OXFORD

Symposium Article

Repeatability of Adaptive Radiation Depends on Spatial Scale: Regional Versus Global Replicates of Stickleback in Lake Versus Stream Habitats

Antoine Paccard*, Dieta Hanson*, Yoel E. Stuart, Frank A. von Hippel, Martin Kalbe, Tom Klepaker, Skúli Skúlason, Bjarni K. Kristjánsson, Daniel I. Bolnick, Andrew P. Hendry, and Rowan D. H. Barrett

From the Redpath Museum and Department of Biology, McGill University, Montreal, QC H3A 0C4, Canada (Paccard, Hanson, Hendry, and Barrett); Department of Integrative Biology, University of Texas at Austin, Austin, TX 78712 (Stuart); Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ (von Hippel); Department of Evolutionary Ecology, Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany (Kalbe); University of Bergen, Department of Biology, PO Box 7800, 5020 Bergen, Norway (Klepaker); Department of Aquaculture and Fish Biology, Hólar University College, Háeyri 1, 551 Sauðárkrókur, Iceland (Skúlason and Kristjánsson); Ecology and Evolutionary Biology, University of Connecticut, 75 N. Eagleville Road, Unit 3043, Storrs, CT 06269-3043 (Bolnick).

Address correspondence to Antoine Paccard at the address above, or e-mail: antoine.paccard@mcgill.ca.

*Both authors contributed equally to the work.

Abstract

The repeatability of adaptive radiation is expected to be scale-dependent, with determinism decreasing as greater spatial separation among "replicates" leads to their increased genetic and ecological independence. Threespine stickleback (*Gasterosteus aculeatus*) provide an opportunity to test whether this expectation holds for the early stages of adaptive radiation—their diversification in freshwater ecosystems has been replicated many times. To better understand the repeatability of that adaptive radiation, we examined the influence of geographic scale on levels of parallel evolution by quantifying phenotypic and genetic divergence between lake and stream stickleback pairs sampled at regional (Vancouver Island) and global (North America and Europe) scales. We measured phenotypes known to show lake-stream divergence and used reduced representation genome-wide sequencing to estimate genetic divergence. We assessed the scale dependence of parallel evolution by comparing effect sizes from multivariate models and also the direction and magnitude of lake-stream divergence vectors. At the phenotypic level, parallelism was greater at the regional than the global scale. At the genetic level, putative selected loci showed greater lake-stream parallelism at the regional than the global scale. Generally, the level of parallel evolution was low at both scales, except for some key univariate traits. Divergence vectors were often orthogonal, highlighting possible ecological and genetic constraints on parallel evolution at both scales. Overall, our results confirm that the repeatability of adaptive radiation decreases at increasing spatial scales. We suggest that greater environmental heterogeneity at larger scales imposes

1

different selection regimes, thus generating lower repeatability of adaptive radiation at larger spatial scales.

An unanswered question surrounding adaptive radiation is the extent to which the process is deterministic versus stochastic. That is, if the "tape of life" were rewound and started over multiple independent times, to what extent would a similar set of species evolve when presented with a similar set of environmental niches (Gould 1990; Blount et al. 2018)? This thought experiment cannot be made real outside of the laboratory (Craig MacLean 2005), and so evolutionary biologists have tended to substitute geographical space for evolutionary time by asking: To what extent do related species evolve in parallel in similar sets of environments that are spatially isolated and evolutionarily independent? A difficulty in addressing this question is that relatively few adaptive radiations are replicated in space; many are instead unique, such as Darwin's finches in the Galapagos Islands (Grant 1981) and the Hawaii honeycreepers (James 2004); or are relatively rare, such as the radiations of cichlids in the 3 Great Lakes of Africa (Kocher 2004). A few adaptive radiations are sufficiently replicated to make inferences about repeatability; and these cases generally indicate that adaptive radiations have substantial repeatable components, such as the same ecomorphs of *Anolis* lizards evolving on multiple islands in the Greater Antilles (Losos 2009) or the same ecomorphs of spiders evolving on multiple Hawaiian Islands (Gillespie 2004; Gillespie et al. 2018). Yet these same radiations often show substantial nondeterministic components, such as "missing" ecomorphs of lizards or spiders on some islands or strong contingency depending on physical properties of the system (Brawand et al. 2014).

