Evolutionary Ecology Research, 2016, **17**: 437–457

Coarse- and fine-grained phenotypic divergence among threespine stickleback from alternating lake and stream habitats

Rebecca Izen[1,2], Yoel E. Stuart[2], Yuexin Jiang[2,3] and Daniel I. Bolnick[2]

[1]*Genetics and Developmental Biology Center, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA,* [2]*Department of Integrative Biology, University of Texas at Austin, Austin, Texas, USA and* [3]*Airbnb, San Francisco, California, USA*

ABSTRACT

Background: Habitat characteristics can vary over small spatial scales at which gene flow is expected to swamp any effect of divergent natural selection. However, fine-grained ('microgeographic') adaptive divergence may still be feasible if individuals exhibit dispersal behaviours that improve the match between their phenotype and habitat. For example, threespine stickleback (*Gasterosteus aculeatus*) from lake and stream habitats maintain differences across a narrow ecotone because of non-random gene flow. However, it is unknown whether dispersal bias might also contribute to even finer-scale divergence within habitats, in response to microhabitat variation within lakes and within streams.

Question: Does stickleback morphology co-vary with flow regime within stream populations, controlling for distance from adjoining lake populations?

Data: We sampled stickleback along a transect through alternating lake and stream habitats. Within each stream, multiple traps were set at 50 m intervals. We recorded microhabitat data (flow rate and depth) at each trap. We measured morphology (gill rakers, head shape, fin shape, standard length) of more than 900 stickleback captured from these traps.

Analysis: We used multivariate analyses of covariance and linear models to test for: (1) phenotypic divergence between lake and stream stickleback, (2) divergence among stream sites as a function of their distance from an adjoining lake, and (3) covariation between local flow regime (at each trap) and the morphology of stickleback captured from that trap.

Conclusions: Fish from different flow regimes within a stream show phenotypic variation that is not due to clinal transitions from lake to stream. We found covariation between local flow regime and either fin morphology or gill raker length in different streams. The total effect size of stream microhabitat on morphology was greater than the effect size of habitat (lake vs. stream), for overall multivariate data and for a subset of univariate traits. These findings imply that local adaptation can occur on a finer spatial scale than is typically expected, perhaps as a result of non-random dispersal.

Keywords: cline, *Gasterosteus aculeatus*, habitat preference, matching habitat choice, microgeographic variation, threespine stickleback.

Correspondence: R.M. Izen, National Institutes of Health, 10 Center Drive, Building 10, Room 6C103, Bethesda, MD 20892, USA. email: rebecca.izen@gmail.com

INTRODUCTION

The scale of local adaptation is widely thought to reflect a balance between divergent natural selection and homogenizing gene flow (Lenormand, 2002; Slatkin, 1985). Accordingly, many empirical studies have shown that gene flow constrains, but may not entirely prevent, population divergence and local adaptation (Hendry *et al.*, 2001, 2002; Bolnick and Nosil, 2007; Bolnick *et al.*, 2008; Kaeuffer *et al.*, 2012). However, other studies suggest that phenotypic and genetic divergence may be very strong across remarkably small spatial scales (Skelly, 2004; Antonovics, 2006; Edelaar *et al.*, 2008; Bolnick *et al.*, 2009), where gene flow is expected to be most homogenizing. Examples include abrupt clines and genetic divergence in peppered moths (Saccheri *et al.*, 2008), toxin-tolerant grasses (Antonovics, 2006), *Rhagoletis* apple flies (Filchak *et al.*, 2000), crater-lake cichlid fishes (Elmer *et al.*, 2009), and many other organisms (Ehrlich and Raven, 1969; Selander and Kaufman, 1975).

Such 'microgeographic' variation (Richardson *et al.*, 2014) may be widespread, but often overlooked because we tend not to look for genetic divergence where we do not expect to find it [e.g. within an organism's 'dispersal neighbourhood' (Wright, 1946)]. Thus, tests for microgeographic divergence can help clarify the minimum spatial scales at which adaptation can occur. Such tests may also reveal the evolutionary processes underlying adaptation, because some combination of strong selection, habitat choice, and phenotypic plasticity is required to counteract the blurring effect of dispersal. The adaptationist interpretation is particularly compelling when we observe correlations between environmental factors and organismal traits on fine spatial scales. Such adaptive microgeographic divergence can affect many key biological processes, including speciation, the maintenance of genetic variation, extinction risk (Richardson *et al.*, 2014), and ecosystem services (Schindler *et al.*, 2010).

Empirical studies of microgeographic variation have adopted one of three approaches. The first approach is to test for trait–environment correlations in a metapopulation inhabiting a mosaic of distinct habitats. For instance, Richardson and Urban (2013) showed that salamander populations from a mosaic of ponds (with or without particular predators) were able to diverge despite high gene flow. The second approach is to locate exceptionally steep phenotypic or genetic clines across a sharp environmental gradient (ecotone). For example, grasses spanning an ecotone of soil toxicity along mine tailings exhibit dramatic heritable differences in metal tolerance that are well within the range of gene flow expected from pollen and seed dispersal (Antonovics, 2006). Both these approaches rely on categorizing the environment into discrete habitat types (e.g. predator or toxin presence/absence). A third approach is to test whether phenotypes (or genotypes) co-vary with subtle quantitative measures of the environment that may not follow a simple monotonic cline.

Here, we combine two of these approaches to study phenotypic divergence within and among lake and stream populations of threespine stickleback (*Gasterosteus aculeatus*). By examining divergence across ecotones separating adjoining lake and stream populations of stickleback, we evaluate the magnitude of divergence between spatially distinct populations from ecologically very divergent habitat categories separated by hundreds of metres. Then, by looking within streams, we can test for phenotypic covariance with habitat variables such as flow and depth on the order of metres to tens of metres. By combining these two approaches into one study, we can compare the effect of microhabitat (within a habitat category) against between-habitat divergence. If microgeographic adaptation is possible across discrete habitats at such small spatial scales, might more continuous environmental variation also drive finer microgeographic adaptive divergence within habitats at even smaller scales?

Study system

Parapatric lake and stream populations of threespine stickleback are widely used in studies of local adaptation to discrete habitat types (Reimchen *et al.*, 1985; Hendry *et al.*, 2002; Reimchen and Nosil, 2006; Koskella *et al.*, 2012; Thrall *et al.*, 2012; Feulner *et al.*, 2015; Roesti *et al.*, 2015). Many lake-resident stickleback are limnetic foragers: they have smaller streamlined bodies for sustained swimming, and longer and more numerous gill rakers for feeding on zooplankton (though the extent of limnetic feeding and morphology varies among lakes) (Bell, 1982; Berner *et al.*, 2009; Kaeuffer *et al.*, 2012). In contrast, stickleback in streams have a more classically 'benthic' phenotype, adapted for foraging on large substrate-dwelling macroinvertebrates (Lavin and McPhail, 1986, 1993). Stream fish typically are larger, have deeper bodies thought to aid in manoeuvring, and have fewer and shorter gill rakers. Lake–stream differences in gill rakers and body shape are at least partly heritable, though plasticity does contribute to divergence as well (Oke *et al.*, 2016).

Although phenotypic divergence is observed in nearly all such pairs, gene flow occurs between adjoining lake and stream populations. Mark–recapture studies have found that stickleback are quite mobile: individuals may move up to ~200 m within days to weeks [median of 40 m (Bolnick *et al.*, 2009; Moore and Hendry, 2009)]. Many lake–stream pairs consist of continuous suitable habitat with few if any physical barriers to movement. Stream stickleback are often found just tens to hundreds of metres from the nearest lake. Thus, even relatively distant stream sites (e.g. >1 km away from a lake) are in principle well within a per-generation dispersal radius.

Consistent with this high potential for gene flow, phenotypic divergence is constrained by migration between lake and stream habitats (Hendry and Taylor, 2004). Transects along multiple streams reveal that both genetic and phenotypic divergence increase as one moves from the lake (Berner *et al.*, 2009). These clines are particularly abrupt for inlet streams (often <50 m). Water currents flowing into the lake prevent upstream gene flow by lake fish, which are less inclined to swim upstream against a current (Bolnick *et al.*, 2009; Jiang *et al.*, 2015). In contrast, outlet streams' clines are often quite gradual, spanning hundreds of metres to kilometres (Hendry *et al.*, 2002; Berner *et al.*, 2009), perhaps because lake fish are more readily swept downstream to generate gene flow.

