

Natural selection drives patterns of lake–stream divergence in stickleback foraging morphology

D. BERNER,* D. C. ADAMS,†,‡ A.-C. GRANDCHAMP* & A. P. HENDRY*

*Redpath Museum & Department of Biology, McGill University, Montreal, QC, Canada

†Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA, USA

‡Department of Statistics, Iowa State University, Ames, IA, USA

Keywords:

determinism;
divergent selection;
evolutionary constraint;
Gasterosteus aculeatus;
limnetic–benthic differentiation;
line of least resistance;
parallel evolution;
phenotypic covariance.

Abstract

To what extent are patterns of biological diversification determined by natural selection? We addressed this question by exploring divergence in foraging morphology of threespine stickleback fish inhabiting lake and stream habitats within eight independent watersheds. We found that lake fish generally displayed more developed gill structures and had more streamlined bodies than did stream fish. Diet analysis revealed that these morphological differences were associated with limnetic vs. benthic foraging modes, and that the extent of morphological divergence within watersheds reflected differences in prey resources utilized by lake and stream fish. We also found that patterns of divergence were unrelated to patterns of phenotypic trait (co)variance within populations (i.e. the ‘line of least resistance’). Instead, phenotypic (co)variances were more likely to have been shaped by adaptation to lake vs. stream habitats. Our study thus implicates natural selection as a strong deterministic force driving morphological diversification in lake–stream stickleback. The strength of this inference was obtained by complementing a standard analysis of parallel divergence in means between discrete habitat categories (lake vs. stream) with quantitative estimates of selective forces and information on trait (co)variances.

Introduction

Parallel evolution, i.e. the repeated evolution of similar phenotypes under similar environmental circumstances, is considered strong evidence for the operation of natural selection (Endler, 1986; Schluter, 2000). Parallel evolution is typically explored by comparing two or more replicate lineages (e.g. different species or conspecific populations) with respect to their phenotypic divergence between distinct habitat types. The deterministic role of shared selective pressures relative to other evolutionary forces (e.g. historical contingency and gene flow) is then inferred from the concordance among lineages in the nature of habitat-related divergence within those lineages (McPeck, 1995; Losos *et al.*, 1998; Rundle *et al.*, 2000; McGuigan *et al.*, 2003; Jastrebski & Robinson, 2004; Langerhans & DeWitt, 2004; Langerhans *et al.*, 2004; McKinnon *et al.*, 2004; Ostbye *et al.*, 2006).

Correspondence: Daniel Berner, Redpath Museum & Department of Biology, McGill University, 859 Sherbrooke St W., Montreal, QC, Canada H3A 2K6.

Tel.: +1 514 398 4086; fax: +1 514 398 3185;
e-mail: daniel.berner@mail.mcgill.ca

We here wish to address three potential ambiguities in this approach to inferring evolutionary process from diversification patterns.

The first ambiguity arises because populations are typically assigned to discrete habitat categories (e.g. ‘lake’ vs. ‘stream’ and ‘high predation’ vs. ‘low predation’) without quantification of the ecological conditions actually experienced by each population. The potential concern of this approach is that local selective forces may vary among replicate ‘habitats’ and thereby cause deviations from expected parallelism in patterns of phenotypic divergence. Ignoring this variation may result in an underestimation of natural selection’s deterministic power. We address this issue by complementing a typical analysis based on two discrete habitat categories (lake and stream) with quantitative information on ecological conditions within each habitat. This allows us to examine how divergence between discrete habitats within replicate lineages is modified by variation in divergent selection.

The second ambiguity concerns an alternative explanation for concordant patterns of phenotypic divergence in mean trait values. Rather than being driven primarily

by repeated patterns of divergent selection, trajectories of divergence may be biased by the trait (co)variance structure within populations. The reason is that evolution might be the easiest (or the most rapid) in the direction where traits show the highest genetic variances and covariances (Arnold, 1992; Björklund, 1996; Schluter, 1996; Arnold *et al.*, 2001). A specific prediction derived from this idea and addressed in our study is that the major axis of diversification in trait means among populations should line up with the major axis of trait (co)variation among individuals within those populations (the 'genetic line of least resistance', Schluter, 1996).

The third ambiguity, related to the second, concerns whether trait (co)variances are shaped by habitat-related selective conditions instead of simply influencing evolutionary responses to selection (Schluter, 1996; Arnold *et al.*, 2001; McGuigan, 2006). For instance, if selection tends to be more stabilizing within one habitat but more disruptive within the other habitat, we expect to find repeated differences in variances and possibly covariances between the habitats. Empirical studies, however, have usually focused only on responses in trait means, and how both the phenotypic and genetic (co)variance structure are influenced by selection remains little understood (Roff, 2000; Stepan *et al.*, 2002; McGuigan, 2006).

Our study

The goal of our study was to address the above ambiguities by exploring patterns of phenotypic divergence in the foraging morphology of threespine stickleback (*Gasterosteus aculeatus*) fish. Stickleback are particularly well suited for this task because morphologically divergent populations can be found in lakes and streams (habitats) within multiple watersheds (replicate lineages). Previous work has shown that stream-resident fish usually have deeper bodies and fewer gill rakers than do lake fish, differences that have an additive genetic basis (Lavin & McPhail, 1993; Hendry *et al.*, 2002). Several lines of evidence also suggest that these differences are an adaptive response to divergent selection. First, some differentiation in these traits can be maintained despite substantial gene flow (Hendry & Taylor, 2004; Moore *et al.*, 2007). Second, the differences have evolved repeatedly in many watersheds independently colonized by anadromous (sea-run) ancestors (Thompson *et al.*, 1997; Hendry & Taylor, 2004). Third, the trait differences match functional expectations related to different foraging opportunities and swimming modes (details below). Nevertheless, the consistency of multivariate divergence has not been formally quantified in lake–stream stickleback. The importance of deterministic selection relative to other evolutionary forces thus remains to be evaluated.

In a first step, we therefore quantify similarity and dissimilarity in multivariate phenotypic divergence

between lake and stream stickleback in eight watersheds, each representing an independent evolutionary lineage. We then test for an association between habitat-related divergence and quantitative estimates of ecological conditions (diet) that probably influences the strength of divergent selection between the habitats. In a second step, we test whether morphological divergence between lake and stream fish within watersheds is biased along the line of least resistance within the resident populations (e.g. Schluter, 1996). In a third step, we test whether adaptation to lake and stream environments is associated with changes in trait (co)variances. As we will show, the integration of these different analyses implicates natural selection as a strong determinant of morphological diversification in lake–stream stickleback.

Materials and methods

Stickleback samples

Our study is based on fish collected from eight watersheds on Vancouver Island, BC, Canada. The stickleback in these different watersheds almost certainly derive from independent post-glacial colonization by anadromous ancestors (Thompson *et al.*, 1997; Hendry & Taylor, 2004), and they therefore serve as evolutionary replicates. Within each watershed, fish were sampled from the lake and from the outlet stream, yielding 16 samples in total (Table 1). Water distances between the lake and stream sites varied from 1540 to 4400 m in the different watersheds.

The stickleback for morphological analysis were collected with unbaited minnow traps or dip nets in the spring of 2006, except for the Misty watershed where collections were made in the spring of 2003. Of the captured fish, we haphazardly selected and retained

Table 1 Description of the sites sampled for the analysis of stickleback foraging morphology.

