doi:10.1111/evo.13110

Partitioning the effects of isolation by distance, environment, and physical barriers on genomic divergence between parapatric threespine stickleback

Jesse N. Weber,[1,2] Gideon S. Bradburd,[3] Yoel E. Stuart,[1] William E. Stutz,[1,4] and Daniel I. Bolnick[1,5]

[1]*Department of Integrative Biology, University of Texas at Austin, Austin, Texas 78712*

[2]*Division of Biological Sciences, University of Montana, Missoula, Montana 59801*

[3]*Department of Integrative Biology, Michigan State University, East Lansing, Michigan 48824*

[4]*Department of Ecology and Evolutionary Biology, University of Colorado at Boulder, Boulder, Colorado 80309*

[5]*E-mail: danbolnick@utexas.edu*

Received May 31, 2016
Accepted October 15, 2016

Genetic divergence between populations is shaped by a combination of drift, migration, and selection, yielding patterns of isolation-by-distance (IBD) and isolation-by-environment (IBE). Unfortunately, IBD and IBE may be confounded when comparing divergence across habitat boundaries. For instance, parapatric lake and stream threespine stickleback (*Gasterosteus aculeatus*) may have diverged due to selection against migrants (IBE), or mere spatial separation (IBD). To quantitatively partition the strength of IBE and IBD, we used recently developed population genetic software (BEDASSLE) to analyze partial genomic data from three lake-stream clines on Vancouver Island. We find support for IBD within each of three outlet streams (unlike prior studies of lake-stream stickleback). In addition, we find evidence for IBE (controlling for geographic distance): the genetic effect of habitat is equivalent to geographic separation of ~1.9 km of IBD. Remarkably, of our three lake-stream pairs, IBE is strongest where migration between habitats is easiest. Such microgeographic genetic divergence would require exceptionally strong divergent selection, which multiple experiments have failed to detect. Instead, we propose that nonrandom dispersal (e.g., habitat choice) contributes to IBE. Supporting this conclusion, we show that the few migrants between habitats are a nonrandom subset of the phenotype distribution of the source population.

KEY WORDS: **Cline, *Gasterosteus*, gene flow, isolation by distance, isolation by environment, microgeographic divergence, parapatry, resistance.**

Isolation by distance (IBD) refers to the common tendency for geographically distant populations to be more genetically differentiated than nearby ones (Wright 1943). Populations may also diverge when intervening landscape features inhibit movement between them, even if their habitats are identical and the distances are small (Isolation by Resistance; IBR (McRae 2006). Both IBD and IBR are neutral processes, which can be exaggerated by divergent natural or sexual selection that removes immigrants between habitats or removes immigrant alleles in later generations (Fisher 1930; Nosil et al. 2009; Aeschbacher and Bürger 2014). This adaptive reduction in the effective rate of gene flow can contribute to a pattern of "isolation by environment" (IBE; (Wang and Summers 2010; Sexton et al. 2014; Wang and Bradburd 2014). Consequently, a strong pattern of IBE is typically interpreted as evidence that divergent selection is maintaining population differentiation in the face of possible dispersal (Schluter 1998; Kawecki

This article corresponds to Emilie R. (2016), Digest: Sticklebacks do not just go with the flow: Genetic differentiation between lake and stream populations due to more than just geographic distance. Evolution. Doi:10.1111/evo.13138

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

and Ebert 2004). However, an often-overlooked alternative is that IBE can arise from divergent habitat choice or other forms of biased dispersal (Armsworth and Roughgarden 2008; Bolnick and Otto 2013).

A common problem in testing for IBE is that contrasting habitats may also be separated by distance and/or barriers to dispersal. For example, divergence between lake and river populations of fish may reflect their differing habitats (IBE), spatial separation (IBD), or dispersal barriers that limit movement along the river (IBR). Recently developed analytical methods partition the often-confounded patterns of IBD, IBE, and IBR when explaining genetic divergence across a landscape (Bradburd et al. 2013; Wang et al. 2013). The basic strategy is to use IBD as a null hypothesis null hypothesis against which IBE (or IBR) can be tested. Until recently the standard approach was to use partial Mantel tests, but these suffer from inflated error rates (Guillot and Rousset 2013). A recently published alternative (implemented in the software BEDASSLE) estimates the effect of habitat contrasts on genetic divergence, scaled relative to the effect of geographic distance (Bradburd et al. 2013). This unit equivalency makes it easier to quantitatively compare how much genetic divergence depends on spatial separation per se (IBD), versus environmental suppression of migration (IBE). Here, we apply BEDASSLE to estimate the relative strength of IBD versus IBE in three replicate pairs of parapatric lake and stream threespine stickleback (*Gasterosteus aculeatus*) that vary in their physical resistance to stickleback dispersal up/downstream. Although lake-stream stickleback have been intensively studied, we provide the first quantitative comparison of the effects of IBE versus IBD in this model system. To evaluate possible causes of this IBE, we then introduce a new analytical approach to test for possible cases of nonrandom dispersal.

STUDY SYSTEM

Prior studies documented genetic divergence between parapatric populations of lake and stream threespine stickleback (*Gasterosteus aculeatus*) (Lavin and McPhail 1993; Thompson et al. 1997; Hendry and Taylor 2004; Moore and Hendry 2005; Berner et al. 2009; Kaeuffer et al. 2012; Moser et al. 2012; Roesti et al. 2012; Hendry et al. 2013; For a complete list and summary see Table S1). Lake and stream stickleback consistently differ in morphology (Berner et al. 2009), sensory systems (Jiang et al., unpubl. ms.), behavior (Jiang et al. 2015), and immunity (Scharsack et al. 2007). These differences are both heritable and plastic (Berner et al. 2011; Lucek et al. 2014; Oke et al. 2015) and are associated with genome-wide differentiation (Roesti et al. 2012; Chain et al. 2014; Feulner et al. 2015). Gene flow between contiguous lake and stream populations constrains this divergence (Hendry et al. 2002; Hendry and Taylor 2004; Deagle et al. 2012; Taugbøl et al. 2014).

This repeated pattern of (often) heritable phenotypic divergence between parapatric lake and stream stickleback is widely interpreted as evidence for divergent natural selection. This interpretation is boosted by evidence for genomic regions exhibiting particularly strong differentiation relative to neutral expectations, and evidence for parallel evolution of some (but not all) phenotypes (Ravinet et al. 2012; Ferchaud and Hansen 2015; Feulner et al. 2015; Roesti et al. 2015). However, two limitations persist that weaken this conclusion.

First, comparing single sampling sites in both a lake and stream site cannot distinguish between isolation by distance and adaptive divergence. This is particularly true for more distant lake-stream comparisons (e.g., some parapatric comparisons actually entail sites many kilometers apart). Might two lakes, or two streams, separated by a comparable distance, be equally divergent? Of 73 papers we know of describing lake-stream divergence, only five studies obtained clinal genetic samples along the length of a given stream (Table S1). Consequently, although adaptive divergence-with-gene-flow seems highly likely, only a minority of studies have actually tested for isolation by distance (IBD; Moore et al. 2007; Ravinet et al. 2012; Lucek et al. 2013, 2014). None of these studies found significant IBD within a given stream (Moore and Hendry 2005), suggesting that gene flow among sites within streams is very common. Note that no studies used multiple sites within a lake. Consequently, we still lack a formal comparison of IBD versus IBE (except for one partial Mantel test of genetic divergence across habitat types vs distance [Moore et al. 2007], which as noted above is prone to false positives).

A second limitation is that the strong circumstantial evidence for adaptive divergence (inferred from parallel evolution and/or regions of high genomic divergence; see Table S1), is not backed up by clear experimental or observational support for reciprocal divergent natural selection in nature. Lake and stream stickleback have been subjected to multiple transplant experiments and mark-recapture studies to test for local adaptation or divergent selection (Hendry et al. 2002; Räsänen and Hendry 2008; Raeymaekers et al. 2010; Räsänen et al. 2012; Rolshausen et al. 2015; Stutz and Bolnick unpubl. ms.). Although most studies have found a growth, survival, or infection-resistance advantage of one ecomorph over the other, almost without exception this selection is asymmetric rather than divergent (summarized in Table S1). For instance, Moser et al. (2015) showed a strong advantage of stream fish over lake fish in the stream, but lacking a mirror image experiment in the lake were unable to conclude that there is divergent selection. An unpublished study by Stutz and Bolnick similarly found an advantage of stream fish in both habitats, which would not lead to genetic divergence. Other studies have failed to detect phenotypic selection against lake-like phenotypes in one or more streams (e.g., Rolshausen et al. 2015). It is certainly possible that such studies simply have been unlucky (conducted in years when

habitat-specific selection is low or absent), examined a life-history stage or trait on which selection is weak, or been underpowered.

Alternatively, divergent selection might be less important than it seems. First, studies may overestimate the effect of IBE if they neglect to account for IBR, which also facilitates genetic divergence at small spatial scales. Many streams containing stickleback are characterized by pool-riffle habitat with occasionally high-velocity water that may inhibit dispersal even over short distances. Past studies of lake-stream divergence have not typically quantitatively accounted for variation in resistance along the full length of a stream. Second, IBE can evolve without natural selection, if there is habitat choice or other forms of nonrandom dispersal (Bolnick and Otto 2012).