Despite the many papers on lake–stream divergence, little is known about adaptive variation within lakes or within streams. Each habitat type is usually treated as if it were largely homogeneous. In reality, streams (and lakes) are heterogeneous habitats, with substantial variation in depth, substrate, flow regime, and vegetation structure. If biased dispersal (or selection or plasticity) can generate abrupt clines between lake and streams, perhaps stickleback traits might also co-vary with microhabitat traits, over a fine spatial scale within a given habitat. Such covariation between morphology and flow regime has been demonstrated once before, in a single stream (Moore and Hendry, 2005). However, in that study microhabitat measures were confounded with the distance from the nearest lake, making it difficult to distinguish between microhabitat effects and clines arising from gene flow. Consequently, there is still a need for studies of within-stream heterogeneity at finer spatial scales that control for the effect of distance from the adjoining lake.

We therefore tested for microgeographic, within-stream phenotypic divergence in the form of trait–environment correlations (controlling for clinal variation due to proximity to the upstream lake). For comparison, we measured phenotypic divergence between discrete lake and stream habitats. We show that microgeographic environmental variation within the

streams generates as much morphological divergence as the more obvious lake-versus-stream habitat boundary.

METHODS

Study site

We examined a transect of alternating lake and stream habitats in the west fork of the Amor de Cosmos watershed on northern Vancouver Island, British Columbia. The large oligotrophic Amor Lake (362 ha) is the site of highest elevation. The outlet stream that drains from Amor Lake is shallow and rocky, with appreciable current (~8 $cm \cdot s^{-1}$) as it drops ~25 m over its 1.2 km length (Fig. 1). Much of the upper reach of Amor outlet stream has fast turbulent current and rocky substrate, and is uninhabited by stickleback. The lower 350 m of the stream is slower, containing slow-moving water and muddy substrates interspersed with faster-moving currents through narrow channels with sandy substrates.

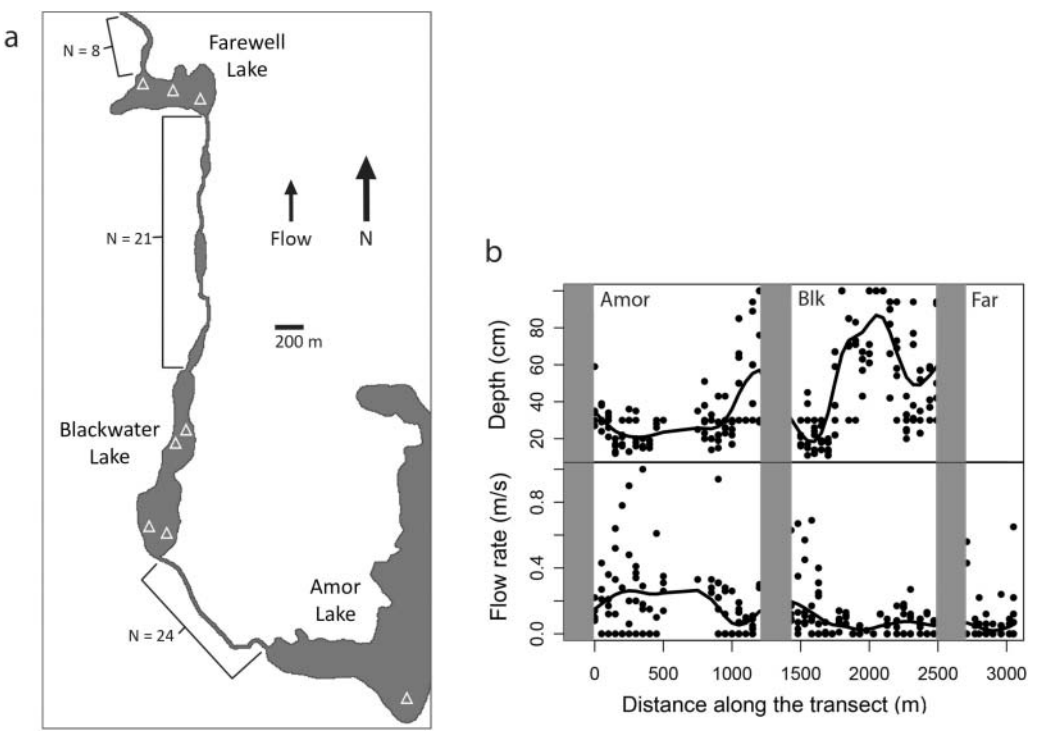

Fig. 1. (a) Map of the research site. Stickleback were collected along a transect of alternating lake and stream habitats. Flow proceeds north from Amor Lake into Amor stream, then Blackwater Lake, which drains into Blackwater stream, Farewell Lake, and then into Farewell stream. Triangles indicate lake sample sites. Streams were sampled at ~50 m intervals, and brackets indicate the number of sample sites in each stream. There were multiple traps at each site. (b) Trap flow rates ($m \cdot s^{-1}$) and depths (cm) along the transect, with a loess spline fit for illustration. The grey rectangles represent lakes (the width is not to scale), and the points indicate stream sites. Depth data were not obtained from Farewell outlet stream.

The final high-current channel empties into the long, thin Blackwater Lake (37.2 ha). The start of Blackwater's outlet stream is clearly delineated by a log jam, at which current begins. The Blackwater outlet stream is 1.2 km long, consisting of alternating areas of rocky high-flow stream-bed, moderate-flow sandy channels, pools, and still-water marsh. Near the midpoint of the stream there is a large 200 m long and >1 m deep pool, after which there is a series of pool-riffle microhabitats before the stream enters Farewell Lake (20.7 ha). Farewell's outlet stream (>10 km) also alternates pools and riffles and includes some small cascades. In summary, all three streams have: (1) shallow high-flow regions at their upper and lower interfaces with adjoining lakes, (2) heterogeneous flow rates along their entire length, with alternating pools and riffles, and (3) dense stickleback populations along their entire length (except the upper two-thirds of the stream between Amor de Cosmos and Blackwater Lakes).

Gene flow is a reasonable expectation between almost all sites in this system. The only barriers in the lower two streams are areas of moderate water flow and temporary beaver dams. Consequently, genetic divergence was very weak between Blackwater and Farewell Lakes [$F_{ST} = 0.03$ based on 6 microsatellites (Caldera and Bolnick, 2008)] despite 1.2 km of alternating stream habitat between them, with numerous fast-flowing riffles. Blackwater Lake and its inlet stream also exhibited minimal genetic divergence [$F_{ST} = 0.008$ (Bolnick *et al.*, 2009)]. The only insurmountable barrier to stickleback movement is the steep section of stream between Amor and Blackwater Lakes: microsatellites revealed significant genetic divergence ($F_{ST} = 0.13$) between Amor Lake stickleback and both downstream lake populations (Caldera and Bolnick, 2008).

Field sampling

In July 2010, we collected stickleback along a continuous transect from Amor Lake to the Farewell outlet stream (Fig. 1, Table 1). A modest sample ($n = 20$) of Amor Lake fish was collected on the south shore, 2.4 km from the outlet stream. Blackwater Lake fish ($n = 133$) were collected at two sites near the inlet and at two sites near the outlet. Farewell Lake fish

Table 1. Geographic location, sampling sites, number of traps, and sample size for the three lake and three stream populations

Population	Sites	Latitude	Longitude	Traps	Sample size
Amor Lake	Near outlet	50°9′19″N	125°31′60″W	20	20
Amor stream	$n = 24$	50°9′40″N	125°34′60″W	93	127
Blackwater Lake	Near inlet	50°10′1″N	125°35′23″W	20	41
	Near inlet	50°10′1″N	125°35′15″W	20	41
	Near outlet	50°10′52″N	125°35′26″W	20	22
	Near outlet	50°10′52″N	125°35′11″W	20	29
Blackwater stream	$n = 21$	50°11′37″N	125°34′60″W	100	328
Farewell Lake	Near inlet	50°11′54″N	125°35′11″W	20	46
	Centre	50°12′6″N	125°35′16″W	20	49
	Near outlet	50°12′12″N	125°35′36″W	20	50
Farewell stream	$n = 8$	50°12′22″N	125°34′60″W	28	104

Note: GPS coordinates for streams are from the stream centre.