Watershed	Code	Habitat	Latitude (N)	Longitude (W)	Sample size
Beaver	Be	Lake	50°35'59.9"	127°19'11.6"	28
		Stream	50°35'16.2"	127°22'43.8"	28
Boot	Bo	Lake	50°03'00.2"	125°32'27.4"	30
		Stream	50°02'25.3"	125°33'53.8"	30
Joe's	Jo	Lake	50°37'32.2"	127°29'06.8"	28
		Stream	50°37'29.6"	127°31'06.9"	27
McCreight	Mc	Lake	50°19'18.0"	125°38'06.2"	30
		Stream	50°19'57.2"	125°39'10.2"	29
Misty	Mi	Lake	50°36'34.9"	127°15'55.4"	30
		Stream	50°36'59.1"	127°17'04.1"	30
Morton	Mo	Lake	50°07'19.2"	125°28'24.7"	27
		Stream	50°07'01.2"	125°26'18.1"	28
Pye	Py	Lake	50°18'09.3"	125°35'27.5"	29
		Stream	50°18'53.0"	125°33'37.6"	29
Robert's	Ro	Lake	50°13'37.5"	125°33'13.1"	28
		Stream	50°14'17.4"	125°34'55.2"	30

27–30 per site (461 fish in total), excluding those less than a year old and those showing any signs of gravidity. The retained fish were killed with an overdose of MS222 and preserved in 95% ethanol, which was replaced at least twice. After 7 months of preservation, we determined individual fresh mass, took morphological measurements and dissected the fish to determine sex. Including sex as a statistical factor (results not shown) did not materially influence any of our conclusions, and so we here present analyses with the sexes pooled.

The above samples were not processed in a way that allowed the optimal preservation of stomach contents. We therefore collected an additional 20 fish at each of the 16 sites in the spring of 2007. The sampling protocol was the same as above, except that we always inspected the traps less than 4.5 h after they were set. In addition, after the fish were killed, we opened their abdominal cavity before transfer to ethanol. These procedures ensured that the stomach contents were preserved in good condition.

Foraging traits and measurement

We studied six morphological traits: three related to gill structure and three related to body depth (Fig. 1). Gill traits were measured (or counted) on each fish's left side at 15–45 \times magnification under a stereomicroscope fitted with an ocular micrometer (maximal precision 0.01 mm). Gill 'arch length' was the length of the ventral bone of the first gill arch from its base to its joint with the dorsal arch bone. Gill 'raker number' (the number of

bony protuberances) was counted on the same ventral arch bone. Gill 'raker length' was the average length of rakers two to four, as counted from the joint with the dorsal arch bone. Body depth traits were measured on the preserved specimens with a digital caliper (precision 0.01 mm). 'Head depth' was the distance from the posterior edge of the cranium to the anterior tip of the ectocoracoid (see Walker, 1997 for details). 'Anterior depth' was the distance from the insertion of the first dorsal spine to the insertion of the pelvic spine (ventral). 'Posterior depth' was the distance from the insertion of the third dorsal spine to immediately posterior to the insertion of the anal spine.

These specific gill traits were chosen because they are clearly under divergent selection between limnetic (zooplankton) and benthic (macro-invertebrate) foraging modes. In particular, previous studies have shown that individuals with more numerous and longer gill rakers feed preferentially and more efficiently on zooplankton, whereas those with fewer and shorter gill rakers feed preferentially and more efficiently on macro-invertebrates (Bentzen & McPhail, 1984; Schluter & McPhail, 1992; Robinson, 2000; Bolnick, 2004). These specific associations have not been tested directly for our study system, but it nevertheless seems very likely that selection favours a more developed gill structure in lake fish than in stream fish. The reason is that lake fish feed more often on zooplankton than do stream fish (present study), and lake and stream fish show the expected phenotypic and additive genetic divergence in gill raker number (Gross & Anderson, 1984; Lavin & McPhail, 1993; Hendry *et al.*, 2002). Both gill raker number and length generally display high heritabilities in stickleback populations (0.34–0.58) (Hagen, 1973; Schluter, 1996; Hermida *et al.*, 2002; Aguirre *et al.*, 2004). Gill arch length has not been studied previously in stickleback but was included here because it influences spacing between the gill rakers, and so presumably impacts the efficiency of capturing small prey (Gerking, 1994).

The body depth traits were chosen because they too are expected to be under divergent selection between lakes and streams. Stream stickleback feed mainly on benthic prey in complex environments, which should select for high manoeuvrability and therefore deeper bodies (Webb, 1984; Blake, 2004). Lake stickleback feed more often on limnetic prey in the open water, which should select for sustained swimming ability and therefore shallower bodies with lower drag. These functional predictions match observations of greater body depth for stickleback in streams when compared with lakes (Reimchen *et al.*, 1985; Lavin & McPhail, 1993; Caldecutt & Adams, 1998; Hendry *et al.*, 2002; Hendry & Taylor, 2004), in benthic when compared with limnetic habitats within lakes (Schluter & McPhail, 1992; Schluter, 1993; Robinson, 2000) and in lakes with complex substrates when compared with lakes with simple substrates (Walker, 1997). Moreover, these body depth differences

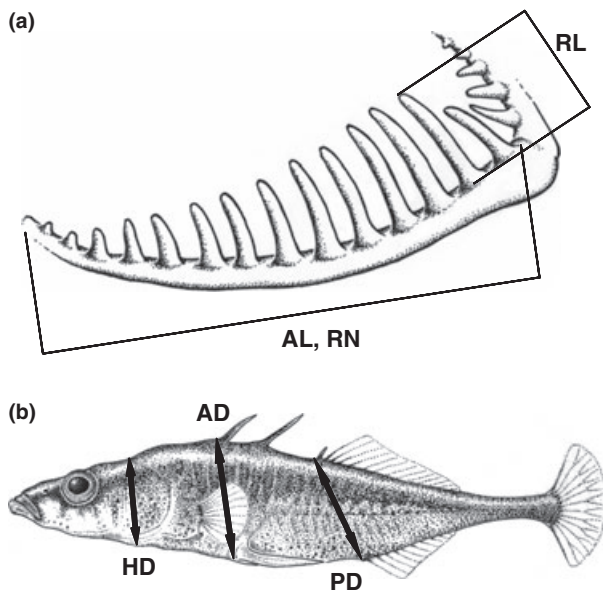


Fig. 1 Stickleback gill raker structure (a) and body depth traits (b). The gill traits are raker length (RL), arch length (AL) and raker number (RN). The body depth traits are head depth (HD), anterior body depth (AD) and posterior body depth (PD).

have a strong additive genetic basis in lake–stream stickleback (Lavin & McPhail, 1993; Hendry *et al.*, 2002) and in other stickleback populations (Schluter *et al.*, 2004; Albert *et al.*, 2008).

All of the measured traits, except for gill raker number, are correlated with body size, which necessitated size standardization. We first ln-transformed all measurements and counts so as to decouple (co)variances from trait means and measurement scales (Bookstein *et al.*, 1985). We next conducted principal components analysis (PCA) with all individuals pooled and retained each individual's score on the first principal component (PC1). This PC1 vector accounted for 77% of the total variance and showed positive loadings for all traits (0.42–0.47) except for raker number (−0.03). Moreover, the PC1 scores were highly correlated with individual fresh mass within all populations (Pearson's $r = 0.90$ – 0.99 , mean 0.96), confirming the suitability of these scores as a generalized body size index (Jolicoeur, 1963). We then regressed each trait against the PC1 scores (all individuals pooled) and used the residuals as new size-standardized morphological traits for all subsequent analyses except for the comparison of (co)variances (see below). Two alternative approaches to control for body size (using PC2–PC6 as size-free morphology, and using PC1 scores as size covariate) produced highly consistent results throughout. We therefore preferred working with PC1 residuals because this greatly facilitated the resampling procedures (see below), and because it allowed expressing all results in terms of the original six traits.

Measurement error was estimated by measuring each trait on two separate occasions for 20 haphazardly selected individuals. Correlations between the first and second measurements were excellent for all traits (all $r \geq 0.978$). For the three body depth traits, we additionally examined how well caliper measurements from 20 haphazardly selected ethanol-preserved specimens agreed with corresponding measurements taken from digital photographs of the same individuals taken at capture. Again, the correlation was very high (all $r \geq 0.992$).