We wished to determine the strength of IBE across the lake-stream habitat boundary relative to the strength of IBD. We analyzed ddRAD genotypes (Peterson et al. 2012) with BEDASSLE (Bradburd et al. 2013) to quantitatively partition the relative effects of distance and environment in generating genetic divergence within each of three parapatric lake-stream pairs. We selected outlet streams, which tend to exhibit more gradual clines than inlets perhaps because water flow might facilitate dispersal of lake fish into the stream. The three outlet streams that we chose varied substantially in water flow rates. We expected that this variation in resistance would lead to differences in the rate of gene flow (IBR), which has not previously been quantitatively tested in lake-stream stickleback, nor distinguished from IBD or IBE. Lastly, we tested for evidence of nonrandom dispersal of morphological variants between the lake and stream, which might represent an alternative mechanism driving genetic diversification at micro-geographic spatial scales (Richardson et al. 2014).

Methods

SAMPLING AND DESCRIPTIONS OF NATURAL POPULATIONS

Threespine stickleback were sampled from three lakes (Comida, Farewell, and Roberts) and their respective outlet streams, on Vancouver Island, BC (Fig. 1, Table S2). One of these lake-stream pairs (Roberts) was previously studied by Berner et al. (2009), who used microsatellites and morphometrics to document clinal lake-stream divergence. Sampling was conducted during the active breeding season (June—early July) in each location. Within each lake we sampled sites near the outlet stream (one site in Comida, three sites in each of the other two lakes at increasing distances from the outlet). Within each stream we sampled multiple sites at increasing distances downstream from the lake (Table S1) for 1.0–1.5 km. At each site, ten unbaited minnow traps were placed within five meters of the target location. Collected fish were preserved in 10% formalin after removing a fin for DNA extraction. Sampling for this project was approved by the University of Texas IACUC (#07100201) and a scientific fish collecting permit from the British Columbia Ministry of the Environment.

The three streams vary in the severity of physical resistance to movement, allowing us to evaluate the effect of IBD and IBE across a variety of barrier strengths. To quantify this variation in resistance, we used a flow meter to measure water velocity at 5 meter intervals down the entire sampled length of each stream (Fig. S1). Flow was measured at one third the total depth, as close to the middle of the stream as we could reach. Flow was also measured at any particularly high-velocity locations between our 5-meter sites.

DNA EXTRACTION AND ddRAD LIBRARY PREPARATION

From the trapped fish, we sampled 382 stickleback for partial genome sequencing via ddRADseq (sample sizes detailed in Table S2) following Peterson et al. (2012). DNA was extracted from fin clips using the Wizard® SV 96 Genomic DNA Purification Kit (Promega Cat# A2371), per manufacturer instructions, and was fluorometrically quantified using Quant-iT™ Picogreen® dsDNA Assays (Life Technologies) and an Infinite® Pro M200 Pro microplate reader (TECAN). We digested DNA with 1 Unit of both NlaIII (NEB Cat# R0125L) and MluCI (NEB Cat# R0538L), and used AMPure SPRI beads (Beckman Coulter) for all enzymatic clean-ups. We quantified digested products, normalized each sample to 30, 60, or 90 ng, and added 5X-mass of "flex"-adapters P1 (48 different barcodes) and P2 (biotin-labeled). For ligations, we added 200 Units of T4 ligase (NEB# M0202L) and 1X-ligase buffer to a total volume of 40 μL.

Next, we pooled samples with equivalent preligation masses and unique P1-adapters. We cleaned and concentrated pools with AMPure SPRI beads, and ran products on a PippinPrep (2% agarose gel; Sage Sciences) using a "tight" extraction of DNA in the range of 371–416 bp (includes 76 bp of adapter sequence). We then used M-270 Streptavidin-coated Dynabeads® (1 μL per ~5 ng of DNA; Life Technologies) to isolate P2-labeled fragments. To flank fragments with Illumina-compatible sequences, we divided each pool into six aliquots, added "flex-P2 primers" to each pool (all aliquots from a single pool received the same P2 primer), and ran 12-cycle PCRs using the Phusion™ Polymerase kit (Life Technologies). After amplification, we combined replicate PCRs, performed a final cleanup, and examined products with a 2100 BioAnalyzer (High Sensitivity DNA chip; Agilent Technologies).

DNA SEQUENCING AND BIOINFORMATICS ANALYSIS

We sequenced ddRAD libraries on an Illumina HiSeq 2500 (Ver. 3 chemistry, PE-2×100bp with index reads; four lanes for Comida, five lanes for Farewell and Roberts together). The higher sequencing effort for Comida samples was because lower

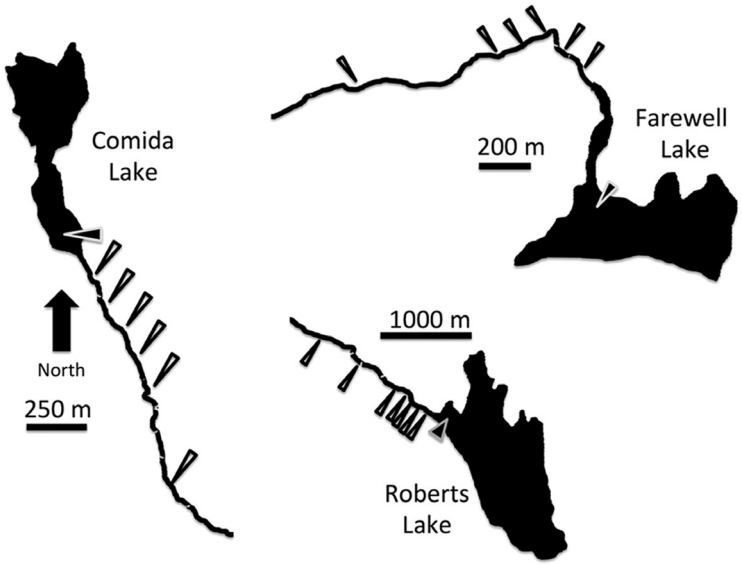

Figure 1. **Maps of the sampling locations from three lake-stream pairs on northern Vancouver Island, British Columbia. Sample sites are indicated by wedges. We plot only outlet, and not inlet, streams. The arrangement of the three lake-stream pairs in the figure is not indicative of their relative geographic locations. In Farewell and Roberts lakes we sampled three locations per lake at increasing distances from the stream.**

quality of extracted DNA required more sequencing effort to obtain roughly comparable numbers of SNPs. We used Stacks (Catchen et al. 2013) to demultiplex individuals, and then aligned raw reads to the stickleback genome (Ver. 77), first with BWA (Li and Durbin 2009) and then Stampy (Lunter and Goodson 2011). For the BEDASSLE analyses, we calculated genotype probabilities and called genotypes separately for each lake-stream population, using Samtools mpileup (v0.1.19; (Li et al. 2009)) and BCFtools (v1.1; (Danecek et al. 2011)), respectively. Using VCFtools (v0.1.12b; (Danecek et al. 2011)) we removed the sex chromosome (linkage group 19), and retained genotypes with qualities (GQ) greater than 20. We retained sites with individual read depths ≥ 12, minor allele frequencies ≥ 0.01, and only sites where $\geq 80\%$ of individuals (for a given lake-stream pair) had genotypes passing these filter requirements. Genotypes were not called at sites where individuals had fewer than 12 reads. BEDASSLE assumes that loci are unlinked, so we also randomly thinned our set of high quality SNPs to a maximum of one SNP per 100 kb. All data required for the following analyses has been archived on DRYAD (doi:10.5061/dryad.q8c13).

ANALYSIS OF GENOMIC DIVERGENCE

We calculated Weir–Cockerham's unbiased F_{ST} (Weir and Cockerham 1984) for all pairwise combinations of sample sites within each lake-stream pair. To visualize genomic divergence, we used the ANGSD software package (Korneliussen et al. 2014) to calculate genotype likelihoods (based on the original GATK approach; McKenna et al. 2010) separately for each lake-stream pair. We only used reads with a mapping quality (MapQ) of at least 30 and nucleotide Phred quality scores ≥ 20. We retained sites with read depths ≥ 8 in at least 80% of individuals and a minor allele frequency ≥ 0.05. We used these likelihoods as inputs for the ngsCovar function in the software package ngsTools (Fumagalli et al. 2013, 2014), and then used the resulting variance/covariance matrix to calculate genomic PCAs. As with the genotype likelihoods, these PCAs were calculated separately for each lake-stream pair. For each pair, we used an ANCOVA to test whether PC1 is a function of habitat and distance downstream.

ANALYSIS OF MORPHOLOGICAL DIVERGENCE

We rinsed formalin-fixed specimens, then stained them with Alizarin red. We measured mass, standard length, gape width, body depth, the number of gill rakers on the first right branchial arch, and the length of the three longest gill rakers. Traits were size-corrected by obtaining residuals from a regression of the log-transformed trait on log standard length. We used a MANOVA to test for shared and unique features of morphological divergence, testing lake-stream pair, habitat, and pair $\times$ habitat interaction effects on size-corrected traits.