($n = 145$) were sampled at three sites: near the inlet, halfway along the shore, and near the outlet. In the streams, we set unbaited minnow traps at 50 m intervals down the entire lengths of Amor ($n = 127$) and Blackwater ($n = 328$) outlet streams, and down the first kilometre of the Farewell outlet stream ($n = 104$). There were multiple traps at each 50 m stream site. We recorded which individual fish came from each trap, and catch rates per trap. In Amor and Blackwater streams, we also collected microhabitat data for each trap: depth (cm) and water velocity ($cm \cdot s^{-1}$, measured with a Flowatch flowmeter) immediately adjacent to each trap. Flow but not depth data were collected for Farewell outlet. Captured fish were euthanized in buffered MS-222 and preserved in 10% neutral buffered formalin.

In the following descriptions, we use 'habitat' to refer generically to all lake versus all stream sites (e.g. a main effect). We use 'pair' to refer to a lake and its outlet (Amor, Blackwater, Farewell), because gene flow from a lake up-current into an inlet tends to be weak. We use 'population' to refer to a particular lake or stream (e.g. Amor Lake, Amor stream, Blackwater Lake, Blackwater stream, Farewell Lake, Farewell stream). We use 'site' to refer to a particular trapping location within a population (multiple traps per site). Lastly, we use 'trap' to refer to the particular location where specific fish were caught from a single trap.

Morphology measurements

We measured standard length (mm) of each fish using digital calipers. Using a dissecting microscope, we counted the gill rakers on the first branchial arch. We photographed the three longest gill rakers at 7.5× magnification, and used imageJ (imagej.nih.gov/ij/) to trace the lengths of those rakers, from which we calculated mean gill raker length. A single individual (R.I.) performed all measurements. We performed body size corrections on gill raker length by regressing ln-transformed raker length on ln-transformed standard length and keeping the residuals. We used a single regression line for the size-correction for all populations (lakes and streams); there was no population × size interaction term. Sex was recorded for individuals with mature gonads, but was not scored for reproductively immature fish. Phenotypic clines in some lake–stream pairs are sex-specific (Moore and Hendry, 2005). For sake of brevity, an analysis of sex effects is beyond the scope of this paper.

We used geometric morphometrics to measure divergence in body shape. We digitized 20 homologous landmarks from the head (anterior extent of maxilla, anterior, dorsal, and posterior extent of orbit, supraoccipital notch, first dorsal spine insertion, insertion point of pelvic spine into the pelvic girdle, junction of gill with ventral surface, quadrate-articular jaw joint; Fig. 2) and the remaining body (second and third dorsal spine insertions, posterior insertion of the dorsal fin, dorsal and ventral insertions of the caudal fin, posterior extent of the caudal peduncle, posterior and anterior insertions of the anal fin, anterior tip of the pelvic girdle, dorsal and ventral insertions of the pectoral fin) with tpsDig2. We performed a generalized Procrustes analysis on the landmarks using the *geomorph* package in R (Adams and Otárola-Castillo, 2013), to obtain aligned x and y coordinates.

Our analysis focused on head shape landmarks, for several reasons. First, head shape is associated with variation in feeding mechanics such as the propensity for ram-feeding or suction-feeding (Willacker *et al.*, 2010). Given that lake and stream fish typically consume different types of prey, we had a strong hypothesis that head shape would be important for adaptive divergence between lake and stream fish, and among fish in different stream microhabitats. Second, head shape also incorporates a measurement of body depth

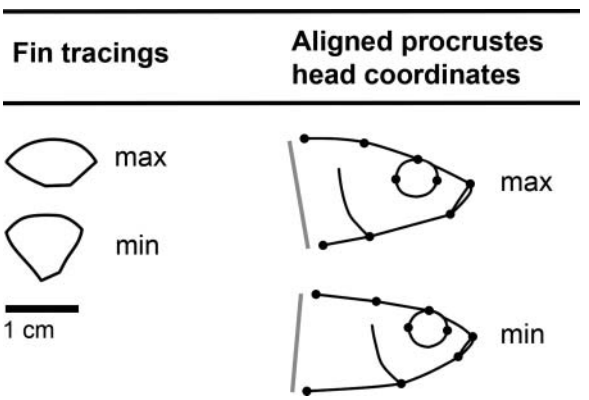

Fig. 2. (Left) Variation in fin shape captured by the discriminant function of fin measurements as a function of habitat (lake/stream). The illustrated deformations represent tracings of fins with the minimum (lake-like) and the maximum (stream-like) discriminant axis scores. The scale bar corresponds to fins only. (Right) Variation in head shape captured by the first principal component of variation in head landmarks after generalized Procrustes analysis. The illustrated deformations are plots of aligned Procrustes coordinates, averaged for the four lowest-scoring and the four highest-scoring heads. Head shape is scale-free.

(insertion of the first dorsal spine to the insertion of the pelvic spine into the pelvic girdle), a known component of manoeuvring and sustained swimming that differs between lake and stream habitats (Walker, 1997; Hendry *et al.*, 2011). Lastly, unlike body shape, head shape was unaffected by either bending of the vertebral column during formalin preservation, or by reproductive state in females. We retained the first principal component of variation among the nine aligned head coordinates as a proxy for overall variation in head shape.

Stickleback are pectoral swimmers, so pectoral fin shape affects sticklebacks' ability to manoeuvre and maintain their position in flowing water (Walker, 2004), and their propensity to swim up- or down-current (Jiang *et al.*, submitted). We therefore posited that fin shape might differ between lake and stream habitats, and between stream microhabitats. To measure pectoral fin morphological divergence, we removed specimens' right pectoral fins and fanned them out alongside the left side of the body in the same photographs taken for geometric morphometrics. We omitted 205 fins from the analysis that showed previous damage, tore, or did not fan out properly in the photograph. We used imageJ to measure surface area, lengths of the longer edge ray, the shorter edge ray, width across all of the rays at the base of the fin, and distance between tips of the rays at the distal edge of the fin. We used the fin characters to generate a discriminant function axis of pectoral fin morphology based on lake or stream habitat. The axis was dominated by aspect ratio (fin length/width), a key factor in pectoral fin rowing efficiency [LD1 loadings: base width = 23, tip width = 18, long edge ray = 3.5, short edge ray = −6.2, surface area = −31, longest ray = −2.2 (Walker and Westneat, 2002)]. Fins with a higher aspect ratio create less drag, so they and are better adapted for sustained swimming and power strokes during limnetic foraging (Walker and Westneat, 2002; Walker, 2004). Lower aspect ratio fins generate more drag, so they are well suited for fine-scale manoeuvring and station-keeping during benthic foraging (Walker, 2004). We retained individuals' scores along this lake/stream discriminant function axis as a proxy for overall variation in fin shape.

We wish to emphasize that we are measuring phenotypic divergence from wild-caught fish that may include both heritable and plastic sources of variation. Several studies have experimentally quantified lake–stream divergence in common-garden stickleback. Their results demonstrate that many traits have both heritable and plastic components, though the magnitude of these effects varies among phenotypes (Lucek *et al.*, 2014), and among lake–stream pairs (Oke *et al.,* 2016). Gill raker length, for instance, is widely viewed as highly plastic.

Between-habitat analyses

To test for morphological divergence between lake and stream habitats, we used a multivariate analysis of variance (MANOVA) of morphology (gill raker length, gill raker number, head shape, fin shape, and standard length) against habitat type (lake vs. stream) with pair as a factor. Subsequently, we used analysis of variance (ANOVA) on each trait separately to identify individual traits contributing to lake versus stream differences. In these trait × trait analyses, we included pair as a factor, with a pair × habitat interaction effect. Tukey HSD tests identified particular population comparisons underlying significant effects in the MANOVA or ANOVA models.