Analysing multivariate divergence

We first subjected our morphological data set to multivariate analysis of variance (MANOVA) with habitat (lake or stream), watershed and their interaction as factors. This analysis indicated a significant interaction, but did not allow us to examine how the magnitude and orientation of lake–stream divergence differed among particular watersheds. For this, we used phenotypic change vector analysis (PCVA) (Adams & Collyer, 2007; Collyer & Adams, 2007). For each watershed, we first used a general linear model to calculate the lake and stream population centroids (multivariate least-squares means). Connecting the lake and stream centroids within a watershed then yielded the habitat-related PCV within

that watershed. We then compared PCVs for each of the 28 possible watershed pairings with respect to vector magnitude (length difference between the vectors) and orientation (angle between the vectors). Statistical significance for the differences in these vector attributes was determined by comparing the observed values to corresponding random distributions. These distributions were generated by applying the residual randomization protocol described in detail in the statistical appendices of Adams & Collyer (2007) and Collyer & Adams (2007). Residual randomization (and all resampling tests below) was carried out with R 2.6.1 (R Development Core Team 2006).

Patterns of multivariate divergence among populations were visualized by first performing PCA on the 16 observed centroids (i.e. using the multivariate population means as data points). We then plotted population means along PC1 and PC2, with standard errors calculated by projecting individuals onto these axes. Finally, we used eigenvalues and trait loadings to assess the importance and orientation in trait space of PC1 and PC2. Visualization using canonical variate axes (CVs) yielded qualitatively similar patterns, but these are not presented here because the distortion of trait space in CV analyses makes interpretation more difficult.

The above multivariate analysis of morphology was supplemented by a univariate test for differences in gill raker spacing. For this, we expressed raker spacing for each individual as the length of the arch bone divided by the number of gill rakers, and used ANOVA with habitat, watershed and the interaction as factors.

Estimating the strength of divergent selection

We used information on limnetic vs. benthic prey types found in the stomachs of stickleback to quantify the strength of divergent selection on foraging morphology between lake and stream habitats. This indirect approach was chosen because attempts to quantitatively compare food availability would be compromised by the necessarily different sampling procedures (plankton tows are not possible in many streams and benthic sampling proceeds very differently in flowing vs. static water). Conveniently, however, previous work on stickleback has documented a strong correlation between foraging morphology (raker traits and body depth) and the exploitation of limnetic vs. benthic prey resources, both among and within populations (Gross & Anderson, 1984; Schluter & McPhail, 1992; Robinson, 2000). Furthermore, the functional link between morphology and foraging performance on limnetic vs. benthic prey has been substantiated by the identification of foraging trade-offs in laboratory experiments (Bentzen & McPhail, 1984; Robinson, 2000). As these studies strongly suggest, prey-mediated selection is the primary driver of foraging traits in stickleback (and fish in general, see Robinson & Wilson, 1994; Skúlason & Smith, 1995). Hence, even

though we do not expect that stomach content data provide a precise description of local prey availability, they should nevertheless encapsulate a key selective factor acting on foraging morphology. We further assume that stomach content, representing only a snapshot of an individual's foraging habits, provides reliable information on individual long-term foraging. Indeed, recent work on stickleback indicates repeatability across years and strong agreement with inferences from stable isotope analysis (Bolnick *et al.*, 2008).

Prey items were identified using a stereomicroscope at 15–45× magnification. Following Schluter & McPhail (1992), prey were classified as limnetic (open water), benthic (in or on the substrate) or 'other' (potentially in the open water or on the substrate). The identification and assignment of individual prey items to these categories were based on Pennak (1989) and Thorp & Covich (2001). The most frequently encountered benthic prey included diptera larvae (Chironomidae and Ceratopogonidae), chydorid cladocera, mayfly and caddis fly larvae and ostracoda. The most frequently encountered limnetic prey included nonchydorid cladocera, calanoid copepods and emerging mayflies and diptera. 'Other' prey included mainly cyclopoid copepods.

All statistical analyses of stomach content data are based on the number of limnetic prey relative to the number of limnetic and benthic prey combined (hereafter 'proportion of limnetic prey' or PLP). This was carried out for simplicity, for ease of biological interpretation and for agreement with previous work (Schluter & McPhail, 1992). We first tested PLP differences among sites by using ANOVA with habitat, watershed and the interaction as factors. We next calculated the difference in average PLP between lake and stream fish within each watershed. We then constructed an 8 × 8 distance matrix describing the difference between each pair of watersheds in the difference in PLP between lake and stream fish within those watersheds. Response matrices were created analogously for the angle between PCVs (similarity among watersheds in the orientation of divergence), for the absolute length of PCVs (similarity among watersheds in the overall magnitude of divergence) and for the magnitude of divergence along PC1 extracted from the population centroids. The latter response matrix was included because PC1 was found to capture morphological variation most probably associated with foraging performance. The association between the predictor (PLP) matrix and each of the three morphological response matrices was then tested in separate Mantel tests based on 1999 matrix permutations (Manly, 2007).

Lines of least resistance

The line of least resistance was originally defined as the major axis of the additive genetic trait (co)variance matrix (Schluter, 1996). Estimates of genetic (co)vari-

ances, however, are notoriously difficult to obtain (Lynch & Walsh, 1998) and are unavailable for our lake–stream populations. Our study therefore uses the phenotypic line of least resistance as a surrogate for its genetic counterpart. Even though a perfect correspondence between the phenotypic and genetic (co)variance structure is unlikely, the former generally approximates the latter quite well (Cheverud, 1988; Roff, 1995, 1996; Schluter, 1996; Roff *et al.*, 1999; Badyaev & Hill, 2000; Bégin & Roff, 2004). This holds particularly for traits with substantial heritability (Lande, 1979), which is the case in our stickleback system (see above). Hence, the phenotypic line of least resistance should be useful to explore the hypothesis of variational bias to phenotypic diversification.

To obtain estimates of the phenotypic line of least resistance, we extracted the major axis (PC1) of (size-standardized) trait (co)variance for each population separately using PCA. These vectors were scaled to unit length and their trait loadings averaged across the lake and stream population within each watershed. This yielded watershed-specific estimates of the average phenotypic line of least resistance. We preferred to use these average values rather than estimates from the stream populations only, because it is uncertain whether the stream fish represent the true ancestors of the lake populations studied. This decision, however, did not influence our conclusions in any way because both approaches yielded almost identical results (details not presented). We next tested whether the line of least resistance within each watershed was associated with the direction of divergence between lake and stream centroids within that watershed. We did so by calculating the smaller angle between the average line of least resistance and the PCV for a watershed, and then evaluated the significance of this angle by comparison with a random distribution.

The random distributions were generated by separately bootstrapping (resampling with replacement) lake and stream samples 2000 times. For each resampling iteration, we calculated the watershed-specific average line of least resistance and the PCV, as well as the angle between them (as described above). The two-tailed (hence conservative) lower 95% confidence limit for the angle was then estimated by the empirical lower 2.5 percentile of the random angle distribution (Manly, 2007). The null hypothesis of directional concordance of the PCV with the line of least resistance within a watershed was rejected when the lower confidence limit of the random angle distribution was greater than zero. Parametric confidence limits produced similar results and are not reported here.

Habitat-related divergence in (co)variances

For the accurate estimation of phenotypic (co)variance matrices, size standardization was performed within each population separately. For this, we applied PCA directly

to ln-transformed measurements, regressed each trait against the resulting PC1 scores (body size) and used the residuals to calculate the (co)variance matrix. We then investigated habitat-related differences in (co)variances with the jackknife MANOVA approach (Roff, 2002; Bégin *et al.*, 2004). Briefly, this approach converts population-level (co)variances to individual (co)variance pseudo-values by sequentially deleting single individuals from a population and recalculating (co)variances according to the jackknife procedure (Manly, 2007). The resulting pseudovalues represent approximate random variables that can be organized as data columns and analysed in factorial designs. For significance testing, we performed MANOVA on the pseudovalues with habitat, watershed and the interaction as factors. This was followed by an inspection of univariate significances to identify (co)variances that contributed most strongly to variation associated with the habitat factor (Roff, 2002).