PARTITIONING THE EFFECTS OF IBD AND IBE

For an initial visual evaluation of IBD and IBE, we plotted F_{ST} as a function of distance separating each pair of sites, distinguishing between within- or between-habitat contrasts. Then, we used BEDASSLE to simultaneously model the effects of IBD (increasing F_{ST} with distance) and IBE (elevated F_{ST} for across-habitat

effects). Initially, we ran BEDASSLE's Markov chain Monte Carlo (MCMC) algorithm with various random-walk tuning parameters to identify values that led to efficient mixing (see Bradburd et al. (2013) details). We subsequently used these tuning parameters to run two independent MCMC-searches per population to estimate IBD and IBE parameters that best explain variation in allele frequencies across sampling sites. Replicate MCMC runs consisted of ~10 million iterations sampled every 5000 generations. We focus our analyses on parameters aE (a proxy for IBE), aD (a proxy for IBD), and the ratio aE/aD. This ratio may be interpreted as the distance (in meters) at which IBD matches the effect size of IBE. We used diagnostic graphical functions implemented in BEDASSLE to confirm parameter convergence. We examined the fit between observed F_{ST} values and those derived from posterior predictive simulations drawn from 1000 sampled iterations of the MCMC.

In this analysis (and all others in this article), we treat habitat as a categorical variable. We are aware that this approach overlooks the potential for adaptive genetic structure within a given habitat type. Certainly, there exist environmental differences between allopatric lakes, or between allopatric streams, that create adaptive divergence between habitat replicates. Such heterogeneity could generate variation in aE among replicate pairs, but will not substantially affect our within-pair analyses. There remains the possibility, however, that IBE could arise within a given habitat category due to microhabitat variation (e.g., pool-riffle differences; Izen et al. 2016). If this within-habitat variation is spatially auto-correlated (e.g., changes monotonically along the stream), BEDASSLE would incorrectly interpret the resulting genetic divergence as an effect of IBD. We observed no obvious habitat variables that could create such bias, but this possibility must be kept in mind.

TESTING FOR AN EFFECT OF WATER FLOW ON GENETIC ISOLATION (IBR)

To estimate the strength of IBR, we quantified physical resistance to movement along the length of each lake-stream transect using the flow rates described above. We assume here that the primary deterrent to fish dispersal is water velocity (or correlated turbulence). This is justified by the biomechanical observation that the energetic cost of movement increases exponentially as a function of flow rate (Tudorache et al. 2007). We acknowledge, however, that stickleback may use boundary layers along the substrate, or shoreline refuges, to avoid the highest flow regions where we measured current velocity. Thus, our estimates represent an upper-bound estimate of the difficulty of upstream movement as we do not map velocity in fine detail across the full cross-section of the stream at each point.

To calculate resistance, we interpolated the flow rate at every meter along each stream and used this to estimate the energetic expenditure E (i.e., O_2 consumption) of an average-mass stickleback swimming against the current assuming $E = 5.48e^{0.05v}$, where v is flow velocity (from Fig. 2 of Tudorache et al. (2007)). Resistance between any two adjacent sample locations was calculated as the summed expenditure associated with the intervening water velocities. Because stickleback frequently swim against the current even when they are being displaced downstream (Jiang et al. 2015), we expect this energetic cost applies to both up- and down-stream movement, though not necessarily equally. In a circular flow-tank study, Jiang et al. (2015) showed that the individual stickleback that were displaced the farthest "down-stream" actually expended the most swimming effort, because they repeatedly tried to swim against the current, then were swept back down stream again each time. Thus, down-stream movement may, counter-intuitively, be as-costly or more-costly than remaining in place or moving up-stream.

Because of the highly skewed distribution of resistance, we used Spearman rank correlations to test whether the F_{ST} between adjacent stream sites is positively related to the resistance between those sites.

DO MIGRANTS EXHIBIT BIASED PHENOTYPES?

A three-way comparison of individuals' capture locations, genotypes, and phenotypes can be used to detect migrants, and to test whether those migrants are a random sample of phenotypes from their source population.

We used linear discriminant function analyses (DFA, using *lda* in R) to characterize the major axis of between-habitat phenotypic divergence separately for each lake-stream pair. The DFA was trained using only the most geographically extreme samples from each pair (lake fish and stream fish found ≥500 m downstream), to exclude any possible hybrid zone. We then applied the resulting discriminant axis to all fish from that lake-stream pair to obtain a posterior probability that each individual's morphology came from the lake or stream population.

Similarly, we applied a DFA to genomic data to characterize the major axis of genomic lake-stream divergence, again separately for each pair. The input data here was the first five genomic PCA axes for each lake-stream pair (Jombart et al. 2010). As with the morphological data, we trained the DFA using just the lake and farthest downstream samples, then applied the resulting axis to calculate posterior probabilities of genomic lake or stream identity.

Next, we identified migrant individuals based on mismatches between individuals' capture location and genotype (genomic DFA posterior probability of being from the lake or stream). We used this to calculate the number of immigrants into (or, equivalently, emigrants from) each habitat, for each lake-stream pair. We used a χ^2 test to evaluate whether the number of immigrant and resident genotypes differed between habitats, within

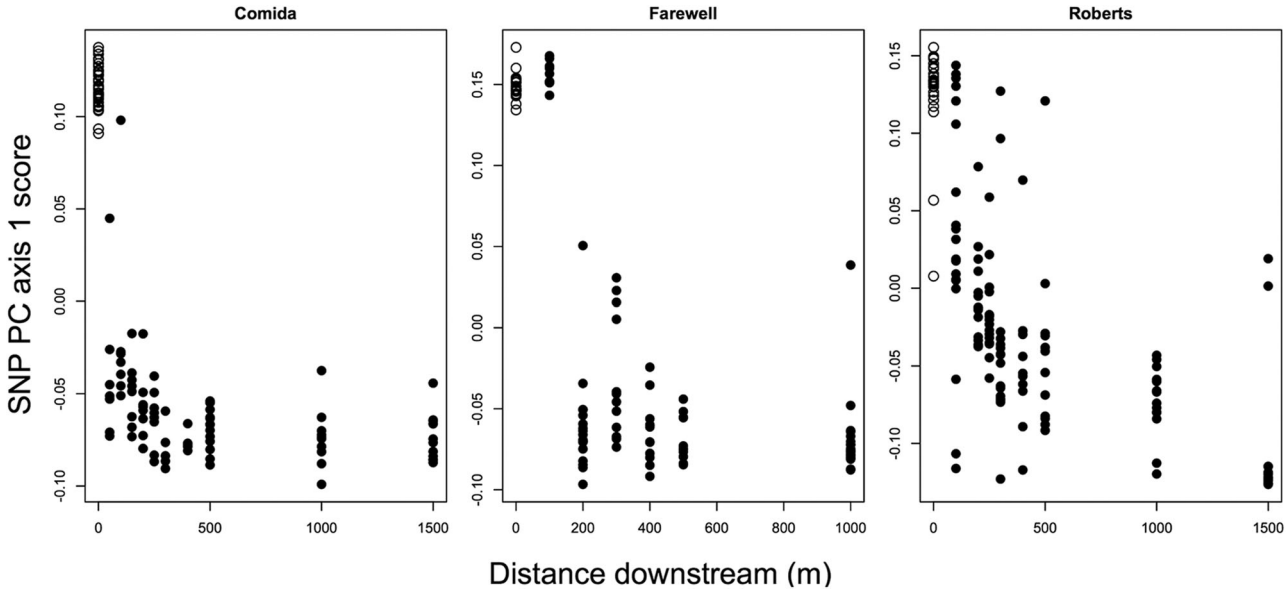

Figure 2. Genetic principal component axis 1 scores as a function of spatial location within each of the three lake-stream pairs. Fish captured in a lake are plotted as open circles, and stream fish are plotted as filled circles.

each lake-stream pair, which would indicate a directional up- or down-stream bias in the overall flux of migrants. Next, for each genotypic group (lake or stream) within each pair, we used an ANOVA to test whether the morphology (DFA axis scores) of emigrants differed from that of nonemigrants (e.g., lake residents vs lake-to-stream emigrants). Significant differences suggest that the phenotypes of dispersing individuals are a biased sample of the pool of potential migrants. Such bias could reflect either phenotype-dependent dispersal, or postdispersal adaptive plasticity.

Results

CONSTRAINTS ON FISH DISPERSAL WITHIN EACH STREAM

Farewell stream has the most substantial barrier to stickleback movement: a small rocky cascade 180 m downstream from the lake had a peak flow rate of 250 cm/sec. This cascade is fully submerged in particularly high-precipitation years but remains high-flow year-round. Multiple other areas of high flow (~100 to 150 cm/sec) occur down Farewell stream. Roberts Lake has moderate resistance, with alternating pools and riffles along the full length of the stream and maximum velocity of 130 cm/sec. Comida Stream presents no physical barriers to fish movement (maximum 7 cm/sec).