Because fish were sampled from one to four sites per lake and from eight to 24 sites per stream, it is possible that fish from a given sample site are biologically non-independent pseudo-replicates. To evaluate this concern, we tested whether there is significant among-site variation within each lake, and within each stream, separately. Next, we ran a mixed effect linear model (lme4 package in R) with habitat and lake–stream pair (and their interaction) as fixed effects, and sample site as a random effect. The mixed-model approach largely recapitulated the results of the linear model (without site random effects), but had little power to detect interaction effects (within which site effects were non- or weakly significant) because we had only one sample site within Amor Lake, and only three and four sites in Blackwater and Farewell Lakes, respectively. Thus, the mixed-model approach confers little advantage (accounting for weak spatial structure within lakes), at a substantial cost of power, so we focus on fixed-effect models in our results.

Within-habitat analyses

To investigate even finer-scale, within-habitat divergence, we tested whether morphology was correlated with microhabitat variation within streams. A multivariate analysis of covariance (MANCOVA) tested whether morphological traits (gill raker number and length, head shape, fin shape, and standard length) were associated with flow rate ($cm \cdot s^{-1}$), depth (cm), the interaction of flow rate and depth, distance from the upstream lake (m), and lake–stream pair. With the exception of 'pair', all predictor variables were continuous covariates. We included distance from the upstream lake to account for possible clines arising from immigration from the lake into its outlet stream. Because upstream migration tends to be weak, we did not account for distance to the downstream lake (Bolnick *et al.*, 2009; Jiang *et al.*, 2015). Distance and microhabitat are not confounded in this study (regression effect of distance on water flow: $r = -0.085$, $P = 0.096$, d.f. = 379; effect on depth: $r = 0.026$, $P = 0.614$, d.f. = 379). To test microhabitat effects for univariate traits, we performed ANCOVAs in which each morphological trait was regressed on the first principal component axis for microhabitat

measures (64% of variance, loadings: flow rate = −0.71, depth = 0.71), with a main effect of population (Amor or Blackwater), and a population × PCA interaction.

We then determined the relative effects of habitat versus microhabitat by comparing effect size estimates from the MANOVA and MANCOVA and models mentioned above. Effect size was calculated as partial variance explained (η^2 = 1 − Wilks' lambda = SS effect / [SS effect + SS error]). Although the two models had different underlying data sets (stream-only vs. lake and stream), total multivariate variances were comparable between the two data sets (stream-only = 62.1; lake and stream = 57.5). This similarity enabled our comparison of the effects of within- and between-habitat environmental variation on morphology. We also compared the habitat versus microhabitat effect sizes for univariate traits, using ANOVAs testing habitat and pair effects on each trait (all fish), and ANCOVAs testing microhabitat and population effects on each trait (stream fish). Effect sizes were measured as the percent variance explained by the ANOVA or ANCOVA, relative to the total variance of the focal trait. For fin shape, the previously described discriminant function axis between lake and stream habitats would have over-emphasized the effect of between-habitat environmental variation. To prevent this bias, we substituted the first principal component of variation in fin shape (88% of variance, loadings: base width = 0.38, tip width = 0.41, long edge ray = 0.41, short edge ray = 0.40, surface area = 0.43, longest ray = 0.42) in models used for the effect size comparisons.

Due to specimen warping and incidental damage during dissection, some morphometric variables were missing for some individuals. However, data were available for nearly all individuals (<2% missing) for linear measures such as gill raker measurements and body size. Head shape measurements were missing in 15% of the data set, and 20% of fin shape measurements were missing. Thus, we imputed missing values to permit multivariate analyses using the Amelia II package in R (Honaker *et al.*, 2011). Expected maximization algorithm chains converged after five Amelia runs. The five replicate imputations produced quantitatively similar statistical inferences (i.e. MANOVA and MANCOVA results). We present the imputed MANOVA with the median effect size of habitat, and the imputed MANCOVA with the median effect size of microhabitat (sum of flow rate, depth, and flow × depth interaction). For all univariate trait × trait analyses, we used the raw data with missing values removed. Univariate ANOVAs and regressions using imputed values yielded the same conclusions as analyses that omitted imputed values. This suggests that the use of imputed values is unlikely to skew results of the multivariate analyses (MANCOVAs). Residuals were normal (as evaluated by qqnorm in R), both before and after imputation.

RESULTS

Between-habitat divergence

Morphology differed significantly between lake and stream habitats (habitat: $\eta^2 = 0.21$, $F = 47$, d.f. = 858, $P < 0.001$; lake–stream pair: $\eta^2 = 0.14$, $F = 28$, d.f. = 858, $P < 0.001$). Univariate analyses revealed that some morphological trait differences between habitats matched those of other lake–stream pairs from previous studies, while others traits did not. Stream stickleback had marginally fewer gill rakers compared with lake fish (Table 2). Stream individuals' fins were shorter and wider (typically considered an adaptation for better manoeuvring), whereas lake sticklebacks' fins had a higher aspect ratio (associated with sustained swimming and power strokes; Table 2, Fig. 2). Atypically, the stream fish

Table 2. Morphological divergence between lake and stream samples

	Stream		Lake		Stream vs. Lake			
Trait	*n*	mean	*n*	mean	*t*	*P*	d.f.	R^2
Head PC1	461	−0.004	278	0.005	−3.83	<0.001	737	0.018
Pectoral fin DFA score	358	0.252	176	−0.513	8.32	<0.001	532	0.113
Gill raker length (cm)*	557	0.028	292	−0.054	7.65	<0.001	847	0.063
Gill raker number	559	20.29	289	20.47	−1.75	0.08	846	0.002
Standard length (cm)	559	4.62	293	4.78	−3.51	<0.001	850	0.013

Note: Means and sample sizes of the traits are provided from one-way ANOVAs with habitat category as a factor. *t*-values, *P*-values, degrees of freedom, and adjusted R^2 are provided.
*Gill raker lengths were allometrically standardized by retaining residuals from a linear regression on standard length, prior to analysis.

were shorter, with longer gill rakers and shallower, less elongated heads relative to lake fish (Table 2). Similar inferences were obtained using mixed effect linear models testing for univariate trait dependence on habitat and lake–stream pair, with sample site as a random effect.

Post-hoc Tukey HSD tests clarified the causes of these significant model effects, by identifying which population combinations were divergent for a given trait. At the beginning of the transect, the stickleback in Amor Lake had traits characteristic of limnetic populations (Fig. 3). The transition to a benthic head and shorter gill rakers in the Amor outlet stream matched canonical lake–stream divergence (Fig. 3). At the next lake–stream transition, Blackwater, stickleback did not regain their limnetic character in the lake. Instead, Blackwater Lake stickleback were even more benthic than their inlet stream neighbours (Fig. 3). Moving further down the watershed, Blackwater outlet stream fish were significantly different from both their up- and downstream lake neighbours (Blackwater and Farewell Lakes). Surprisingly, the stream fish were morphologically relatively limnetic compared with the benthic-like lake fish (Fig. 3). The limnetic phenotype in Amor Lake contributed little to the overall lake versus stream comparison, because comparatively few fish were captured from Amor ($n = 20$, Table 1), whose fish tend to be very trap-shy (D.I. Bolnick, personal observation).