Results

Divergence in foraging morphology

The major axis of morphological diversification among our stickleback populations (PC1 based on population centroids) explained 69% of the variation and showed that stream fish usually had reduced gill raker structures (especially raker length and number) and increased body depth (head, anterior and posterior) relative to their lake counterparts (Fig. 2; univariate descriptive statistics are given in Appendix 1). The second axis of morphological

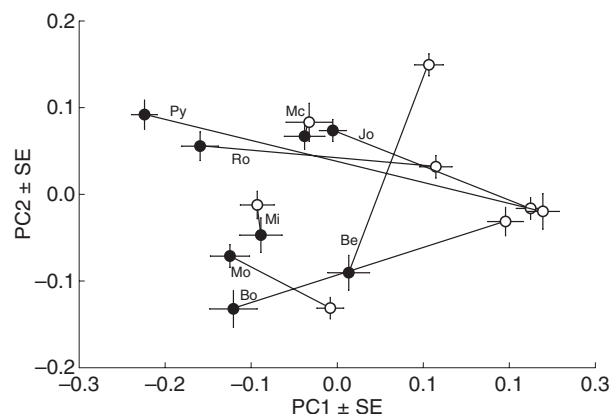


Fig. 2 Phenotypic change vectors connecting lake (black symbols) and stream (white symbols) population centroids within each of the eight watersheds (full watershed names in Table 1). The centroids are plotted along their first two principal components (PCs) and error bars are standard errors calculated from individuals projected on those axes. PC1 accounts for 69% (eigenvalue) of the total variance among sample means and has the following trait loadings: raker number -0.52 , raker length -0.70 , arch length -0.17 , head depth 0.21 , anterior body depth 0.29 , posterior body depth 0.29 . PC2 accounts for 24% of total variance and has the following loadings (same order): -0.39 , 0.50 , -0.65 , -0.26 , 0.03 and 0.31 .

divergence among populations (PC2, 24% of the variance) was most strongly associated with opposing changes in gill arch length and raker length. Morphological variation along this second axis, however, showed no obvious association with habitat.

Multivariate analysis of variance indicated significant differences in morphology between lake and stream habitats ($F_{5,441} = 112.5$, $P < 0.0001$) and among watersheds ($F_{35,1858} = 26.6$, $P < 0.0001$), with a significant interaction ($F_{35,1858} = 20.0$, $P < 0.0001$). All six morphological traits contributed significantly to the diversification between habitats, among watersheds and to the interaction, as indicated by the univariate significances (all $P < 0.0001$, details not presented). ANOVA revealed, however, that raker spacing did not differ among habitats ($F_{1,7} = 0.84$, $P = 0.39$) or watersheds ($F_{7,7} = 1.63$, $P = 0.27$; interaction: $F_{7,445} = 2.24$, $P = 0.03$), indicating that changes in raker number were associated with changes of similar magnitude in gill arch length.

Phenotypic change vector analysis indicated that vectors of lake–stream divergence sometimes differed ($P < 0.05$) among watersheds either in magnitude only (five of 28 pairwise comparisons), in orientation only (six of 28) or in both magnitude and orientation (15 of 28) (Table 2). The only nonsignificant contrasts for *both* magnitude and angle were those between Robert's and Boot watersheds, and between Robert's and Joe's. The vector-based analysis thus made clear that the strength and orientation of habitat-related change in foraging morphology differed greatly among watersheds.

Stomach contents

Analysis of variance revealed that the proportion of limnetic prey (PLP) in stickleback stomachs was much lower (often close to zero) for stream stickleback than for lake stickleback ($F_{1,296} = 97.2$, $P < 0.0001$; Fig. 3). This proportion also differed among watersheds ($F_{7,296} = 19.5$, $P < 0.0001$) and showed a significant interaction between habitat and watershed ($F_{7,296} = 21.1$, $P < 0.0001$). The interaction occurred mainly because two watersheds (McCreight and Morton) did not show differences in diet between lake and stream habitats.

Mantel tests indicated a strong association between divergence in stomach contents and the magnitude of divergence along PC1 ($r = 0.442$, $P = 0.026$) (Fig. 4). A similar relationship was suggested for the overall magnitude of divergence in PCVs ($r = 0.423$, $P = 0.072$). By contrast, divergence in stomach contents was not related to differences in the orientation of phenotypic divergence (angle between PCVs, $r = -0.159$, $P = 0.701$).

The phenotypic line of least resistance and habitat-related (co)variances

The phenotypic line of least resistance within each individual population (PC1 extracted from the [co]var-

Table 2 Results from the phenotypic change vector analysis (PCVA). The upper right semi-matrix gives for each pairwise watershed contrast the observed difference in the magnitude of divergence (lengths of PCVs, first row) and the observed difference in orientation of divergence (angle between PCVs, second row). The lower semi-matrix gives the corresponding *P*-values for magnitude and orientation (*P* < 0.05 in bold), based on 1999 randomizations. A significant *P*-value indicates that two PCVs differ statistically in the corresponding vector attribute, which is mostly the case in the watersheds studied. For full watershed names see Table 1.

Watershed	Be	Bo	Jo	Mc	Mi	Mo	Py	Ro
Be	–	0.065	0.019	0.246	0.223	0.135	0.211	0.015
		52.9	91.7	40.2	37.8	94.1	83.6	71.3
Bo	0.191	–	0.084	0.310	0.288	0.200	0.147	0.049
	0.001		39.4	64.5	79.5	45.2	33.1	25.6
Jo	0.700	0.087	–	0.226	0.204	0.116	0.231	0.035
	< 0.001	0.007		92.0	112.2	13.0	13.8	24.3
Mc	< 0.001	< 0.001	< 0.001	–	0.022	0.111	0.457	0.261
	0.008	< 0.001	< 0.001		55.9	93.0	87.5	76.6
Mi	< 0.001	< 0.001	< 0.001	0.639	–	0.088	0.435	0.239
	0.011	< 0.001	< 0.001	< 0.001		109.9	109.8	90.8
Mo	0.005	< 0.001	0.021	0.023	0.070	–	0.346	0.151
	< 0.001	0.003	0.511	< 0.001	< 0.001		21.3	22.9
Py	< 0.001	0.003	< 0.001	< 0.001	< 0.001	< 0.001	–	0.196
	< 0.001	0.021	0.449	< 0.001	< 0.001	0.152		21.8
Ro	0.746	0.310	0.478	< 0.001	< 0.001	0.004	< 0.001	–
	< 0.001	0.082	0.099	< 0.001	< 0.001	0.116	0.139	

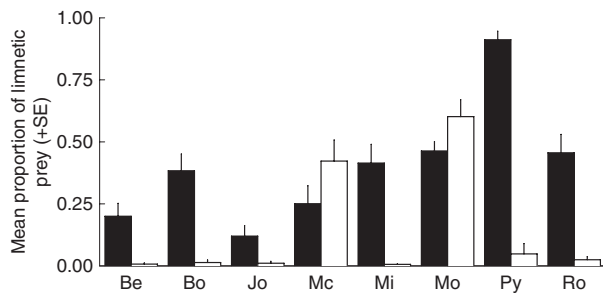


Fig. 3 The mean proportion of limnetic prey is greater for lake (black bars) than for stream (white bars) stickleback in all watersheds except McCreight (Mc) and Morton (Mo). Error bars are one standard error.

