma << 0$) the effect of unequal directional selection becomes comparatively weak and the populations' equilibrium allele frequencies converge again. The result is that there is an intermediate level of NFDS that can cause population genetic differences despite directional selection (**b**), when NFDS might otherwise undermine population genetic differences arising from DS (**a**).

Extended Data Table 1 | Mixed effect binomial GLMs testing proposed effects on stickleback survival

Model	Effect	Estimate	Std Error	Z	P
A AIC: 271.60	**Origin**	**2.08**	**0.48**	**4.32**	**<0.0001**
	Destination	**1.74**	**0.46**	**3.77**	**0.0002**
	Origin:Destination	-0.62	0.67	-0.93	0.3523
	Enclosure variance	0.256			
B AIC: 266.49	**Origin**	**2.41**	**0.54**	**4.44**	**0.00001**
	Destination	**2.94**	**0.5**	**3.85**	**0.00012**
	Origin:Destination	-0.79	0.69	-1.14	0.25602
	Majority	**-0.967**	**0.39**	**-2.48**	**0.01306**
	Enclosure variance	0.475			
C AIC: 265.80	**Origin**	**2.06**	**0.42**	**4.88**	**<0.0001**
	Destination	**1.6**	**0.39**	**4.1**	**<0.0001**
	Majority	**-0.92**	**0.38**	**-2.41**	**0.0159**
	Enclosure variance	0.4056			
D AIC: 263.39	**Origin**	**3.05**	**1.16**	**3.49**	**0.0005**
	Destination	**2.05**	**0.73**	**2.82**	**0.0048**
	Majority	-0.01	0.59	-0.02	0.9867
	Majority:Origin	**-2.55**	**1.22**	**-2.09**	**0.0367**
	Majority:Destination	-0.65	0.78	-0.83	0.4061
	Enclosure variance	0.3463			
E AIC: 262.77	**Origin**	**2.53**	**0.56**	**4.47**	**<0.0001**
	Destination	**1.9**	**0.51**	**3.7**	**0.0002**
	Origin:Destination	-0.7	0.71	-0.99	0.3232
	Majority	-0.67	0.41	-1.63	0.1033
	Size deviation	**1.27**	**0.54**	**2.34**	**0.0195**
	Enclosure variance	0.532			

We consider a set of five models beginning with a simple test of reciprocal local adaptation (A), then adding an effect of majority/minority status (B), retaining majority status but removing the origin–destination interaction (C), a post hoc model in which we added the majority–origin interaction and the majority–destination interaction (D), then finally considering the effect of adding the absolute body size deviation from the cage mean (E). All models include a random effect of cage. Models are ordered in increasing support from AIC, though models B–E are all within $\Delta AIC < 4.0$. We provide effect size estimates for each term in each model, the standard error of these effect size estimates, a *Z*-statistic measuring how many standard errors the estimate is from a null expectation of zero effect, and *P*-values measuring the probability of observing this large a *Z* value or greater if the null hypothesis were true. the null hypothesis that the effect in question is zero. Bold font indicates effects whose *P*-value is below the traditional threshold of 0.05.

Extended Data Table 2 | Mixed models testing effects of native ecotype, and locally common *MHCIIb* genotype, on standardized parasite loads

Models	Model contents:						AIC	Compare	χ^2	df	P
1	Infection load =	Native ecotype +	Native MHC +	Destination +	Native ecotype:Destination +	Native MHC:Destination	306.07				
2	Infection load =	Native ecotype +	Native MHC +	Destination +	Native ecotype:Destination +		311.09	2 vs 1	7.03	1	**0.0080**
3 (best model)	Infection load =	Native ecotype +	Native MHC +	Destination +		Native MHC:Destination	305.98	3 vs 1	1.92	1	0.1663
4	Infection load =	Native ecotype +	Native MHC +	Destination			315.19	4 vs 2	11.21	1	**0.0008**
								4 vs 3	6.1	1	**0.0135**
5	Infection load =	Native ecotype +	Native MHC				313.32	5 vs 4	0.12	1	0.7226
6	Infection load =	Native ecotype +		Destination			318.41	6 vs 4	5.21	1	**0.0224**
7	Infection load =		Native MHC +	Destination			320.91	7 vs 4	7.72	1	**0.0055**
8	Infection load =	Native ecotype +					316.41	8 vs 5	5.09	1	**0.024**
9	Infection load =		Native MHC +				319.19	9 vs 5	7.86	1	**0.005**
10	Infection load =			Destination			319.21	10 vs 6	2.8	1	0.0941
11	Infection load =						317.69	11 vs 8	3.28	1	0.0703

We used linear models to test whether standardized total infection load was a function of native versus foreign ecotype (fixed effect), and native versus foreign *MHCIIb* genotype (a continuous variable; positive values denote individuals with alleles that are diagnostic for the cage habitat), with transplant habitat (destination, a fixed effect). Cage was included as a random effect in all the models. We omit the origin effect because native/foreign status is confounded with the origin–destination combination. We present models from most to least complex, using sequential χ^2 tests to compare successively simpler models to determine the statistical significance of individual effects. Model 3 is optimal by AIC criteria. df, degrees of freedom.