Within-stream variation

We next examined morphological variation at a fine spatial scale along individual stream transects (Fig. 4). As distance downstream increased, we did not observe the expected monotonic clinal transition from lake- to stream-like phenotypes (Fig. 4). Rather, we found irregular phenotypic change with distance, with distinct patterns for different traits in different streams. For example, consider the change in gill raker length when moving from Blackwater Lake into its outlet, then to Farewell Lake (Fig. 4). Blackwater Lake contains short-raker (benthic) fish, which quickly transition to longer rakers as one first enters the stream. Gill rakers then drop down again towards the centre of the stream, before rising again to a limnetic form just before one enters Farewell Lake (which has benthic short-raker fish). There is thus a U-shaped cline in that stream, where the middle (farthest from either lake) most closely resembles the two lakes' forms. The reason for this U-shape appears to

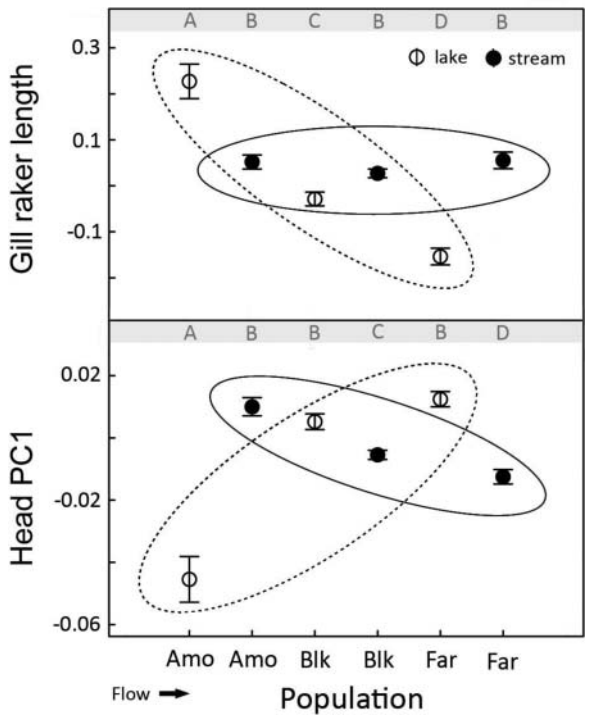

Fig. 3. Divergence in size-adjusted gill raker length (cm) and head shape among the three lake and three stream populations. Trait means and standard errors are displayed for the populations. Significance letters, shown in grey above the plots, were assigned using Tukey HSD tests with a 99% confidence level. The ellipses are drawn to emphasize variation within habitat categories.

be a large area of wider, shallower-flowing stream habitat near the middle of the stream (Fig. 1B).

Fish morphology was correlated with local stream hydrology (flow rate and depth at the trap where each fish was caught; Table 3), controlling for distance from the upstream lake. Morphology of stream fish did co-vary with distance from the upstream lake (MANOVA effect size: $\eta^2 = 0.20$). However, microhabitat (flow rate, depth, and flow × depth effects) collectively had a similar effect size ($\eta^2 = 0.18$, Table 3). Note that effect size is calculated as partial variance explained ($\eta^2 = 0.1$ – Wilks' lambda), which does not sum to 100%. Distance from the upstream lake was not a significant predictor of flow rate or depth (see Methods), supporting microhabitat and distance as stand-alone covariates in the within-stream MANCOVA model.

Microhabitat effects in the stream-only data set had similar total variance to the larger data set of the lake versus stream comparison. The total variance for the whole data set (sum of eigenvalues) was 57.7, compared with 61.8 for the stream-only data (Table 3). Consequently, we can approximately compare the effect sizes for microhabitat versus macrohabitat. The effect size of microhabitat ($\eta^2 = 0.18$) was stronger than the effect of discrete habitat category (lake vs. stream, $\eta^2 = 0.10$, Table 3).

The results are more complicated when we compare the habitat and microhabitat effects on univariate traits (Table 4). Habitat (lake vs. stream) had significant effects on head shape,

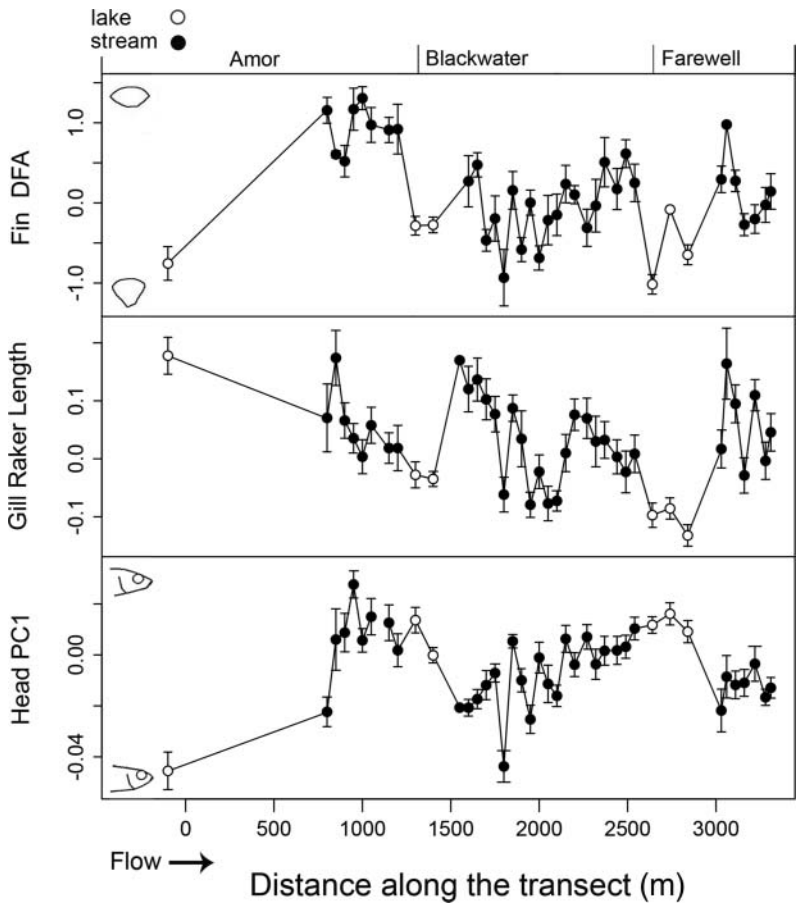

Fig. 4. Trait divergence patterns along the study transect from Amor Lake to the Farewell outlet stream. Means and standard errors for the morphological traits are derived from multiple traps at each sample site. The fin DFA (discriminant function axis) was calculated as a discriminant function between lake and stream habitats (see Fig. 2 for an illustration).

fin shape, gill raker length, and standard length (all $P < 0.001$) but not gill raker number. Within-stream microhabitat (flow and depth PC1) had significant effects on fin shape ($P < 0.05$), gill raker length ($P < 0.001$), and standard length ($P < 0.001$), but neither head shape nor raker number. In the Amor outlet, fins became increasingly benthic-like in slower and deeper sections of the stream (Fig. 5), in line with results from the lake–stream comparison. Gill raker length was unaffected by flow rate and depth in the Amor outlet (Fig. 5). However, in the Blackwater outlet, gill raker length decreased in slower and deeper sections of the stream (becoming more benthic in more benthic habitat), while fin shape was unaffected by flow rate and depth (Fig. 5).

Comparing the univariate effect sizes (Table 4), the relative effect of habitat versus microhabitat differed among traits. Habitat explained three times more variance in head shape than did microhabitat. In contrast, habitat and microhabitat explained fairly similar amounts of variance in gill raker length (only 27% higher effect size of habitat). Two traits

Table 3. Comparison of lake/stream and within-stream multivariate models

Model	η^2	F	d.f.	P
MANCOVA Microhabitat (total variance 61.8%)				
Flow rate ($cm \cdot s^{-1}$)	0.033	2.95	427	0.012
Depth (m)	0.11	10.9	427	<0.001
Distance (m)	0.20	21.6	427	<0.001
Population	0.021	1.82	427	0.11
Flow × Depth	0.034	2.98	427	0.012
MANOVA Lake vs. Stream (total variance 57.7%)				
Habitat	0.10	19.5	858	<0.001
Population	0.084	15.8	858	<0.001

Note: The within-stream MANCOVA is based on morphological traits in the stream-only data set, with microhabitat traits (flow rate, depth, flow rate × depth) and distance from the upstream lake as covariates and lake–stream pair (Amor, Blackwater, Farewell) as a factor. The lake/stream MANOVA is based on the same morphological traits for lake and stream fish combined, with habitat (lake or stream) and lake–stream pair as factors. Total data set variance (sum of eigenvalues), effect size (partial variance explained), approximate F-values, degrees of freedom, and P-values are provided. Partial variance explained is η^2 = 1 − Wilks' lambda = SS effect / [SS effect + SS error].