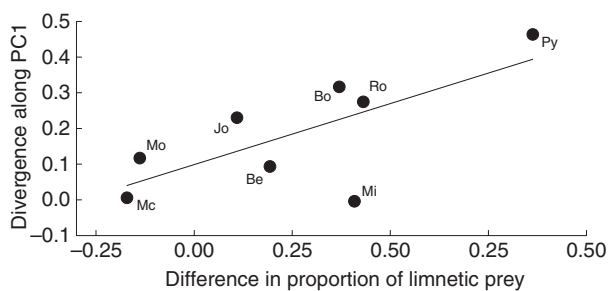


Fig. 4 Lake and stream stickleback within watersheds are more divergent morphologically (PC1 from Fig. 2) when they are also more different in the proportion of limnetic prey utilized (from Fig. 3).

ian matrix) was very strongly driven by variance in raker length (Table 3). (Full phenotypic [co]variance and correlation matrices averaged over all populations are provided in Appendix 2.) The angle between the average line of least resistance and the PCV within a watershed

Table 3 The phenotypic line of least resistance (PC1 of the phenotypic [co]variance matrix) averaged over all 16 populations. This axis was consistently driven very strongly by variance in raker length, as illustrated by the comparatively very high loading of this trait and by the low standard error (in parenthesis).

Trait	Loading on PC1
Raker number	–0.024 (0.015)
Raker length	0.966 (0.005)
Arch length	0.062 (0.022)
Head depth	0.058 (0.009)
Anterior depth	–0.002 (0.004)
Posterior depth	–0.056 (0.007)

was usually substantial (mean: 50.4 °, range: 13–83 °; Table 4), and the lower 95% confidence limit for the angle was always greater than zero. Hence, the orientation of lake–stream divergence was consistently unrelated to the phenotypic line of least resistance.

Multiple analyses of variance performed on jackknife pseudovalues for trait (co)variances indicated significant differences associated with habitat ($F_{21,425} = 2.19$, $P = 0.002$), watershed ($F_{147,2848} = 2.49$, $P < 0.0001$) and their interaction ($F_{147,2848} = 1.79$, $P < 0.0001$). Three (co)variances displayed significant univariate differences between the habitats. Variances in raker length ($F_{1,445} = 4.72$, $P = 0.030$) and in posterior depth ($F_{1,445} = 4.51$, $P = 0.034$) were greater in stream stickleback than in lake stickleback, as was the (negative) covariance between these two traits ($F_{1,445} = 7.33$, $P = 0.007$) (Fig. 5). However, there was also considerable variation among watersheds in the extent of lake–stream divergence in trait (co)variances. Differences in the (co)variance structure due to watershed and the interaction were primarily associated with variances and covariance in the gill structure and head depth (details not shown).

Table 4 Observed angle (degrees) for each watershed between the PCV and the average line of least resistance. The angles can range from 0° (vectors parallel) to 90° (vectors orthogonal). Associated lower 95% confidence limits were estimated by the lower empirical 2.5 percentiles of the random distribution of angles. Confidence limits are greater than zero for all watersheds, indicating that PCVs do not align with the line of least resistance.

Watershed	Observed angle	Lower confidence limit
Beaver	82.6	66.9
Boot	55.3	45.3
Joe's	23.7	16.4
McCreight	82.7	17.3
Misty	71.1	24.8
Morton	13.2	9.2
Pye	47.3	33.9
Robert's	26.9	17.0

Discussion

General patterns of lake–stream divergence

The basic outcome of our analysis was that the main axis of variation among the 16 stickleback populations tended to polarize lake and stream fish. In particular, stream fish usually had fewer and shorter gill rakers than lake fish, as well as deeper bodies (head, anterior and posterior). The results for gill raker number and anterior body depth parallel those obtained in previous analyses of lake–stream populations (Reimchen *et al.*, 1985; Lavin & McPhail, 1993; Hendry *et al.*, 2002; Hendry & Taylor, 2004), whereas the other traits have not been examined before. These differences probably have a strong genetic

basis (as opposed to reflecting phenotypic plasticity), as has been found in previous work on lake–stream stickleback (Gross & Anderson, 1984; Lavin & McPhail, 1993; Hendry *et al.*, 2002) and in other stickleback systems (Hatfield, 1997; Peichel *et al.*, 2001; Schluter *et al.*, 2004; Albert *et al.*, 2008).

The features of lake–stream divergence that are shared across watersheds strongly suggest that habitat-related divergent selection has influenced morphology. The standard approach of using parallel evolution in discrete habitat classes as a means of inferring the role of deterministic selection therefore works in a general sense for lake–stream stickleback (see also Hendry & Taylor, 2004). We here contend, however, that the above simple analysis and inference is incomplete because it still suffers from the ambiguities described in the Introduction and elaborated in the next sections.

Causes of diversification

Our vector-based approach showed that despite similarities in phenotypic responses, lake–stream divergence differed substantially in magnitude and orientation among watersheds (Table 2, Fig. 2). This result emphasizes the need to scrutinize the potential causes of watershed-specific patterns. As a first ambiguity, we have suggested that local selection pressures may not be adequately captured by the 'lake' vs. 'stream' dichotomy. Our approach to investigating this possibility was to incorporate quantitative information on prey-mediated divergent selection on foraging-related traits, as inferred from stomach content analysis. In this regard,

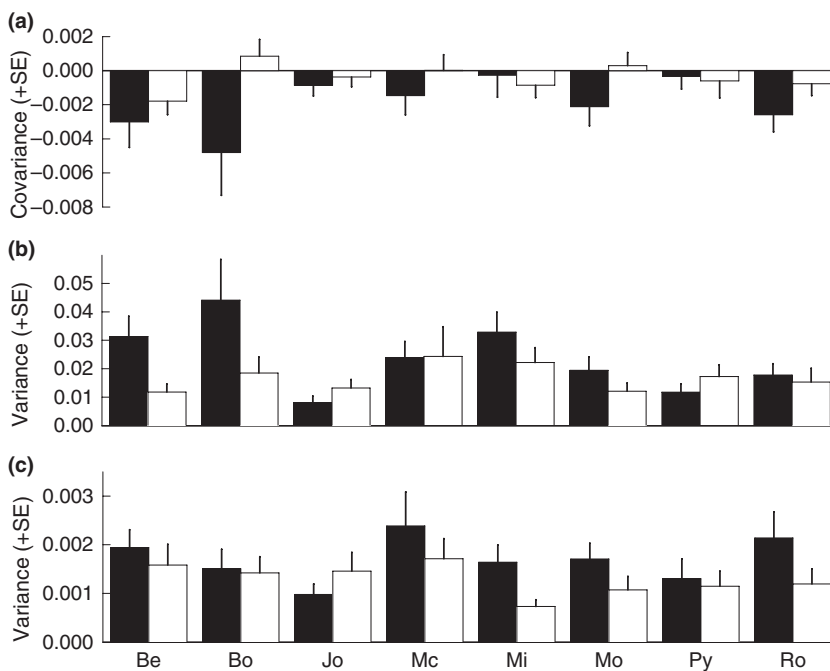


Fig. 5 Lake stickleback (black bars) exhibit greater (co)variances in foraging traits than stream stickleback (white bars). Displayed are the (co)variances that showed a univariate $P < 0.05$ for the habitat factor in MANOVA. These are: (a) raker length – posterior body depth covariance, (b) raker length variance, (c) posterior body depth variance. Error bars are jackknife standard errors.

we identified striking variation among watersheds in the degree to which lake and stream stickleback differed in foraging on benthic vs. limnetic food types. This variation among watersheds was strongly correlated with the magnitude of morphological divergence within watersheds. For example, the watersheds with the least lake–stream difference in diet (McCreight and Morton) also showed low morphological divergence between lake and stream fish (Fig. 4). At the opposite extreme, the watershed with the greatest lake–stream difference in diet (Pye) also showed the greatest phenotypic response. As described above, previous work strongly suggests that prey resources drive morphology (rather than morphology determining foraging *independently* of local prey availability) in stickleback (Bentzen & McPhail, 1984; Gross & Anderson, 1984; Schluter & McPhail, 1992; Robinson, 2000) and in fish generally (Robinson & Wilson, 1994; Skúlason & Smith, 1995). Our results thus indicate that divergence in foraging morphology is driven by the strength of divergent selection.