GENOMIC AND MORPHOLOGICAL DIFFERENTIATION BETWEEN HABITATS

Our genomic PCA datasets contained 7224 SNPs in Comida, 22,624 in Farewell, and 20,137 in Roberts (Table S2). After thinning datasets for the BEDASSLE analyses, we retained a median of 1654 SNPs per individual in Comida, 2325 in Farewell, and 2380 in Roberts (Table S2), with average SNP densities of 4.891/mb, 7.404/mb, and 7.484/mb, respectively. This large sampling of genomic sites provided information for both rare and common variants in all three systems, although we did exclude very rare alleles in both analyses (minor allele frequencies <0.01 and <0.05 in the BEDASSLE and PCA analyses, respectively). The lower SNP density in Comida reflects technical rather than biological processes. These samples underwent several freeze-thaw cycles before ddRAD library construction, which led to shorter starting genomic fragments and a decrease in shared loci among individuals. This technical difficulty did not appear to affect genotype calls. For example, Comida's heterozygosity is higher than in Farewell or Roberts (Fig. S2), which is opposite of what would be expected with allelic dropout.

Parapatric lake and stream populations exhibit moderate genome-wide F_{ST} values ranging from ≈ 0.02 to ≈ 0.14. Genomic PCAs of stickleback collected from each pair (Fig. 2) show that genetic differentiation occurs both between habitats (i.e., lake vs stream) and within habitats (i.e., among sites along the length of a given stream transect). In Comida, polymorphism associated with PC1 explains 7.63% of the total genetic variance, and co-varies strongly with habitat ($t = 49.8$, $P < 0.0001$). One only has to move 50 m from Comida Lake into its outlet stream to find strong divergence in PC1 scores (Fig. 2A), with only a single genetically lake-like individual sampled 100 m downstream. PC1 scores of Farewell fish (4.68% of variance) also differed strongly between lake and stream sites ($t = 15.4$, $P < 0.0001$). However, this divergence occurred between 100 m and 200 m downstream

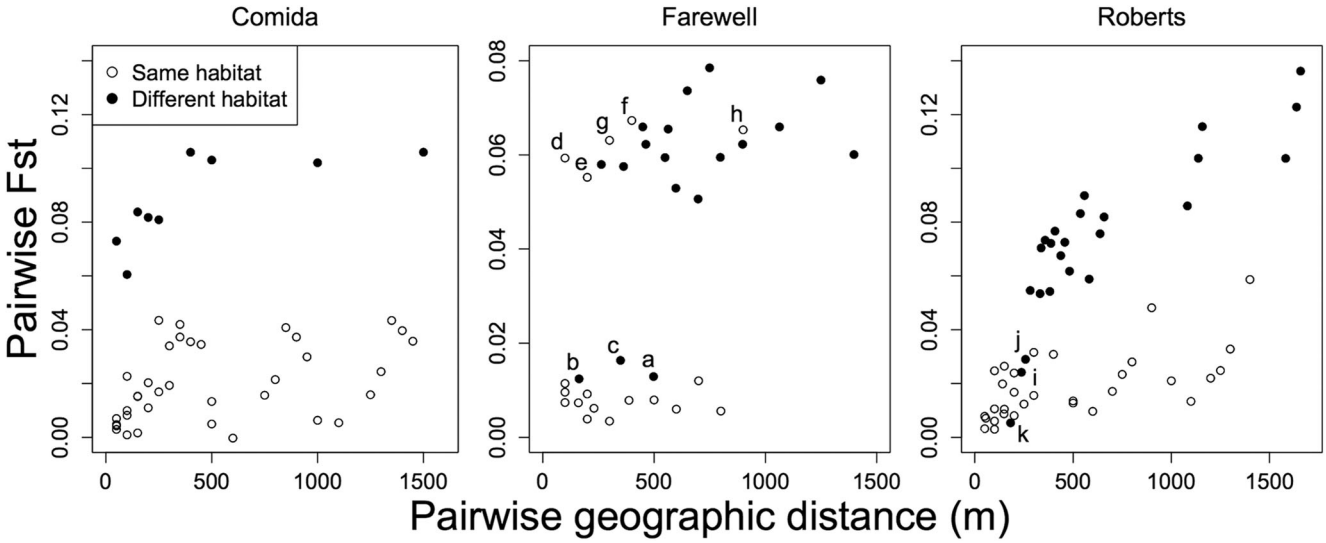

Figure 3. **Pairwise genetic divergence between sampling locations as a function of geographic distance (meters) and whether the comparison entails same or different habitats. We sampled multiple lake sites in Farewell and Roberts, providing additional resolution for these distances and allowing some within-lake comparisons. Outliers identified by BEDASSSLE analysis are indicated with letters. Farewell outlier pairs are L-100m (A–C) and 100 m × all stream sites except 1.5 km (D–H). Roberts outliers: are L-100 m (I–K).**

at a small cascade (0.5 meter elevation drop over ∼1 meter, with a peak flow rate of 275 cm/s). Above this barrier (at 100 m) stream-captured fish had lake-like PC1 scores, and below it (200 m) fish were typical stream genotypes. In Roberts Lake, PC1 (7.18% of variance) exhibits significant habitat association ($t = 8.3, P < 0.0001$), but the transition is less abrupt than the other two lake-stream pairs. Stream sites close to Roberts Lake contain likely first-generation immigrants and admixed individuals.

As is true for most lake-stream pairs, we observed morphological divergence between each parapatric population pair (Comida: Pillai's trace = 0.619, $P < 0.0001$; Farewell: Pillai's trace = 0.383, $P < 0.0001$; Roberts: Pillai's trace = 0.375, $P < 0.0001$). We find a strong correlation between morphology and genomic discriminant axis scores in all three pairs (Comida $r = 0.73$, Farewell $r = 0.56$, Roberts $r = 0.39$, all $P < 0.0001$), but this correlation results from shared spatial structure and is lost when habitat is included as a cofactor in the model (all $P > 0.3$).

ESTIMATING THE JOINT EFFECTS OF IBE AND IBD

In all three lake-stream pairs, population divergence (F_{ST}) occurs within the first ∼200 meters from the lake-stream boundaries (Figs. 2 and 3). In the Comida transect, F_{ST} is strong even between the lake and the first stream site (a mere 50 meters downstream), with weak additional divergence further downstream. Thus, visual inspection implies that IBE is strong and IBD is weak. This is remarkable given that this outlet stream has no areas of high-flow or other physical barriers to fish dispersal between habitats.

In the Farewell pair, we observe negligible genetic structure between sampling sites in Farewell Lake (F_{ST} ranges from 6.18×10^{-3} to 7.86×10^{-3}), but divergence arises rapidly between 100 and 200 meters downstream (Fig. 3B), where the small cascade is located (180 m). Farther downstream, F_{ST} continues to increase weakly with distance. This again implies both a weak effect of IBD and IBE, but the latter is confounded by the close proximity between the cascade and lake-stream transition. We address this in more detail below.

In the Roberts pair, we see a classic signature of both IBD and IBE (Fig. 3C): F_{ST} is greater between than within habitats, but increases with distance in each type of comparison. Notably, we detected weak but statistically significant IBD among the three samples within the lake: the site farthest from the stream (158 m away) is differentiated ($F_{ST} = 0.02$) from the site closest to the stream (80 m, which may receive more immigrants). Likewise, genetic divergence between stream sites increases with distance. In addition to this IBD, there is substantial IBE between the lake and stream (F_{ST} between the lake sites and the first stream sample ranges from 0.0054 to 0.0290; Fig 4).

BEDASSLE provided quantitative measures of IBD and IBE effects on population structure, indicating that both processes jointly contribute to lake-stream divergence. The output provides marginal posterior probability distributions for aD, aE, and aE/aD (respectively corresponding to IBE, IBD, and IBE/IBD). Summaries of these distributions are listed in Table 1. In all three lake-stream pairs, posterior probabilities for aD are greater than zero, suggesting that IBD exists within pairs. Comida has the lowest values of aD, at only $\sim 6e^{-5}$, while aD was approximately three times larger in Farewell, and intermediate in Roberts. This rank-ordering of IBD effects mirrors the rank-order of maximum