Table 4. Comparison of effect sizes of habitat (lake vs. stream) versus microhabitat (within-stream flow and depth) for univariate phenotypic traits

	Lake and stream data set (% variance explained)			Stream-only data set (% variance explained)		
	Variance	Habitat	Habitat × Pair	Variance	Microhabitat	Microhabitat × Population
Head PC1	0.00085	2 ***	13.4 ***	0.00069	0.65	1
Fin PC1	5.26	0.53	5.8 ***	5.46	4.2 ***	0.45
Gill raker length (cm)	0.024	6.5 ***	5.5 ***	0.022	5.1 ***	1.7 *
Gill raker number	2.07	0.34	0.59	2.05	0.66	0.052
Standard length (mm)	49.8	1.4 ***	5 ***	51.4	4.2 ***	0.41

Note: Effect sizes are measured by the percent of explained variance (SS effect/SS total). The habitat effect is estimated using ANOVA testing each trait for effects of habitat, lake–stream pair (Amor, Blackwater, Farewell), and a habitat × pair interaction, using all sampled fish. The microhabitat effect is estimated using ANCOVA on stream fish alone, with a microhabitat covariate (first PCA axis of flow and depth), population (Amor or Blackwater streams, because depth data were lacking in Farewell), and a microhabitat × population interaction. For scale, we report the trait variance for each data set. *P < 0.05, ***P < 0.001.

exhibited stronger effects of microhabitat, which explained three times more variance in standard length and eight times more variance in pectoral fin shape (PC1). Gill raker number had no significant effect of either habitat or microhabitat. Thus, the MANOVA finding of comparable habitat and microhabitat effects may be the result of combining

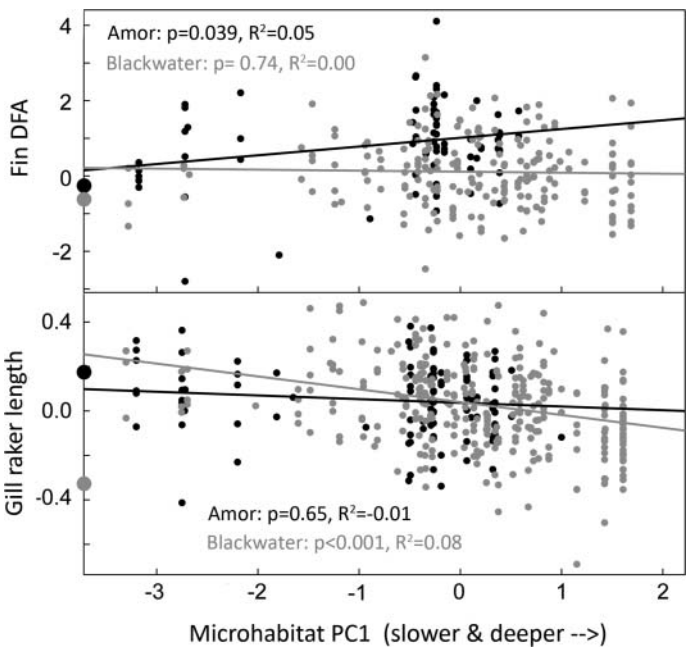

Fig. 5. Plots of fin shape and size-adjusted gill raker length (cm) against the first microhabitat principal component in the Amor (black) and Blackwater (grey) outlet streams. For reference, lake means are plotted on the y axis. *P*-values, r^2, and best-fit lines are provided from linear models of morphology against stream microhabitat (flow and depth). Farewell Lake and outlet stream are omitted because we lack depth measures from the stream.

multiple phenotypes, some of which co-vary more strongly with habitat, and with microhabitat. A caveat is that the univariate microhabitat effects may be an underestimate, because we used the first principal component axis of flow rate and depth, omitting 33% of the microhabitat variation.

DISCUSSION

Many studies of phenotypic variation and local adaptation treat a landscape as a mosaic of discrete habitats. This approach, however, can obscure fine-scale variation within those habitats. Our results present an example of microgeographic variation that is nested within such discrete habitat divergence. We find that microhabitat variation within stream populations explains more phenotypic variance (overall, and for certain univariate traits) than does a main effect of lake versus stream habitat. The implication is that phenotypic divergence can occur at finer spatial scales, and in response to more subtle environmental variation, than we often expect based on the homogenizing potential of gene flow. Consequently, researchers may generally be well advised to complement between-habitat analyses (where divergence is expected) with intra-habitat tests of trait–environment correlations.

Lake–stream divergence

We examined stickleback morphology along a transect of alternating lake and stream habitats in a single watershed, while also considering microhabitat variation within the stream habitat. At a lake–stream level, we found significant morphological differences between stickleback inhabiting these alternate habitat categories, as expected from prior research (Lavin and McPhail, 1993; Hendry *et al.*, 2002; Kaeuffer *et al.*, 2012; Jiang *et al.,* 2015). Some traits (gill raker number and fin shape) matched our typical expectation to find more limnetic- and benthic-like fish in the lake and stream, respectively. Other traits did not match this expectation, because two of the three lakes are relatively small mesotrophic basins with relatively benthic-like fish. Previous studies of lake and stream stickleback have also noted that some lake–stream pairs exhibit atypical divergence for certain traits, such that parallel evolution is not wholly consistent in this system (Berner *et al.*, 2009; Kaeuffer *et al.*, 2012). It appears that lake populations are responsible for much of the idiosyncrasy in lake–stream divergence: all three stream populations were on average morphologically fairly similar, whereas Amor Lake was highly divergent from the other two lakes both morphologically as well as genetically (Caldera and Bolnick, 2008). This may reflect either separate colonization of the Amor Lake fish, or adaptation to their much larger (more limnetic) lake.

Given this significant habitat effect, it is clear that phenotypic divergence can proceed despite some gene flow between the populations. This is consistent with prior studies that found phenotypic divergence between lake and outlet stream stickleback over a scale of hundreds of metres to kilometres. And yet, this result is also somewhat surprising: almost all of our stream sites are within about 500 m of a lake, about 2.5 times the distance that stickleback have been observed to move in just a few weeks (Moore and Hendry, 2009). There are also no physical barriers to fish movement up- or downstream. We therefore anticipate a high rate of gene flow, and indeed F_{ST} is very low between some of the populations studied here. Most notably, Farewell and Blackwater Lakes exhibit no significant genetic divergence (Caldera and Bolnick, 2008), and little morphological divergence (this study). Nonetheless, both lakes contained fish that were morphologically significantly different from their intervening stream population.

Although some of the observed divergence may be environmentally induced, most phenotypes and most lake–stream pairs exhibit some heritable differences (Lucek *et al.*, 2014; Oke *et al.*, 2016). Of the populations studied here, we previously found heritable differences in neuromast number, body size, and immune function between stickleback from Blackwater Lake and its inlet stream (Bolnick *et al.*, 2015; Jiang *et al.*, submitted), which exhibit no significant F_{ST} for five of six microsatellites examined (Bolnick *et al.*, 2009). Nonetheless, we cannot at present confidently determine whether the divergence we document is genetic or the result of adaptive (or perhaps mal-adaptive) phenotypic plasticity.

Within-stream divergence

Observing lake–stream divergence despite gene flow, we wondered whether stickleback also show microhabitat-based divergence at finer spatial scales. We found significant morphological variation among capture sites within a given stream. Most remarkably, this within-stream variation did not adhere to a simple clinal transition that would be expected from migration–selection balance across an ecotone boundary (e.g. from lake-like traits near the lake–stream interface to increasingly stream-like traits with distance from the lake)

(Hendry and Taylor, 2004; Hendry *et al.*, 2013). Instead, we found irregular, non-monotonic variation in morphology along each stream (Fig. 4). For instance, in the middle stream (between Blackwater and Farewell), the most lake-like phenotypes were found in the middle of the stream and the least lake-like phenotypes were found closest to the upper and lower lakes. Such deviations from clinal trends are attributable in part to microhabitat variation in flow rate and depth. Depending on which stream we examined, either gill raker length or fin shape co-varied with the first principal component axis of flow and depth. Importantly, because microhabitat is not correlated with distance from the upstream lake [unlike a previous analysis from another stickleback lake–stream pair (Moore and Hendry, 2005)], the microhabitat effect is not merely a by-product of distance from the nearest lake (which itself has a comparable and significant effect).

Although depth and flow co-vary, when we considered them separately we found that depth generally has a larger effect on morphology than does flow rate; however, depth and flow interact. For example, at low depths, flow rate had no appreciable effect on fin shape. In contrast, at deeper traps, flow rate was positively associated with stream-like fin shape. The cause of this interaction remains unclear. One possibility is methodical: turbulence and boundary effects may generate more variable flow rates in shallow water, leading to noisier flow meter measurements. Or, we speculate that shallow sites (subject to water level fluctuations) contain only transient visitors, whereas deeper sites represent more important permanent habitats to which individuals adapt via selection, plasticity, or habitat choice.