Another possible contributor to variation in lake–stream divergence is gene flow (Hendry *et al.*, 2002; Hendry & Taylor, 2004; Moore *et al.*, 2007). In particular, the Misty watershed showed by far the greatest deviation between the observed magnitude of lake–stream morphological divergence and that expected based on the difference in diet (Fig. 4). This result agrees perfectly with evidence for strong maladaptation in the Misty outlet stream owing to very high gene flow from Misty Lake (Hendry *et al.*, 2002; Moore *et al.*, 2007). Repeating the PC1 Mantel test with the Misty watershed excluded produced a very strong correlation between morphological divergence and the expected magnitude of diet-based divergent selection (matrix correlation $r = 0.769$, $P = 0.007$; standard correlation $r = 0.92$). The stream populations in the other watersheds may have been less susceptible to gene flow (as opposed to some of these studied by Hendry & Taylor, 2004) because we here selected stream sites that were farther from the lakes. Our data from the Misty watershed further support the view that stomach content reflects local prey availability, rather than reflecting phenotype-specific foraging independent of local prey availability. Because morphological lake–stream divergence in the Misty watershed is negligible (Fig. 2; Appendix 1), both populations should use similar prey resources if foraging was determined primarily by the phenotype. Instead, stomach content differed dramatically, indicating that the lake and stream populations indeed experience contrasting prey resources but have not strongly adapted to them.

In summary, our study makes clear that divergent selection mediated by limnetic vs. benthic foraging conditions strongly determines morphological differences among stickleback populations inhabiting lake and stream environments. Some of this effect could be seen

by grouping populations into the simple ‘lake’ and ‘stream’ habitat categories. However, quantitative measurements of a key selective factor (diet) within those habitats substantially improved the habitat–morphology association and thus revealed an even stronger effect of selection in driving morphological diversification. The reason is that local foraging conditions accounted for *both* similarities and differences among watersheds in the magnitude of lake–stream divergence. Additional improvements could be made by incorporating information on gene flow. In short, studies of the causes of morphological diversification greatly benefit from the quantitative measurement of multiple evolutionary forces.

Lake–stream divergence and the line of least resistance

The second ambiguity we raised in the Introduction relates to another potential contributor to similarity among multiple evolutionary responses: phenotypic divergence may be biased along the ‘lines of least resistance’ (Schluter, 1996). Such constraints were not evident in our study given that the main axis of lake–stream divergence within each watershed was consistently different in orientation from the phenotypic line of least resistance for populations within that watershed. Although Schluter’s (1996) original prediction was related to genetic (co)variances, a reasonable correspondence between the genetic and phenotypic line of least resistance is likely for our study. A first reason is that phenotypic (co)variances generally approximate underlying genetic (co)variances quite well (Cheverud, 1988; Roff, 1996; Schluter, 1996; Roff *et al.*, 1999; Badyaev & Hill, 2000; Bégin & Roff, 2004) and may sometimes even provide a more accurate estimate of the latter (Roff, 1995; Shaw *et al.*, 1995). Furthermore, in our study the phenotypic line of least resistance was particularly strongly driven by variance in gill raker length. The same trait also exhibited the highest genetic variance and heritability among five morphological traits in the lake population examined by Schluter (1996, table 1). Taken together, it is very likely that our phenotypic analysis provides insight into evolutionary bias associated with the genetic trait (co)variance structure.

Our result indicates that patterns of adaptive diversification are not easily predicted based on knowledge about trait (co)variation, a conclusion also reached in a number of other recent studies (Merilä & Björklund, 1999; Badyaev & Hill, 2000; McGuigan *et al.*, 2005; Berner & Blanckenhorn, 2006; Brakefield & Roskam, 2006). Our conclusion thus conflicts with Schluter’s (1996) report of persistent evolutionary bias to diversification in stickleback. The two studies, however, are not directly comparable because different sets of traits and different stickleback populations were considered.

Although we provide strong evidence against persistent constraints to diversification, it nevertheless remains possible that the early stages of divergence do proceed along the lines of least resistance, but that this bias is lost over time. However, as divergence among our stickleback populations has evolved in less than 12 000–15 000 years (Clague & James, 2002), any such initial constraints, if present, were lost rapidly.

Habitat-related changes in (co)variances

The third ambiguity relates to whether phenotypic (co)variances are the product of selection, rather than historical legacies (Schluter, 1996; Arnold *et al.*, 2001; McGuigan, 2006). Our analysis suggests that this might well be the case for lake–stream stickleback. In particular, when variances for gill raker length and posterior depth (and covariances between these traits) differed between habitats, they were generally lower for stream populations. Again, a contribution of phenotypic plasticity to these changes cannot be ruled out entirely but is unlikely to be substantial. For instance, even though a modest plastic response in foraging traits to limnetic vs. benthic food treatments in the laboratory was observed by Day *et al.* (1994), this result could not be reproduced in later studies (Day & McPhail, 1996; D. Berner, unpublished data). A genetic basis to the observed changes in (co)variances is therefore probable. Moreover, we will now explain how previous work on lake stickleback, combined with our stomach content data, provides a strong hypothesis for how divergent selective conditions have shaped trait (co)variances.

Lake stickleback can show substantial within-population variation in resource use along a continuum ranging from limnetic to benthic foraging modes (Schluter & McPhail, 1992; Robinson, 2000; Bolnick, 2004). This variation coincides with marked differences in foraging morphology: limnetic specialists have more numerous and longer gill rakers and shallower bodies than do benthic specialists. This individual specialization appears maintained by persistent disruptive selection owing to frequency-dependent competition for shared resources (Bolnick, 2004; Svanbäck & Bolnick, 2007). A likely consequence of individual specialization on limnetic vs. benthic prey is elevated phenotypic variance in foraging traits. Specialization should also maintain positive covariances between synergistically selected traits (e.g. raker number and raker length) and negative covariances between antagonistically selected traits (e.g. raker length and body depth).

As the stomach content data indicate, most of our lakes indeed provided both limnetic and benthic prey, and hence ample opportunity for individual specialization on these different resources. This was not true for most stream sites, which were instead characterized by more uniform benthic diets. We would therefore expect reduced individual specialization and hence lower

(co)variances in foraging-related traits in stream than lake populations. This prediction is supported by our analysis, and we therefore suggest that different selective conditions between lake and stream habitats have not only shaped foraging trait means (see above), but also their (co)variances. It is further possible that gill raker length, a key trait in individual specialization (Robinson, 2000; Bolnick, 2004), is subject to particularly strong disruptive selection in lakes, thus maintaining a particularly high variance relative to other traits. Addressing these hypotheses more directly with larger samples and information on selective conditions acting within local populations is a promising avenue for future research.

Conclusions

Stickleback inhabiting lake and stream habitats in multiple watersheds have primarily diversified along the limnetic vs. benthic foraging axis. This pattern was certainly evident when grouping populations into discrete lake and stream categories, but was made much clearer by incorporating quantitative estimates of variation in a key selective factor, and information on gene flow. In addition, patterns of lake–stream divergence within watersheds were unrelated to the phenotypic line of least resistance within those watersheds. This result adds to the chorus of studies arguing against major constraints to diversification over moderate and long time frames. Instead, it appears that (co)variances among morphological traits are themselves shaped by local selective conditions. Our analyses thus highlight the power of natural selection in shaping the morphological diversification of threespine stickleback inhabiting lakes and streams.

Acknowledgments

Field work was supported by Jean-Sébastien Moore, Kate Hudson and Dan Bolnick. Derek Roff kindly provided the jackknife code and Frédéric Guillaume, Amy Schwartz and two anonymous reviewers gave helpful comments on earlier manuscript drafts. Western Forest Products Inc. provided accommodation in the field. DB is supported by the Swiss National Science Foundation, the Janggen-Pöhn Foundation, the Roche Research Foundation and the Stiefel-Zangger Foundation. DCA is supported in part by NSF grant DEB-0445768. AH is supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada. We are most grateful to all these persons and institutions.