A remaining complication is that repeated adaptive radiations with a given group of organisms are not entirely independent—genetically or ecologically. With respect to genetic nonindependence, postcolonization gene flow occurs among islands for *Anolis* lizards (Losos 2009) and Hawaiian spiders (Gillespie 2004), and independent radiations of African cichlids appear to be using some of the same ancestral polymorphisms (Loh et al. 2012). With respect to ecological nonindependence, adaptive radiations are often replicated on only regional spatial scales (e.g., among islands in Caribbean or Hawaii), where overall ecological conditions are relatively similar. Hence, the resulting genetic or ecological nonindependence could be primary factors shaping deterministic responses. Stated in the same counterfactual sense as "replaying the tape of life," if a few ancestral species had colonized very different geographic locations, to what extent would a similar adaptive radiation take place? Intuition would suggest that such "replicate" adaptive radiations separated by large geographical scale would yield lower repeatability than would replicate adaptive radiations at a more regional scale—as a result of greater genetic and ecological independence.

The multiscale adaptive radiation thought experiment described above is not inevitably counterfactual. Threespine stickleback (*Gasterosteus aculeatus*) are found throughout much of the northern hemisphere, where they show many replicate adaptive radiations in freshwater ecosystems arising from independent colonization events from a common marine ancestor (Colosimo et al. 2005). Although the resulting populations are all placed under the same Latin binomial, many of these populations show large genetic differences that

are primarily shaped by adaptation to different environments as opposed to interactions between close relatives, as in the Hawaiian spiders or Anolis lizards (Gillespie 2004; Losos 2009; Schluter et al. 2010; Peichel and Marques 2017; Gillespie et al. 2018). For example, the F_{ST} between parapatric lake-stream stickleback is greater than 0.10 in each of many watersheds colonized postglacially by marine stickleback (Kaeuffer et al. 2012; Roesti et al. 2012; Stuart et al. 2017). Hence, different freshwater populations of stickleback can be thought of as representing the early stages of adaptive radiation (Schluter and McPhail 1992; Schluter 1996; Taylor and McPhail 1999). With this stickleback system, one can compare regional replicate adaptive radiations (stickleback in similar habitats in different watersheds in the same geographical area) with global replicate adaptive radiations (stickleback in similar habitats in different watersheds in distant geographical areas). We therefore use this system to ask whether genetic and phenotypic divergence is more repeatable (similar magnitude and direction across replicates) at regional than at global scales.

Study System

Parapatric pairs of lake and stream stickleback have been used frequently to study the repeatability/predictability/parallelism/convergence of adaptive divergence (e.g., Reimchen et al. 1985; Lavin and McPhail 1993; Deagle et al. 1996; Thompson et al. 1997; Reusch et al. 2001; Hendry and Taylor 2004; Aguirre 2009; Lucek et al. 2010, 2013; Eizaguirre et al. 2011; Kaeuffer et al. 2012; Roesti et al. 2012; Ravinet et al. 2013; Stuart et al. 2017; Weber et al. 2017). In this study system, stickleback in a lake and its adjoining inlet or outlet stream (hereafter, lake-stream population pairs) often evolve similar patterns of phenotypic divergence in independent watersheds. Most notably, lake stickleback generally evolve shallower bodies and more gill rakers than do stream stickleback, a pattern driven at least in part by consistent differences in diet between habitats (Berner et al. 2008; Kaeuffer et al. 2012). These phenotypic differences tend to be genetically based (Sharpe et al. 2008; Berner et al. 2011; Hendry et al. 2011; Moser et al. 2016; Oke et al. 2016) and are likely adaptive, given their close association with variation in prey availability (Berner et al. 2008; Kaeuffer et al. 2012). However, nonparallel aspects of divergence are known where some traits show low levels of parallelism (Stuart et al. 2017).