The trait–depth correlations that we document are just one example of an under-appreciated but perhaps quite general phenomenon of within-habitat phenotypic differentiation. A prior study found that individual stickleback with more exploratory behaviour were more likely to be found in open, rather than sheltered, areas of a river (Pearish *et al.*, 2013). In breeding salmon, deeper-bodied fish tend to be found deeper within a given pool (Hendry *et al.*, 2001). In trout, there is heritable variation in foraging personality between pools and riffles within a given stream (Bridcut and Giller, 1995), and many intertidal organisms (e.g. *Littorina* snails) exhibit dramatic microgeographic phenotypic variation across elevational gradients of a few metres (Reimchen, 1981; Sundberg, 2008). These and many other examples (Richardson *et al.*, 2014) highlight the potentially very general nature of microhabitats leading to stable microgeographic phenotypic variation. The large theoretical literature on migration–selection balance (Slatkin, 1985) leads us to not expect, and not look for, such fine-scale divergence in mobile organisms. Therefore, microgeographic variation such as we document here may be widely overlooked.

Comparing habitat to microhabitat effects

Few if any prior studies in stickleback or in other organisms have explicitly contrasted the effect of habitat category against the effect of microhabitat variation within a category. In the present study, the multivariate effect size of microhabitat within stream ($\eta^2 = 0.18$) was stronger than the effect of crossing the lake–stream boundary ($\eta^2 = 0.10$). This comparison of microhabitat to habitat effect sizes is approximate because they come from different models with the microhabitat analysis using a subset of the total data. However, the full and subset data have very comparable multivariate variances, implying that differences in η^2 are also comparable.

Considering univariate traits separately, habitat and microhabitat effect sizes were similar for gill raker length. Fish size (standard length) and fin shape varied more with

microhabitat, whereas head shape varied more across lake/stream comparisons. The habitat effect sizes in our study are comparable to the lake–stream habitat effects from multiple lake–stream pairs in a prior study (Berner *et al.*, 2009). For instance, our lake–stream effect on body size is equivalent to the average habitat effect from eight other lake–stream pairs. In contrast, gill raker length exhibited relatively weak divergence in our study.

Mechanisms of habitat and microhabitat divergence

Phenotypic divergence between parapatric habitats is usually interpreted as an adaptive response to divergent natural selection in the face of gene flow. However, this is unlikely to be a sufficient explanation for our microgeographic variation within streams. This is because stream stickleback readily travel up to 150 m in a few days (Bolnick *et al.*, 2009) or up to 200 m in a few weeks (Moore and Hendry, 2009). That is, per-generation dispersal neighbourhoods will be much larger than the spatial scale of divergence documented here. Assuming that gene flow is likely high at this fine scale, selection would need to be exceptionally strong to generate the observed phenotypic divergence (Bolnick and Otto, 2013). Such strong selection is inconsistent with the modest selection strengths found in most lake–stream experimental transplant studies of stickleback (Hendry *et al.*, 2002; Rolshausen *et al.*, 2015; but see Moser *et al.*, 2016), suggesting that other mechanisms of divergence are at play in our system.

Phenotypic plasticity can cause microspatial variation if individuals first disperse and then adjust their phenotype to suit their new habitat. Phenotypic plasticity has been described for many traits, in many different populations of stickleback, including lake and stream stickleback (Day *et al.*, 1994; Day and McPhail, 1996; Wund *et al.,* 2008; Lucek *et al.*, 2014; Foster *et al.*, 2015; Mazzarella *et al.,* 2015; Oke *et al.*, 2016). However, no study has yet tested for heritable phenotypic differences among sites within a stream (or within a lake), or plastic responses to such microspatial variation. If plasticity is important, it would have to act impressively quickly relative to the speed at which individuals diffuse across the landscape and among microhabitats. Based on prior dispersal estimates, stream stickleback typically move ~40 m within a few days (Moore and Hendry, 2005; Bolnick *et al.*, 2009). In our stream survey, such dispersal kernels would mean that most fish encounter the full range of available microhabitats within a week of movement.

Non-random gene flow is another possible explanation for microgeographic divergence. Gene flow can actually play a positive role in local adaptation when it facilitates spatial sorting of genotypes into their respective preferred habitats (Armsworth and Roughgarden, 2005; Edelaar and Bolnick, 2012; Bolnick and Otto, 2013; Berner and Thibert-Plante, 2015). Lake and stream stickleback do vary in their dispersal ability and habitat preferences (Bolnick *et al.*, 2009; Jiang *et al.*, 2015), facilitating non-random sorting of phenotypes and alleles among habitats. For instance, fin shape affects swimming mechanics in flowing water (Walker and Westneat, 2002) and lateral line sensory systems affect sticklebacks' ability to detect water flow (Wark and Peichel, 2009; Jiang *et al.*, submitted). Both traits differ between Blackwater Lake and stream fish, contributing to divergent swimming behaviour in flowing water (Jiang *et al.*, 2015). This variation in swimming behaviour may, in turn, facilitate adaptive divergence between habitats, or among microhabitats within the stream.

The causes and implications of this microgeographic variation for adaptation are intriguing but remain poorly understood. First and foremost, we do not yet know whether this microgeographic variation is heritable. Common-garden rearing experiments using families bred from within and among different locations within a stream would clarify

this point. If the divergence is heritable, we would have evidence for evolutionary divergence at a remarkably fine scale within stream fish. We are intrigued by the possibility that spatial segregation of individuals across microhabitats can alter a variety of species interactions (diet, predation risk, infection status) in ways that can generate spatial variation in eco-evolutionary feedbacks.

CONCLUSIONS

We found both coarse-scale (lake–stream) and fine-scale (within-stream) phenotypic variation along an alternating transect of habitats. Importantly, the fine-scale variation was stronger than coarse-scale trait divergence. These results provide a case study illustrating within-habitat microgeographic variation (Richardson *et al.*, 2014). A corollary of this result is that coarse categorization of habitat (e.g. lake vs. stream) can lead us to overlook more subtle quantitative variation within habitats.

ACKNOWLEDGEMENTS

Special thanks go to Julia Saltz and Scott Egan for discussions and comments on the manuscript. W.E. Stutz, K. Ballare, and R. Carlson assisted with field collections. This work was supported by the David and Lucille Packard Foundation, the Howard Hughes Medical Institute, and NSF grant DEB-1144773 to D.I.B.

REFERENCES

Adams, D.C. and Otárola-Castillo. E. 2013. Geomorph: An R package for the collection and analysis of geometric morphometric shape data, *Meth. Ecol. Evol.*, **4**: 393–399.

Antonovics, J. 2006. Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. *Heredity*, **97**: 33–37.

Armsworth, P.R. and Roughgarden, J.E. 2005. The impact of directed versus random movement on population dynamics and biodiversity patterns. *Am. Nat.*, **165**: 449–465.

Bell, M.A. 1982. Differentiation of adjacent stream populations of threespine sticklebacks. *Evolution*, **36**: 89–199.

Berner, D. and Thibert-Plante, X. 2015. How mechanisms of habitat preference evolve and promote divergence with gene flow. *J. Evol. Biol.*, **28**: 1641–1655.

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

Bolnick, D.I. and Nosil, P. 2007. Natural selection in populations subject to a migration load. *Evolution*, **61**: 2229–2243.

Bolnick, D.I. and Otto, S.P. 2013. The magnitude of local adaptation under genotype-dependent dispersal. *Ecol. Evol.*, **3**: 4722–4735.

Bolnick, D.I., Caldera, E. and Matthews, B. 2008. Migration load in a pair of ecologically divergent lacustrine stickleback populations. *Biol. J. Linn. Soc.*, **94**: 373–387.

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

Bolnick, D.I., Shim, K.C. and Brock, C.D. 2015. Female stickleback prefer shallow males: sexual selection on nest microhabitat. *Evolution*, **69**: 1643–1653.