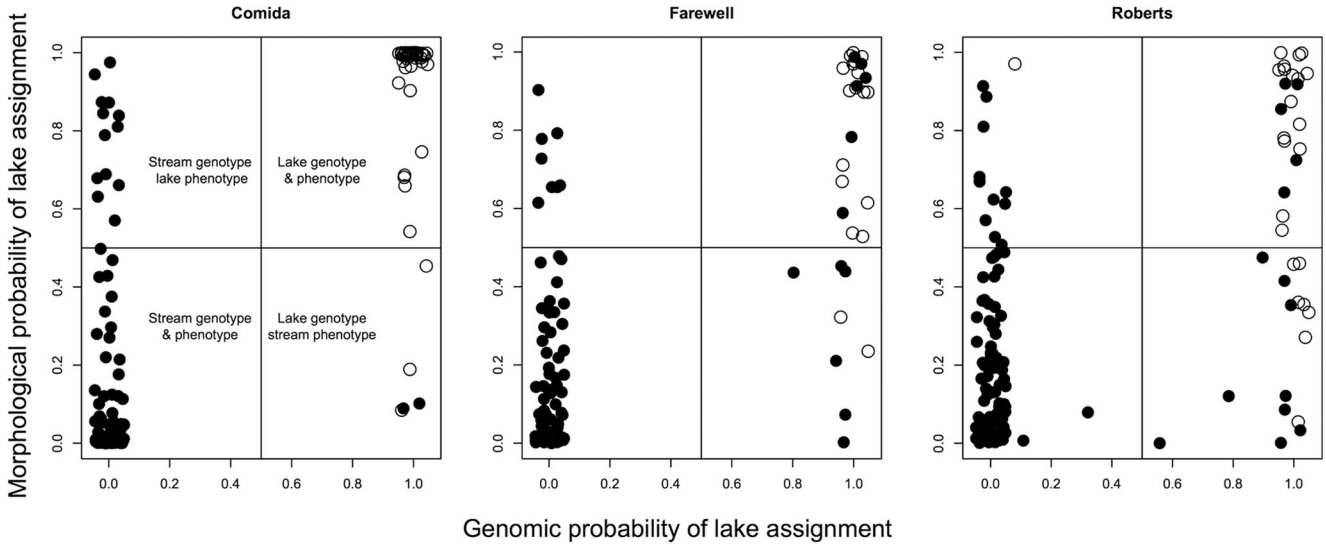

Figure 4. Habitat classification by genotype (PCA 1–10) and by phenotype. The *x* and *y* axes are the posterior probabilities of individuals being assigned to the lake ecotype based on genotype and phenotype, respectively. Linear discriminant function predictions were trained using lake fish and stream fish from ≥500 meters downstream. Individuals to the right of the vertical line would be classified as lake fish based on their genotypes; individuals to the left would be classified as stream fish. Individuals above (below) the horizontal line would be classified as lake (stream) fish based on their morphology. Open circles are fish captured from within the lake, filled circles are fish captured from within the stream. Thus, resident stream fish are predominantly filled circles in the lower left quadrant, and resident lake fish are open circles in the upper right quadrants. Emigrants are open circles in the left half, or filled circles in the right half. Some random horizontal jitter is added to the genotype scores to visually separate overlapping data points.

flow rates, though with only three pairs we cannot statistically test this association. This informal association suggests that resistance contributes to the BEDASSLE estimates of IBD, which we evaluate below. Note that BEDASSLE's aD quantifies the IBD process generating the positive slopes in Fig. 3, however this estimator cannot be directly converted into units of F_{ST} so should not be interpreted as summary statistic for Fig. 3.

In all three systems, our estimate of IBE (aE) was three orders of magnitude stronger than aD. The lake-stream transition in Comida causes genetic differentiation equivalent to moving ~2500 m within a habitat, and the boundary in Farewell and Roberts are both equivalent to moving ~1500 m. The particularly large aE/aD ratio in Comida reflects both the strong effect of habitat (Fig. 2) and the comparatively weak effect of IBD.

The model estimated by BEDASSLE provides a repeatable and good fit to our data. Independent MCMC runs provided relatively similar estimates of aD and aE in all systems, and the more important estimate of aE/aD showed high correspondence between runs (Table 1). We confirmed that the posterior predictive population genetic simulations derived from the BEDASSLE-fit model closely resemble our actual data (Fig. S3). The few sample sites where the Farewell and Roberts models inaccurately estimated F_{ST} are physically proximate to areas with high flow rate, which is not accounted for in the BEDASSLE model.

EFFECT OF WATER FLOW ON GENETIC ISOLATION (IBR)

Regression of genetic distances (F_{ST}; only between adjacent sample locations) against resistance between sample sites revealed an intriguing pattern. In the Comida lake-stream pair, F_{ST} ranged from 0.0009 to 0.0729. Flow rates were undetectably slow at all but one location (7 cm/sec; Fig. S1), and consequently there was no significant effect of water flow on F_{ST} between successive sites ($F_{1,7} = 0.002$, $P = 0.96$). The F_{ST} across the one stream segment with detectable flow ($F_{ST} = 0.015$ across a 100 m distance) is substantially less than the transition the lake and first stream site (50 m apart F_{ST} of 0.0729).

In the Farewell lake-stream pair, F_{ST} between adjoining sites ranged from 0.0061 to 0.0124. Divergence is strongest across the 100–200 m region where resistance is high due to the small cascade. Linear regression of F_{ST} on resistance is significant ($F_{1,6} = 423.1$, $P < 0.0001$, adjusted $R^2 = 0.98$) but utterly dominated by this one high-leverage contrast where the cascade is located. When we instead use Spearman rank correlation to mitigate the strong leverage of the cascade, resistance has a marginally nonsignificant effect on genetic divergence (rho = 0.68, $P = 0.062$). To put this into context, the genetic divergence attributable to this cascade (between the 100 to 200 m sites $F_{ST} = 0.059$) is nearly an order of magnitude larger than the genetic divergence between the closest lake/stream sites (–60 to 100 m along our

Table 1. Effects of IBD, IBE, and IBR in each of three stickleback lake-stream pairs.

Lake/stream Pair	MCMC Replicate	aE (95% credible int)	aD (95% credible int)	aE/aD (95% credible int)	IBR effect (Spearman's ρ & *P*)
Comida	**1**	**0.122** (0.121–0.122)	**5.31×10^{-5}** (5.28×10^{-5}–5.34×10^{-5})	**2292** (1913–2803)	$\rho = 0.4107$ $P = 0.272$
	2	**0.166** (0.164–0.169)	**7.18×10^{-5}** (7.08×10^{-5}–7.29×10^{-5})	**2312** (1964–2846)	Max F_{ST} across lake-stream boundary
Farewell	**1**	**0.422** (0.420–0.424)	**2.67×10^{-4}** (2.65×10^{-4}–2.68×10^{-4})	**1591** (1294–1964)	$\rho = 0.6826$ $P = 0.062$
	2	**0.424** (0.421–0.426)	**2.74×10^{-4}** (2.73×10^{-4}–2.76×10^{-4})	**1540** (1291–1951)	Max F_{ST} across max-resistance area
Roberts	**1**	**0.231** (0.229–0.233)	**1.45×10^{-4}** (1.44×10^{-4}–1.46×10^{-4})	**1591** (1419–1807)	$\rho = 0.03107$ $P = 0.933$
	2	**0.232** (0.230–0.234)	**1.46×10^{-4}** (1.45×10^{-4}–1.47×10^{-4})	**1595** (1405–1811)	Max F_{ST} displaced ~150 m downstream of habitat boundary

Columns two through five summarize the results of duplicate BEDASSLE analyses of each parapatric pair. We provide estimates of the means and 95% credible intervals of BEDASSLE model parameters aE, aD, and aE/aD. Estimates were calculated from MCMC posterior distributions, and include replicate MCMC samples (2800 generations each) for all lake-stream systems. The ratio aE/aD indicates the spatial distance (in meters) at which IBD would equal the effect of IBE. In addition, we present a test of isolation by resistance (IBR) carried out by testing for a rank correlation of F_{ST} between adjacent sample locations, with the resistance to dispersal between these sites (estimated from flow rates). We provide the Spearman rank correlation ρ and *P*, as well as a comment indicating what habitat features if any occur along the stream area where F_{ST} is maximal.

cline, $F_{ST} = 0.0062$) or between the most-distant two adjoining points (between the 500 m and 1000 m sites, $F_{ST} = 0.0079$).

In the Roberts pair, adjoining sample locations' F_{ST} ranged from 0.003 to 0.025, and these were uncorrelated with distance (maximum distance of 500 m, rho = 0.420, $P = 0.227$) or resistance (rho = 0.031, $P = 0.933$). Note, again, that the former result is a weak test of IBD because it does not use the full range of geographic distances, being restricted to adjacent sample points. In this cline, the area of maximum genetic divergence (between the 100 m to 200 m sites, $F_{ST} = 0.024$) is neither the stream region with the highest resistance (500–1000 m downstream, $F_{ST} = 0.0135$), longest distances (500–1000 m: $F_{ST} = 0.0135$, or 1000—1500 m, $F_{ST} = 0.0127$), nor the lake-stream boundary itself (0.0054–0.0087, depending on which lake site is used as a reference). Notably, the elevated genetic divergence between the 1000 m and 1500 m sites occurs although they are in the same habitat and have zero resistance between them (the stream is very slow-moving in that region). This illustrates a detectable role of IBD within a single stream, occurring even within a single 500-m stretch of a stream.

MIGRANTS EXHIBIT BIASED PHENOTYPES

Combining genetic and phenotypic discriminant functions, we correctly assigned a high proportion of fish to their habitat-of-capture based on both criteria (Fig. 4; 112/133 fish from Comida, 80/99 in Farewell; 106/131 in Roberts). In Comida, 41 of 44 lake-captured fish were classified as lake fish based on both morphology and genotype, and 71 of 80 stream-captured fish were classified as stream fish by both criteria. In Farewell Lake and Stream, three-way matches between origin, genotype, and morphology were found for 16/18 lake and 64/81 stream fish (in Roberts: 16/24 and 90/107). Averaging across all habitats and pairs, 82% of the fish were both genetically and phenotypically assigned to the habitat from which they were captured.