References

- Adams, D.C. & Collyer, M.L. 2007. Analysis of character divergence along environmental gradients and other covariates. *Evolution* **61**: 510–515.

- Aguirre, W.E., Doherty, P.K. & Bell, M.A. 2004. Genetics of lateral plate and gillraker phenotypes in a rapidly evolving population of threespine stickleback. *Behaviour* **141**: 1465–1483.
- Albert, A.Y.K., Sawaya, S., Vines, T.H., Knecht, A.K., Miller, C.T., Summers, B.R., Balabhadra, S., Kingsley, D.M. & Schluter, D. 2008. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution* **62**: 76–85.
- Arnold, S.J. 1992. Constraints on phenotypic evolution. *Am. Nat.* **140**: S85–S107.
- Arnold, S.J., Pfrender, M.E. & Jones, A.G. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* **112–113**: 9–32.
- Badyaev, A.V. & Hill, G.E. 2000. The evolution of sexual dimorphism in the house finch. I. Population divergence in morphological covariance structure. *Evolution* **54**: 1784–1794.
- Bégin, M. & Roff, D.A. 2004. From micro- to macroevolution through quantitative genetic variation: positive evidence from field crickets. *Evolution* **58**: 2287–2304.
- Bégin, M., Roff, D.A. & Debat, V. 2004. The effect of temperature and wing morphology on quantitative genetic variation in the cricket *Gryllus firmus*, with an appendix examining the statistical properties of the Jackknife-MANOVA method of matrix comparison. *J. Evol. Biol.* **17**: 1255–1267.
- Bentzen, P. & McPhail, J.D. 1984. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): specialization for alternative trophic niches in the Enos Lake species pair. *Can. J. Zool.* **62**: 2280–2286.
- Berner, D. & Blanckenhorn, W.U. 2006. Grasshopper ontogeny in relation to time constraints: adaptive divergence and stasis. *J. Anim. Ecol.* **75**: 130–139.
- Björklund, M. 1996. The importance of evolutionary constraints in ecological time scales. *Evol. Ecol.* **10**: 423–431.
- Blake, R.W. 2004. Fish functional design and swimming performance. *J. Fish Biol.* **65**: 1193–1222.
- Bolnick, D.I. 2004. Can intraspecific competition drive disruptive selection? An experimental test in natural populations of sticklebacks. *Evolution* **58**: 608–618.
- Bolnick, D.I., Caldera, E.J. & Matthews, B. 2008. Evidence for asymmetric migration load in a pair of ecologically divergent lacustrine stickleback populations. *Biol. J. Linn. Soc.* **94**: 273–287.
- Bookstein, F.L., Chernoff, B., Elder, R.L., Humphries, J.M., Smith, G.R. & Strauss, R.E. 1985. Morphometrics in evolutionary biology. *Acad. Nat. Sci. Phila. Spec. Publ.* **15**: 1–277.
- Brakefield, P.M. & Roskam, J.C. 2006. Exploring evolutionary constraints is a task for an integrative evolutionary biology. *Am. Nat.* **168**: S4–S13.
- Caldecutt, W.J. & Adams, D.C. 1998. Morphometrics of trophic osteology in the threespine stickleback, *Gasterosteus aculeatus*. *Copeia* **1998**: 827–838.
- Cheverud, J.M. 1988. A comparison of genetic and phenotypic correlations. *Evolution* **42**: 958–968.
- Clague, J.J. & James, T.S. 2002. History and isostatic effects of the last ice sheet in southern British Columbia. *Quatern. Sci. Rev.* **21**: 71–87.
- Collyer, M.L. & Adams, D.C. 2007. Analysis of two-state multivariate phenotypic change in ecological studies. *Ecology* **88**: 683–692.
- Day, T. & 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. & Schluter, D. 1994. Ecology and genetics of phenotypic plasticity: a comparison of two sticklebacks. *Evolution* **48**: 1723–1734.
- Endler, J.A. 1986. *Natural Selection in the Wild*. Princeton University, Princeton, NJ.
- Gerking, S.D. 1994. *Feeding Ecology of Fish*. Academic, San Diego, CA.
- Gross, H.P. & Anderson, J.M. 1984. Geographic variation in the gillrakers and diet of European sticklebacks, *Gasterosteus aculeatus*. *Copeia* **1984**: 87–97.
- Hagen, D.W. 1973. Inheritance of numbers of lateral plates and gill rakers in *Gasterosteus aculeatus*. *Heredity* **30**: 303–312.
- Hatfield, T. 1997. Genetic divergence in adaptive characters between sympatric species of stickleback. *Am. Nat.* **149**: 1009–1029.
- Hendry, A.P. & 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., Taylor, E.B. & 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.
- Hermida, M., Fernandez, C., Amaro, R. & San Miguel, E. 2002. Heritability and “evolvability” of meristic characters in a natural population of *Gasterosteus aculeatus*. *Can. J. Zool.* **80**: 532–541.
- Jastrebski, C.J. & Robinson, B.W. 2004. Natural selection and the evolution of replicated trophic polymorphisms in pumpkinseed sunfish (*Lepomis gibbosus*). *Evol. Ecol. Res.* **6**: 285–305.
- Jolicoeur, P. 1963. Multivariate generalization of allometry equation. *Biometrics* **19**: 497–499.
- Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* **33**: 402–416.
- Langerhans, R.B. & DeWitt, T.J. 2004. Shared and unique features of evolutionary diversification. *Am. Nat.* **164**: 335–349.
- Langerhans, R.B., Layman, C.A., Shokrollahi, A.M. & DeWitt, T.J. 2004. Predator-driven phenotypic diversification in *Gambusia affinis*. *Evolution* **58**: 2305–2318.
- Lavin, P.A. & McPhail, J.D. 1993. Parapatric lake and stream sticklebacks on northern Vancouver Island: disjunct distribution or parallel evolution? *Can. J. Zool.* **71**: 11–17.
- Losos, J.B., Jackman, T.R., Larson, A., de Queiroz, K. & Rodriguez-Schettino, L. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* **279**: 2115–2118.
- Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.
- Manly, B.F.J. 2007. *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 3rd edn. Chapman & Hall, Boca Raton, FL.
- McGuigan, K. 2006. Studying phenotypic evolution using multivariate quantitative genetics. *Mol. Ecol.* **15**: 883–896.
- McGuigan, K., Franklin, C.E., Moritz, C. & Blows, M.W. 2003. Adaptation of rainbow fish to lake and stream habitats. *Evolution* **57**: 104–118.
- McGuigan, K., Chenoweth, S.F. & Blows, M.W. 2005. Phenotypic divergence along lines of genetic variance. *Am. Nat.* **165**: 32–43.