Bridcut, E.E. and Giller, P.S. 1995. Diet variability and foraging strategies in brown trout (*Salmo trutta*): an analysis from subpopulations to individuals. *Can. J. Fish. Aquat. Sci.*, **52**: 2543–2552.

Caldera, E.J. and Bolnick, D.I. 2008. Effects of colonization history and landscape structure on genetic variation within and among lacustrine populations of three-spine sticklebacks in a watershed. *Evol. Ecol. Res.*, **10**: 1–24.

Day, T. and McPhail, J.D. 1996. The effect of behavioural and morphological plasticity on foraging efficiency in the threespine stickleback (*Gasterosteus* sp.). *Oecologia*, **108**: 380–388.

Day, T., Pritchard, J. and Schluter, D. 1994. A comparison of two sticklebacks. *Evolution*, **48**: 1723–1734.

Edelaar, P. and Bolnick, D.I. 2012. Non-random gene flow: an underappreciated force in evolution and ecology. *Trends Ecol. Evol.*, **27**: 659–665.

Edelaar, P., Siepielski, A.M. and Clobert, J. 2008. Matching habitat choice causes directed gene flow: a neglected dimension in evolution and ecology. *Evolution*, **62**: 2462–2472.

Ehrlich, P.R. and Raven, P.H. 1969. Differentiation of populations. *Science*, **165**: 1228–1232.

Elmer, K.R., Lehtonen, T.K. and Meyer, A. 2009. Color assortative mating contributes to sympatric divergence of neotropical cichlid fish. *Evolution*, **63**: 2750–2757.

Feulner, P.G., Chain, F.J., Panchal, M., Huang, Y., Eizaguirre, C., Kalbe, M. *et al.* 2015. Genomics of divergence along a continuum of parapatric population differentiation. *PLoS Genet.*, **11**: e1004966.

Filchak, K.E., Roethele, J.B. and Feder, J.L. 2000. Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*, **407**: 739–742.

Foster, S.A., Wund, M.A., Graham, M.A., Earley, R.L., Gardiner, R., Kearns, T. *et al.* 2015. Iterative development and the scope for plasticity: contrasts among trait categories in an adaptive radiation. *Heredity*, **115**: 335–348.

Hendry, A.P. and Taylor, E.B. 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., Day, T. and Taylor, E.B. 2001. Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. *Evolution*, **55**: 459–466.

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

Hendry, A.P., Hudson, K., Walker, J.A., Räsänen, K. and Chapman, L. 2011. Genetic divergence in morphology-performance mapping between Misty Lake and inlet stickleback. *J. Evol. Biol.*, **24**: 23–35.

Hendry, A.P., Kaeuffer, R.E., Crispo, E., Peichel, C.L. and Bolnick, D.I. 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.

Honaker, J., King, G. and Blackwell, M. 2011. Amelia II: a program for missing data. *J Stat. Softw.*, **45**: 1–47.

Jiang, Y., Peichel, C.L. and Bolnick, D.I. 2015. Divergent rheotaxis contributes to divergent habitat preferences between lake and stream threespine stickleback. *Evolution*, **69**: 1925–1939.

Jiang, Y., Torrance, L., Peichel, C.L. and Bolnick, D.I. submitted. Heritable variation in lateral line sensory systems mediates rheotactic response of lake and stream stickleback. *J. Exp. Biol.*

Kaeuffer, R., Peichel, C.L., Bolnick, D.I. and Hendry, A.P. 2012. Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution*, **66**: 402–18.

Koskella, B., Lin, D.M., Buckling, A. and Thompson, J.N. 2012. The costs of evolving resistance in heterogeneous parasite environments. *Proc. R. Soc. Lond. B: Biol. Sci.*, **279**: 1896–1903.

Lavin, P.A. and McPhail, J.D. 1986. Adaptive divergence of trophic phenotype among freshwater populations of the threespine stickleback (*Gasterosteus aculeatus*). *Can. J. Fish. Aquat. Sci.*, **43**: 2455–2463.

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

Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.*, **17**: 183–189.

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

Mazzarella, A.B., Voje, K.L., Hansson, T.H., Taugbol, A. and Fischer, B. 2015. Strong and parallel salinity-induced phenotypic plasticity in one generation of threespine stickleback. *J. Evol. Biol.*, **28**: 667–677.

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

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

Oke, K.B., Bukhari, M., Kaeuffer, R., Rolshausen, G., Räsänen, K., Bolnick, D.I. *et al.* 2016. Does plasticity enhance or dampen phenotypic parallelism? A test with three lake–-stream stickleback pairs. *J. Evol. Biol.*, **29**: 126–143.

Pearish, S., Hostert, L. and Bell, A.M. 2013. Behavioral type–environment correlations in the field: a study of three-spined stickleback. *Behav. Ecol. Sociobiol.*, **67**: 765–774.

Reimchen, T.E. 1981. Microgeographical variation in *Littorina mariae* Sacchi and Rastelli and a taxonomic consideration. *J. Conch.*, **30**: 341–350.

Reimchen, T.E. and Nosil, P. 2006. Replicated ecological landscapes and the evolution of morphological diversity among *Gasterosteus* populations from an archipelago on the west coast of Canada. *Can. J. Zool.*, **84**: 643–654.

Reimchen, T.E., Stinson, E.M. and Nelson, J.S. 1985. Multivariate differentiation of parapatric and allopatric populations of threespine stickleback in the Sangan River watershed, Queen Charlotte Islands. *Can. J. Zool.*, **63**: 2944–2951.

Richardson, J.L. and Urban, M.C. 2013. Strong selection barriers explain microgeographic adaptation in wild salamander populations. *Evolution*, **67**: 1729–1740.

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

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

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

Saccheri, I.J., Rousset, F., Watts, P.C., Brakefield, P.M. and Cook, L.M. 2008. Selection and gene flow on a diminishing cline of melanic peppered moths. *Proc. Natl. Acad. Sci. USA*, **105**: 16212–16217.

Schindler, D.E., Hilborn, R., Chasco, B., Boatright, C., Quinn, T.P., Rogers, L.A. *et al.* 2010. Population diversity and the portfolio effect in an exploited species. *Nature*, **465**: 609–612.

Selander, R.K. and Kaufman D.W. 1975. Genetic structure of populations of brown snail (*Helix aspersa*). I. Microgeographic variation. *Evolution*, **29**: 385–401.

Skelly, D.K. 2004. Microgeographic countergradient variation in the wood frog, *Rana sylvatica*. *Evolution*, **58**: 160–165.

Slatkin, M. 1985. Gene flow in natural populations. *Annu. Rev. Ecol. Syst.*, **16**: 393–430.

Sundberg, P. 2008. Microgeographic variation in shell characters of *Littorina saxatilis* Olivi – a question mainly of size? *Biol. J. Linn. Soc.*, **35**: 169–184.

Thrall, P.H., Laine, A.L., Ravensdale, M., Nemri, A., Dodds, P.N., Barrett, L.G. *et al.* 2012. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecol. Lett.*, **15**: 425–435.

Walker, J.A. 1997. Ecological morphology of lacustrine threespine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. *Biol. J. Linn. Soc.*, **61**: 3–50.

Walker, J.A. 2004. Dynamics of pectoral fin rowing in a fish with an extreme rowing stroke: the threespine stickleback (*Gasterosteus aculeatus*). *J. Exp. Biol.*, **207**: 1925–1939.

Walker, J.A. and Westneat, M.W. 2002. Kinematics, dynamics, and energetics of rowing and flapping propulsion in fishes. *Integr. Comp. Biol.*, **42**: 1032–1043.

Wark, A.R. and Peichel, C.L. 2009. Lateral line diversity among ecologically divergent threespine stickleback populations. *J. Exp. Biol.*, **213**: 108–117.

Willacker, J.J., von Hippel, F.A., Wilton, P.R. and Walton, K.M. 2010. Classification of threespine stickleback along the benthic–limnetic axis. *Biol. J. Linn. Soc.*, **101**: 595–608.

Wright, S. 1946. Isolation by distance under diverse systems of mating. *Genetics*, **31**: 39–59.

Wund, M.A., Baker, J.A., Clancy, B., Golub, J.L. and Foster, S.A. 2008. A test of the 'flexible stem' model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *Am. Nat.*, **172**: 449–462.