The converse, fish whose capture site is mismatched to their genotype and/or phenotype, provides an opportunity to estimate (i) the rate of dispersal and introgression between habitats, and (ii) phenotypic bias of dispersing individuals. We begin with the reasonable assumption that a mismatch between capture site versus genotype is indicative of migrants (or, in a minority of instances, F1 or later-generation hybrids). In the Comida Lake-stream pair, 0/44 fish captured in the lake were inferred to be stream genotypes and 2/89 fish captured in the stream were inferred to be lake genotypes. Overall, only 1.5% of Comida samples were inferred to be migrants. Given the paucity of migrants, we had little power to detect up/down-stream bias in movement ($\chi^2 = 0.05$, $P = 0.82$). In the Farewell lake-stream pair, 11.7% of fish exhibited a genotype-habitat mismatch: 0 of 18 fish captured in the lake were inferred to be immigrants from the stream, and 12 of 81 fish captured in the stream were inferred to have come from the lake (10 of these were found in the 100 m site). We observed no significant bias towards up- or down-stream migration ($\chi^2 = 1.8$, $P = 0.18$). In the Roberts lake-stream pair, 10.8% of fish were inferred to be migrants (1/24 fish captured in the lake, 14/107 fish captured in the stream). Again, we observe no significant up- or down-stream migration bias ($\chi^2 = 0.88$, $P = 0.35$) in Roberts lake and stream.

Next, we reexamined each lake-stream pair to ask whether migrants' morphology was a random subsample of source population. In Comida we identified no stream-to-lake migrants, and only two lake-to-stream migrants. Despite the low sample size of emigrants, these two individuals were significantly different than their relatives who remained in the lake (Fig. 4A; $F_{1,44} = 27.7$, $P < 0.001$). Similarly, in Farewell we only found lake-to-stream migrants (10/12 of these above the cascade), and these tended to be morphologically more stream-like than expected (Fig. 5A), but this trend was marginally nonsignificant ($F_{1,28} = 4.06$, $P = 0.054$). The same pattern held for Roberts (Fig. 5B). Lake to stream migrants were significantly more stream-like than their lake relatives ($F_{1,35} = 9.83$, $P = 0.003$), and the single stream emigrant found in Roberts Lake had by far the most lake-like morphology of any genetically stream-like fish ($F_{1,100} = 9.78$, $P = 0.002$). The implication is that emigrants' phenotypes are better predicted by their destination habitat than by their ancestry.

Discussion

Parapatric lake and stream populations of threespine stickleback provide an excellent example of parallel phenotypic and genetic divergence with gene flow and consequently have received substantial attention (Table S1). The speed, parallelism, and magnitude of lake-stream divergence certainly suggest an effect of divergent natural selection. Yet, the majority of studies that reach this conclusion rely on single sample points per habitat and consequently cannot quantitatively test their results against a null expectation of isolation by distance. In addition, evidence for divergent sexual or natural selection is still surprisingly ambiguous. Laboratory mate choice trials revealed only weak and highly asymmetric mate preferences between ecotypes (Raeymaekers et al. 2010), whose breeding seasons overlap extensively (Hanson et al. 2015). Numerous transplant experiments or mark-recapture studies have failed to find evidence of reciprocal local adaptation; both ecotypes do not clearly achieve higher fitness in their respective habitats. Instead, studies have either failed to find a home-site advantage (Moore and Hendry 2009), or found that one ecotype is more fit than the other in one (Moser et al. 2015) or both habitats (Räsänen and Hendry 2008). This ambiguous support for divergent selection and local adaptation could of course reflect low statistical power to detect weak selection or variation in selection across life history stages, seasons, or years. However, we must also consider the possibility that divergent selection is weaker than commonly thought, and that genetic divergence is instead maintained mostly by other processes such as isolation by distance, isolation by dispersal barriers ('resistance'), or non-random dispersal.

Using a new statistical approach implemented in BEDASSLE, we provide the strongest statistical support to-date for the "party-line" view that genetic divergence between lake and stream fish is mostly due to habitat-associated reproductive isolation (IBE). Unlike prior studies, we find evidence of isolation by distance within streams, but conclude that this effect is weak compared to IBE. Resistance is not required for lake-stream divergence (e.g., Comida), but very high-flow areas can become the focal point for lake-stream clines (e.g., Farewell). These inferences do not necessarily prove that there is divergent selection, however, because IBE might also arise from biased dispersal. Consistent with this biased dispersal, we find that emigrants are a phenotypically non-random subset of their source population.

MICROGEOGRAPHIC DIVERGENCE

Recently, interest in "microgeographic" divergence has grown, as evolutionary biologists increasingly appreciate the potential for genetic and phenotypic differentiation within individuals' dispersal neighborhood (Richardson et al. 2014). Our study reveals significant genetic differentiation between habitats over as little as 50 meters (in the case of Comida Lake and Stream), without any physical barriers to movement. By about 200 meters downstream, all three lake-stream pairs exhibited F_{ST} values of 0.06 or higher (up to almost 0.14). For comparison, genome-wide mean F_{ST} for the well-known benthic and limnetic stickleback species pairs range from $F_{ST} = 0.15$ to 0.2 (Diana Rennison, pers. comm.). Thus, in our system, 200 meters' distance across a lake-stream habitat boundary is sufficient to impart about

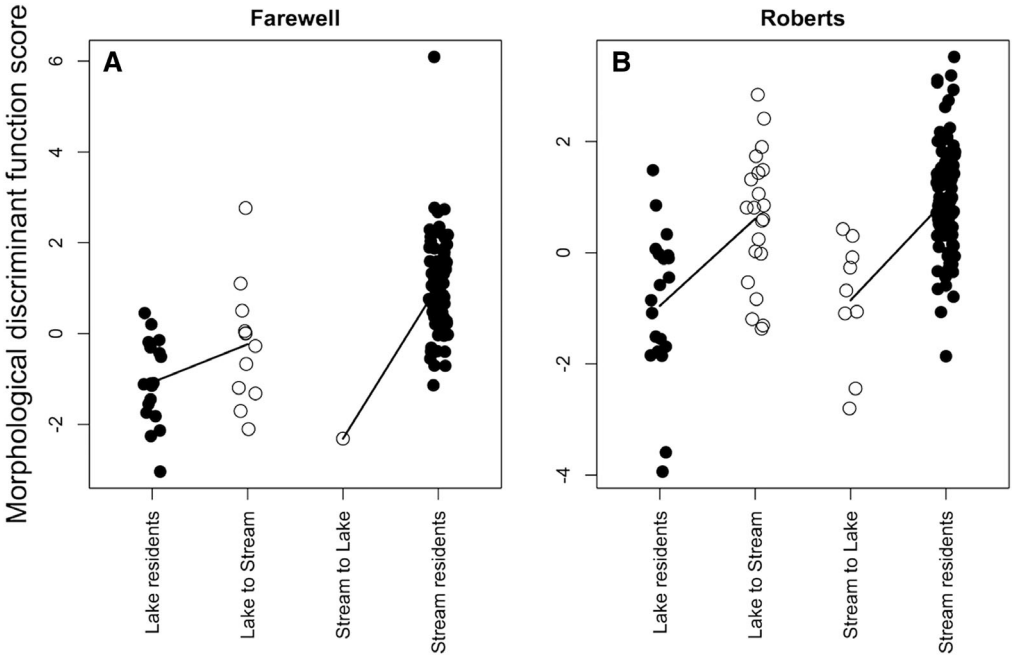

Figure 5. **Emigrants (open circles) are phenotypically different from the residents of their ancestral habitat (filled circles), for (A) Farewell, and (B) Roberts lake-stream pairs. Residents are those individuals whose genetic DFA posterior probabilities match their capture habitat, whereas emigrants are individuals whose genetic DFA posterior probability supports ancestry in the noncapture habitat. Lines connect the phenotypic means of each group. We do not plot data from Comida because there were so few migrants.**

one-third to one-half of the genetic divergence corresponding to widely recognized stickleback species pairs.

This small spatial scale of divergence is appreciably smaller than previous studies of lake-stream stickleback have reported. For example, Berner et al. (2009) also documented clinal genetic divergence between Roberts Lake and Stream, but they described the hybrid zone as being ~550 meters downstream, compared to our ~150. Other lake-stream pairs in their study exhibited genetic structure after progressing longer distances downstream: 2.3 km (Beaver), 0.61 km (Boot), 2.3 km (Joe's), 2.9 km (McCreight), 1.7 km (Pye). Ravinet et al. (2012) also report genetic clines in multiple streams feeding into a lake in Ireland, with divergence arising on spatial scales exceeding a kilometer. Our denser SNP dataset suggests that clines arise at smaller spatial scales than previously appreciated, at least for outlet streams where water flow is argued to facilitate gene flow from the lake into the stream (Hendry and Taylor 2004).

Such divergence at our fine spatial scale is especially striking given that previous mark-recapture studies estimate that stickleback typically move similar distances within the span of days to weeks (Moore and Hendry 2005; Bolnick et al. 2009). Marked fish released into a stream exhibited median and maximal dispersal distances of roughly 50 and 200 meters, respectively in a few days to weeks (Bolnick et al. 2009; Moore and Hendry 2009). This mobility is consistent with our evidence that between 0.7% and 19% of fish in our samples were migrants between habitats (first- or later-generation), and with the weak but nonzero effect of IBD in all three pairs (aD estimates in Table 1). Dispersal should therefore largely erase genetic differentiation at such small spatial scales. Thus, spatial dispersal limitation alone is not sufficient to explain F_{ST} values of 0.06–0.08 between lake fish and stream sites a few hundred meters apart.