- McKinnon, J.S., Mori, S., Blackman, B.K., David, L., Kingsley, D.M., Jamieson, L., Chou, J. & Schluter, D. 2004. Evidence for ecology's role in speciation. *Nature* **429**: 294–298.
- McPeck, M.A. 1995. Morphological evolution mediated by behavior in the damselflies of two communities. *Evolution* **49**: 749–769.
- Merilä, J. & Björklund, M. 1999. Population divergence and morphometric integration in the greenfinch (*Carduelis chloris*) – evolution against the trajectory of least resistance? *J. Evol. Biol.* **12**: 103–112.
- Moore, J.S., Gow, J.L., Taylor, E.B. & Hendry, A.P. 2007. Quantifying the constraining influence of gene flow on adaptive divergence in the lake-stream threespine stickleback system. *Evolution* **61**: 2015–2026.
- Ostbye, K., Amundsen, P.A., Bernatchez, L., Klemetsen, A., Knudsen, R., Kristoffersen, R., Naesje, T.F. & Hindar, K. 2006. Parallel evolution of ecomorphological traits in the European whitefish *Coregonus lavaretus* (L.) species complex during postglacial times. *Mol. Ecol.* **15**: 3983–4001.
- Peichel, C.L., Nereng, K.S., Oghi, K.A., Cole, B.L.E., Colosimo, P.F., Buerkle, C.A., Schluter, D. & Kingsley, D.M. 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* **414**: 901–905.
- Pennak, R.W. 1989. *Fresh-water Invertebrates of the United States*, 3rd edn. Wiley, New York.
- R Development Core Team 2006. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Reimchen, T.E., Stinson, E.M. & 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.
- Robinson, B.W. 2000. Trade offs in habitat-specific foraging efficiency and the nascent adaptive divergence of sticklebacks in lakes. *Behaviour* **137**: 865–888.
- Robinson, B.W. & Wilson, D.S. 1994. Character release and displacement in fishes – a neglected literature. *Am. Nat.* **144**: 596–627.
- Roff, D.A. 1995. The estimation of genetic correlations from phenotypic correlations – a test of Cheverud's conjecture. *Heredity* **74**: 481–490.
- Roff, D.A. 1996. The evolution of genetic correlations: an analysis of patterns. *Evolution* **50**: 1392–1403.
- Roff, D. 2000. The evolution of the *G* matrix: selection of drift? *Heredity* **84**: 135–142.
- Roff, D. 2002. Comparing *G* matrices: a MANOVA approach. *Evolution* **56**: 1286–1291.
- Roff, D.A., Mousseau, T.A. & Howard, D.J. 1999. Variation in genetic architecture of calling song among populations of *Allonemobius socius*, *A. fasciatus*, and a hybrid population: drift or selection? *Evolution* **53**: 216–224.
- Rundle, H.D., Nagel, L., Boughman, J.W. & Schluter, D. 2000. Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**: 306–308.
- Schluter, D. 1993. Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology* **74**: 699–709.
- Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**: 1766–1774.
- Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford University, Oxford.
- Schluter, D. & McPhail, J.D. 1992. Ecological character displacement and speciation in sticklebacks. *Am. Nat.* **140**: 85–108.
- Schluter, D., Clifford, E.A., Nemethy, M. & McKinnon, J.S. 2004. Parallel evolution and inheritance of quantitative traits. *Am. Nat.* **163**: 809–822.
- Shaw, F.H., Shaw, R.G., Wilkinson, G.S. & Turelli, M. 1995. Changes in genetic variances and covariances: *G* whiz! *Evolution* **49**: 1260–1267.
- Skulason, S. & Smith, T.B. 1995. Resource polymorphism in vertebrates. *Trends Ecol. Evol.* **10**: 366–370.
- Steppan, S.J., Phillipps, P.C. & Houle, D. 2002. Comparative quantitative genetics: evolution of the *G* matrix. *Trends Ecol. Evol.* **17**: 320–327.
- Svanbäck, R. & Bolnick, D.I. 2007. Intraspecific competition drives increased resource use diversity within a natural population. *Proc. R. Soc. B* **274**: 839–844.
- Thompson, C.E., Taylor, E.B. & McPhail, J.D. 1997. Parallel evolution of lake-stream pairs of threespine sticklebacks (*Gasterosteus*) inferred from mitochondrial DNA variation. *Evolution* **51**: 1955–1965.
- Thorp, J.H. & Covich, A.P. (eds) (2001) *Ecology and Classification of North American Freshwater Invertebrates*. Academic, San Diego, CA.
- Walker, J.A. 1997. Ecological morphology of lacustrine threespine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. *Biol. J. Linn. Soc.* **61**: 3–50.
- Webb, P.W. 1984. Body form, locomotion and foraging in aquatic vertebrates. *Am. Zool.* **24**: 107–120.

Received 17 February 2008; revised 30 June 2008; accepted 1 July 2008

Appendix 1

Size-standardized trait means and standard errors (in parentheses) for all stickleback populations, based on ln-transformed measurements. Sample sizes are given in Table 1.

Watershed	Habitat	Gill raker number	Gill raker length	Gill arch length	Head depth	Anterior body depth	Posterior body depth
Beaver	Lake	-0.032 (0.010)	-0.039 (0.029)	0.086 (0.011)	0.044 (0.007)	-0.036 (0.008)	-0.049 (0.010)
	Stream	-0.120 (0.012)	0.006 (0.017)	-0.122 (0.008)	0.004 (0.004)	0.030 (0.008)	0.065 (0.010)
Boot	Lake	0.127 (0.015)	0.013 (0.032)	0.100 (0.009)	0.006 (0.007)	-0.031 (0.008)	-0.071 (0.013)
	Stream	-0.070 (0.011)	-0.162 (0.023)	-0.015 (0.010)	0.037 (0.006)	0.053 (0.006)	0.073 (0.010)
Joe's	Lake	-0.018 (0.012)	0.042 (0.014)	-0.062 (0.009)	-0.009 (0.004)	0.019 (0.005)	0.005 (0.008)
	Stream	-0.116 (0.011)	-0.165 (0.017)	-0.023 (0.008)	0.045 (0.006)	0.070 (0.007)	0.055 (0.010)
McCreight	Lake	0.010 (0.013)	0.056 (0.023)	-0.050 (0.009)	-0.018 (0.007)	-0.009 (0.009)	0.019 (0.012)
	Stream	0.002 (0.014)	0.060 (0.030)	-0.059 (0.011)	-0.031 (0.008)	0.004 (0.008)	0.023 (0.010)
Misty	Lake	0.028 (0.011)	0.050 (0.028)	0.062 (0.010)	-0.009 (0.007)	-0.015 (0.007)	-0.076 (0.009)
	Stream	0.035 (0.011)	0.064 (0.022)	0.034 (0.007)	-0.025 (0.006)	-0.015 (0.007)	-0.049 (0.008)
Morton	Lake	0.123 (0.011)	0.045 (0.022)	0.043 (0.007)	0.008 (0.007)	-0.035 (0.008)	-0.046 (0.011)
	Stream	0.097 (0.012)	-0.073 (0.017)	0.063 (0.007)	0.035 (0.005)	0.002 (0.007)	-0.019 (0.007)
Pye	Lake	0.100 (0.012)	0.191 (0.017)	-0.012 (0.012)	-0.104 (0.006)	-0.066 (0.006)	0.003 (0.008)
	Stream	-0.124 (0.011)	-0.177 (0.023)	-0.019 (0.012)	0.047 (0.006)	0.066 (0.007)	0.064 (0.009)
Robert's	Lake	0.018 (0.014)	0.152 (0.022)	0.010 (0.011)	-0.033 (0.007)	-0.070 (0.006)	-0.049 (0.011)
	Stream	-0.066 (0.013)	-0.068 (0.019)	-0.039 (0.009)	0.010 (0.005)	0.032 (0.007)	0.053 (0.010)

Appendix 2

Phenotypic variances and covariances (multiplied by 1000, upper semi-matrix) and correlations (lower semi-matrix, in bold) among the six morphological traits. The values are averages across the 16 populations, with associated standard errors in parentheses.

Trait	Gill raker number	Gill raker length	Gill arch length	Head depth	Anterior body depth	Posterior body depth
Gill raker number	4.220 (0.290)	-0.478 (0.358)	0.792 (0.192)	0.085 (0.149)	-0.001 (0.066)	0.002 (0.120)
Gill raker length	-0.058 (0.041)	20.245 (2.356)	0.921 (0.579)	0.734 (0.231)	-0.013 (0.138)	-1.167 (0.356)
Gill arch length	0.202 (0.053)	0.117 (0.072)	3.337 (0.321)	1.086 (0.139)	0.111 (0.079)	-0.010 (0.128)
Head depth	0.050 (0.064)	0.163 (0.053)	0.570 (0.041)	1.108 (0.120)	0.216 (0.058)	0.012 (0.087)
Anterior body depth	0.009 (0.043)	-0.010 (0.042)	0.092 (0.054)	0.267 (0.062)	0.623 (0.044)	0.336 (0.052)
Posterior body depth	0.005 (0.051)	-0.197 (0.053)	0.020 (0.067)	0.001 (0.057)	0.370 (0.055)	1.495 (0.108)