ISOLATION BY RESISTANCE

One possible explanation for this abrupt lake-stream divergence invokes physical resistance to movement (IBR) due to cascades or areas of exceptionally high flow. We tested this by producing a detailed (5-meter resolution) map of flow rate variation along the 1.0 to 1.5 km length of each of our clines, and testing whether F_{ST} between adjacent sites is positively related to the intervening water flow. To our knowledge, no other study of lake-stream stickleback divergence has reported such a trend. Although we find marginally nonsignificant support for such a trend in Farewell stream, the other two pairs exhibit no such correlation. Moreover, the most spatially abrupt genetic divergence (at the Comida lake-stream boundary) occurs in the pair with almost no detectable water velocity across the entire transect. This demonstrates that physical barriers to migration are not necessary to maintain lake-stream divergence. We also detected no systematic tendency for migration to be higher from the lake into the outlet stream (favored by currents) than upstream from the stream into the lake. The lack of asymmetry is not too surprising, because stickleback exhibit

varying degrees of rheotaxis (a tendency to swim up-current) that might offset any downstream bias (Jiang et al. 2015). Stream fish in particular are effective at using low-flow boundary layers in moving water, to retain their position or move up-current (Jiang et al. 2015).

Our results for Farewell stream deserve some additional consideration: in this instance, the largest incremental change in genetic divergence along the stream coincided with a region with a small cascade, a site of exceptionally high resistance to stickleback movement. Hybrid zone theory suggests that preexisting or emerging genetic clines, arising from selection or IBD, will be anchored at any landscape features that inhibit movement (Barton and Hewitt 1989). Thus, the small cascade in Farewell stream might be a primary cause of divergence (through IBR) or may serve as a focal point for a cline arising from selection against migrants (IBE).

ISOLATION BY ENVIRONMENT

Having ruled out IBD and IBR as sufficient explanations for microgeographic lake-stream divergence, we are left with isolation by environment (IBE) as an alternative. Consistent with this conclusion, estimates of aE were all relatively large, and the 95% credible interval did not approach zero, for all three lake-stream pairs. The ratio aE/aD allows us to interpret this IBE effect in intuitive terms: a geographic distance at which the effect of IBD would equal IBE. In Comida, the habitat boundary is equivalent to 2300 meters of isolation by distance (IBD was weakest in Comida, of the three pairs). In both Roberts and Farewell, the habitat boundary is equivalent to 1500 meters of spatial isolation. We can therefore infer that parapatric lake and stream stickleback experience, or have experienced, reduced gene flow due to their distinct habitats and that IBD is relatively weak compared to IBE.

At least four processes might generate the IBE documented here: (1) recent, secondary contact between previously isolated populations; (2) natural or sexual selection against migrants; (3) nongenetic (i.e., phenotypically plastic) habitat-biased dispersal; and (4) genotype-dependent dispersal. Each of these processes could plausibly create sharp genetic breaks across the lake-stream boundary well within the typical lifetime dispersal range of individual fish, and without barriers to movement. We discuss each in turn below.

Recent secondary contact between previously isolated populations is unlikely in this system. Numerous genetic studies have found close sister relationships between parapatric lake-stream populations on Vancouver Island, consistent with the colonization history of the region dictated by its glacial history (Berner et al. 2009; Hendry et al. 2013; Stuart et al. Submitted).

As noted above, experimental studies of lake and stream stickleback on Vancouver Island have failed to detect reproductive isolation due to reciprocal divergent selection or reciprocal premating isolation (Hendry et al. 2002; Räsänen et al. 2012; Stutz and Bolnick unpubl. ms.). This does not rule out the possibility of weak or temporally variable divergent selection. Intuitively, it seems necessary to invoke at least some divergent selection to explain the IBE that we observe, especially given the number of putative migrants in our sample (individuals with habitat-genotype mismatches). But, selection would have to be exceptionally strong to remove migrants quickly enough to generate divergence across 50 meters as we observe in Comida, when individuals readily disperse this distance within a few days. Past experiments should have been able to detect such strong selection.

Environmentally induced or genotype-dependent dispersal preferences could both plausibly generate the observed sharp genetic divergence between lake and stream fish. Biased dispersal is a fast-acting and effective cause of reproductive isolation that can dramatically reduce migration rates over small spatial scales relative to movement capacity (Bolnick and Otto 2013). Divergent habitat choice has been documented in lake-stream stickleback: experimentally displaced fish preferentially return to their native habitat (Bolnick et al. 2009). This nonrandom dispersal is mediated by individuals' body shape and lateral line sensory system (Jiang et al. 2015; Jiang et al., unpubl. ms.). Our analysis reveals a pattern consistent with this biased dispersal (Fig. 5). Individuals identified as putative migrants (e.g., captured in one habitat but genetically assigned to the other habitat) were morphologically more like the residents of their new habitat, than their relatives who remained home. That is, lake fish that emigrated to the stream had disproportionately stream-like morphology relative to other lake fish. This observational pattern recapitulates a previous experimental result: experimentally displaced stickleback were more likely to switch habitats if they were morphologically predisposed to suit their nonnative habitat (Bolnick et al. 2009). The concordance between our observational result and previous experimental findings strongly suggests that phenotype-biased dispersal occurs in stickleback and facilitates lake-stream divergence. However, an alternative is also plausible: individuals might migrate randomly, and subsequently adjust their phenotype after entering a new habitat to resemble local natives. Lab experiments have repeatedly confirmed that lake and stream stickleback exhibit moderate phenotypic plasticity (Oke et al. 2015; Moser et al. 2016). Further experiments are required to determine the relative order of events in generating these genotype versus trait and habitat mismatches.

SUMMARY

A growing literature partitions the relative effects of geographic distance and environmental differences (Bradburd et al. 2013; Wang and Bradburd 2014). Here, we document modest but significant patterns of both IBD and IBE between stickleback inhabiting continguous lake and stream habitats. Consistent with

the notion of "microgeographic" divergence (sensu (Richardson et al. 2014)), we find that IBE is substantially stronger than IBD. This finding supports long-standing views about the evolution of parapatric lake and stream stickleback, using new quantitative methods. However, we also show that lake-stream divergence occurs at even finer spatial scales than commonly appreciated. Relatively mobile stickleback can maintain strong genetic and phenotypic divergence over distances of less than 50 meters (e.g., in Comida lake-stream pair). Such dramatic microgeographic divergence is clearly maintained whether or not there are physical barriers to movement, and in the face of substantial migration. The natural conclusion is that there must be strong behavioral or selective barriers, or both. We present some data consistent with a behavioral barrier arising from phenotypically biased dispersal.

ACKNOWLEDGMENTS

We thank Scott Hunicke-Smith and the staff at the University of Texas Genome Sequencing and Analysis Facility (GSAF) for assistance. Field and lab work benefitted from help by Claire Patenia, Todasporn Rodbumrung, Chris Harrison, Kimberly Hendrix, Thor Veen, Cole Thompson, Chad Brock, and Newaz Ahmed. The manuscript benefitted from advice of two reviewers. The research was supported by the Howard Hughes Medical Institute and National Science Foundation grants DEB-1144773 to D.I.B. and DEB-1402725 to G.S.B. The authors have no conflict of interest to declare.

DATA ARCHIVING

The doi for our data is 10.5061/dryad.q8c13.

LITERATURE CITED

Aeschbacher, S., and R. Bürger. 2014. The effect of linkage on establishment and survival of locally beneficial mutations. Genetics 197:317–336.

Armsworth, P. R., and J. E. Roughgarden. 2008. The structure of clines with fitness-dependent dispersal. Am. Nat. 172:648–657.

Barton, N. H., and G. M. Hewitt. 1989. Adaptation, speciation and hybrid zones. Nature 341:497–502.

Berner, D., A.-C. Grandchamp, and A. P. Hendry. 2009. Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. Evolution 63:1740–1753.

Berner, D., R. Kaeuffer, A.-C. Grandchamp, J. A. M. Raeymaekers, K. Rasanen, and A. P. Hendry. 2011. Quantitative genetic inheritance of morphological divergence in a lake-stream stickleback ecotype pair: implications for reproductive isolation. J. Evol. Biol. 24:1975–1983.

Bolnick, D., and S. Otto. 2013. The magnitude of local adaptation under genotype-dependent dispersal. Ecol. Evol. 3:4733–4735.

Bolnick, D. I., L. K. Snowberg, C. Patenia, W. E. Stutz, T. Ingram, and O. L. Lau. 2009. Phenotype-dependent native habitat preference facilitates divergence between parapatric lake and stream stickleback. Evolution 63:2004–2016.

Bradburd, G. S., P. L. Ralph, and G. M. Coop. 2013. Disentangling the effects of geographic and ecological isolation on genetic differentiation. Evolution 67:3258–3273.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22:3124–3140.

Chain, F. J., P. G. Feulner, M. Panchal, C. Eizaguirre, I. E. Samonte, M. Kalbe, T. L. Lenz, M. Stoll, E. Bornberg-Bauer, M. Milinski, et al. 2014. Extensive copy-number variation of young genes across stickleback populations. PLoS Genet. 10:e1004830.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, et al. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Deagle, B. E., F. C. Jones, Y. F. Chan, D. M. Absher, D. M. Kingsley, and T. E. Reimchen. 2012. Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. Proc. R Soc. Lond. Ser B 279:1277–1286.

Ferchaud, A.-L., and M. M. Hansen. 2016. The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: three-spine sticklebacks in divergent environments. Mol. Ecol. 25:238–259.

Feulner, P. G., F. J. Chain, M. Panchal, Y. Huang, C. Eizaguirre, M. Kalbe, T. L. Lenz, I. E. Samonte, M. Stoll, E. Bornberg-Bauer, et al. 2015. Genomics of divergence along a continuum of parapatric population differentiation. PLoS Genet. 11:e1004966.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford University Press. New York.

Fumagalli, M., F. G. Vieira, T. S. Korneliussen, T. Linderoth, E. Huerta-Sánchez, A. Albrechtsen, and R. Nielsen. 2013. Quantifying population genetic differentiation from next-generation sequencing data. Genetics 195:979–992.

Fumagalli, M., F. G. Vieira, T. Linderoth, and R. Nielsen. 2014. ngsTools: methods for population genetics analyses from next-generation sequencing data. Bioinformatics 30:1486–1487.

Guillot, G., and F. Rousset. 2013. Dismantling the mantel tests. Methods Ecol. Evol. 4:336–344.

Hanson, D., R. D. H. Barrett, and A. P. Hendry. 2015. Testing for parallel allochronic isolation in lake-stream stickleback. J. Evol. Biol. 29:47–57.

Hendry, A. P., R. E. Kaeuffer, E. Crispo, C. L. Peichel, and D. I. Bolnick. 2013. Evolutionary inferences from exchangeability: individual classification approaches based on the ecology, morphology, and genetics of lake-stream stickleback population pairs. Evolution 67:3429–3441.

Hendry, A. P. and E. B. Taylor. 2004. 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.

Hendry, A. P., E. B. Taylor, and J. D. McPhail. 2002. Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. Evolution 56:1199–1216.

Izen, R., Y. E. Stuart, Y. Jiang, and D. I. Bolnick. 2016. Coarse- and fine-scale phenotypic variation in three-spine stickleback inhabiting an alternating series of lake and stream habitats. Evol. Ecol. Res. 17:437–457.

Jiang, Y., C. L. Peichel, and D. I. Bolnick. In review. Heritable variation in lateral line sensory systems mediates rheotactic response of lake and stream stickleback. J. Evol. Biol.

Jiang, Y., L. Torrance, C. L. Peichel, and D. I. Bolnick. 2015. Divergent rheotaxis contributes to divergent habitat preferences between lake and stream threespine stickleback. Evolution 69:2517–2524.

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11:94.

Kaeuffer, R., D. Bolnick, A. P. Hendry, and C. L. Peichel. 2011. Convergence and non-convergence in ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. Evolution 66:402–418.

Kawecki, T. J., and D. Ebert. 2004. Conceptual issues in local adaptation. Ecol. Lett. 7:1225–1241.

Korneliussen, T., A. Albrechtsen, and R. Nielsen. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15:356.

Lavin, P. A., and J. D. McPhail. 1993. Parapatric lake and stream sticklebacks on northern Vancouver Island—disjunct distribution or parallel evolution. Can. J. Zool. 71:11–17.

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

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and S. Genome Project Data Processing. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Lucek, K., A. Sivasundar, D. Roy, and O. Seehausen. 2013. Repeated and predictable patterns of ecotypic differentiation during a biological invasion: lake-stream divergence in parapatric Swiss stickleback. J. Evol. Biol. 26:2691–2709.

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

Lunter, G., and M. Goodson. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res. 21:936–939.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20:1297–1303.

McRae, B. H. 2006. Isolation by resistance. Evolution 60:1551–1561.

Moore, J.-S., and A. P. Hendry. 2005. Both selection and gene flow are necessary to explain adaptive divergence: evidence from clinal variation in stream stickleback. Evol. Ecol. Res. 7:871–886.

Moore, J.-S., J. L. Gow, E. B. Taylor, and A. P. Hendry. 2007. Quantifying the constraining influence of gene flow on adaptive divergence in the lake-stream threespone stickleback system. Evolution 61:2015–2026.

Moore, J.-S., and A. P. Hendry. 2009. Can gene flow have negative demographic consequences? Mixed evidence from stream threespine stickleback. Philos. Trans. R Soc. B 364:1533–1542.

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

Moser, D., M. Roesti, and D. Berner. 2012. Repeated lake-stream divergence in stickleback life history within a Central European lake basin. PLoS ONE 7:e50620.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. Divergent selection and heterogeneous genomic divergence. Mol. Ecol. 18:375–402.

Oke, K., M. Bukhari, R. Kaueffer, G. Rolshausen, K. Rasanen, D. I. Bolnick, C. L. Peichel, and A. P. Hendry. 2015. Plasticity enhances phenotypic parallelism: evidence from lake-stream stickleback. J. Evol. Biol. 29:126–143.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135.

Raeymaekers, J. A. M., M. Boisjoly, L. Delaire, D. Berner, K. Räsänen, and A. P. Hendry. 2010. Testing for mating isolation between ecotypes: laboratory experiments with lake, stream, and hybrid stickleback. J. Evol. Biol. 23:2694–2798.

Räsänen, K., M. Delcourt, L. J. Chapman, and A. P. Hendry. 2012. Divergent selection and then what not: the puzzle of missing reproductive isolation in Misty lake and stream stickleback. Int. J. Ecol. 2012:902438.

Räsänen, K., and A. P. Hendry. 2008. Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. Ecol. Lett. 11:624–626.

Ravinet, M., P. A. Prodôhl, and C. Harrod. 2012. Parallel and nonparallel ecological, morphological and genetic divergence in lake-stream stickleback from a single catchment. J. Evol. Biol. 26:186–204.

Richardson, J. L., M. C. Urban, D. Bolnick, and D. K. Skelly. 2014. Microgeographic adaptation and the spatial scale of evolution. Trends Ecol. Evol. 29:165–176.

Roesti, M., A. P. Hendry, W. Salzburger, and D. Berner. 2012. Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. Mol. Ecol. 21:2852–2862.

Roesti, M., B. Kueng, D. Moser, and D. Berner. 2015. The genomics of ecological vicariance in threespine stickleback fish. Nat. Comm. 6:8767.

Rolshausen, G., S. Muttalib, R. Kaueffer, K. Oke, D. Hanson, and A. P. Hendry. 2015. When maladaptive gene flow does not increase selection. Evolution 69:2289–2302.

Scharsack, J. P., M. Kalbe, C. Harrod, and G. Rauch. 2007. Habitat-specific adaptation of immune responses of stickleback (*Gasterosteus aculeatus)* lake and river ecotypes. Proc. R Soc. B Biol. Sci. 274:1523–1532.

Schluter, D. 1998. Ecological causes of speciation. Pp. 114–129 *in* D. J. Howard, and S. H. Berlocher, eds. Endless forms. Oxford Univ. Press, New York.

Sexton, J. P., S. B. Hangartner, and A. A. Hoffmann. 2014. Genetic isolation by environment or distance: which pattern of gene flow is most common? Evolution 68:1–15.

Stuart, Y. E., T. Veen, J. N. Weber, D. Hanson, B. K. Lohman, M. Ravinet, C. J. Thompson, T. Tasneem, A. Doggett, R. Izen, N. Ahmed, et al. In revision. Contrasting effects of environment and genetics generate a predictable continuum of parallel evolution. Nat. Ecol. Evol.

Stutz, W. E. and D. I. Bolnick. Manuscript. Lack of local adaptation in a reciprocal transplant experiment between parapatric lake and stream stickleback.

Taugbøl, A., C. Junge, T. P. Quinn, A. Herland, L. A. Vøllestad. 2014. Genetic and morphometric divergence in threespine stickleback in the Chignik catchment, Alaska. Ecol. Evol. 4:144–156.

Thompson, C. E., E. B. Taylor, and J. D. McPhail. 1997. Parallel evolution of lake-stream pairs of threespine sticklebacks (*Gasterosteus*) inferred from mitochondrial DNA variation. Evolution 51:1955–1965.

Tudorache, C., R. Blust, and G. De Boeck. 2007. Swimming capacity and energetics of migrating and non-migrating morphs of three-spined stickleback *Gasterosteus aculeatus* L. and their ecological implications. J. Fish Biol. 71:1448–1456.

Wang, I. J., and G. S. Bradburd. 2014. Isolation by environment. Mol. Ecol. 23:5649–5662.

Wang, I. J., R. E. Glor, and J. B. Losos. 2013. Quantifying the roles of ecology and geography in spatial genetic divergence. Ecol. Lett. 16:175–182.

Wang, I. J., and K. Summers. 2010. Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. Mol. Ecol. 19:447–458.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370.

Wright, S. 1943. Isolation by distance. Genetics 28:114–138.

Associate Editor: I. Lovette
Handling Editor: M. Servedio

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. Maps of the three lake-stream pairs with flow rates indicated by colored dots (different scale bars are used for each panel)
Figure S2. Variation in mean heterozygosity along the length of each of the three lakestream clines
Figure S3. Posterior predictive sampling to evaluate quality of BEDASSLE model fit.
Table S1. (separate file). A summary of findings from most published 36 studies (and a few manuscripts which we know of) describing comparisons of lake versus stream threespine stickleback
Table S2. Summary of sampling locations and sample sizes for each of the three lake stream pairs.
Table S3. Summary of number of sites and sequencing depths for all analyses.