Most analyses of parallel evolution in lake-stream stickleback have taken place on regional geographic scales, where we expect generally similar selective regimes and genetic backgrounds: northern Vancouver Island (Lavin and McPhail 1993; Hendry and Taylor 2004; Berner et al. 2008, 2009; Kaeuffer et al. 2012; Stuart et al. 2017), Haida Gwaii (Reimchen et al. 1985), Ireland (Ravinet et al. 2013), and northern Germany (Reusch et al. 2001). Several studies, however, have analyzed sets of lake-stream pairs within and between regions: Vancouver Island versus Switzerland (Berner et al. 2010), Vancouver Island versus artificial habitats in California (Hendry et al. 2013), Iceland versus Switzerland (Lucek et al. 2014b), and Vancouver Island versus Switzerland versus Ireland (Lucek et al. 2013). These studies revealed important morphological and genetic

variation in the direction and magnitude of lake-stream divergence between regions, but they did not formally compare the extent of parallel evolution across different geographic scales. In the present study, we conduct such a comparison by assessing parallel evolution of lake-stream divergence at 2 very different scales: 6 pairs within Vancouver Island (hereafter, "regional" comparisons) and 6 pairs from worldwide localities that encompass the majority of the threespine stickleback range (hereafter, "global" comparisons).

We estimated phenotypic (non)parallelism by first comparing the effect sizes of multiple univariate and multivariate models. We then evaluated the degree of phenotypic and genetic (non)parallelism by calculating multivariate phenotypic and genetic (neutral and outlier) lake-to-stream divergence vectors. Then, we quantified the differences among vectors in direction (i.e., the angle between any 2 divergence vectors) and in magnitude (i.e., the difference in length between any 2 divergence vectors) (Collyer and Adams 2007; Stuart et al. 2017). Small angles and small differences in length imply highly parallel evolution. With these summary statistics for (non)parallel evolution, we asked the following 2 questions:

1. What is the relative extent of phenotypic and genetic (non)parallelism at global versus regional scales?
2. What are the potential ecological and genetic drivers of phenotypic and genetic (non)parallelism at each scale?

Materials and Methods

Field Collections and Environmental Data

We used minnow traps to collect 39–40 lake and 40 stream stickleback, between 21 May and 28 June 2013, from each of 6 independent population pairs on Vancouver Island (the "regional" samples: Supplementary Table S1A, Figure S1A). We also used minnow traps or nets to collect 40 lake and 40 stream stickleback, between 10 April and 8 August 2014, from each of 6 other population pairs in North America and Europe (global samples: Supplementary Table S1B, Figure S1B). In all collections, we trapped the lake fish at least 100 m away from the junction with any stream (apart from Lake Constance, where the lake fish were captured at the mouth of the stream), and we trapped stream fish at least 100 m away from the junction with any lake. We did not retain any fish younger than 1 year or any obviously gravid females.

We euthanized the fish with an overdose of MS-222 fish anesthetic (Argent Chemical Laboratories, Redmond, WA), and removed the right pectoral fin, which was preserved in 95% ethanol for genetic analysis. We then fixed the fish in 10% neutral buffered formalin (VWR, Radnor, PA). To facilitate gill raker counts, we stained the fish using alizarin red dye. To do so, we first soaked the fish in water for 24 h, then stained the fish in a solution of alizarin red and 0.5% KOH for 24 h, and then performed a second soak in water for 24 h to remove excess stain. We then stored the fish in 40% isopropyl alcohol until they were processed further (see below), approximately 4 weeks later.

To characterize the environment at each site, we measured several habitat variables. Using the Google Earth Pro V7.3 software, we measured total lake area (m²), lake perimeter (m), and mean lake-stream elevation for each pair (m). Using measuring sticks and a tape meter, we also measured stream depth (cm), lake depth (cm), and stream width (cm) at each trap as well as stream flow (cm/s) using a flow meter.

Morphological Measurements

In the laboratory, we photographed the left and ventral surfaces of each fish against a grid marked at 1-cm intervals. The grid was placed at a constant distance from the lens of a Canon G11 digital camera, which was positioned in the same plane (using a surface level) as the grid. Before the photograph was taken, we removed the left pectoral fin of each fish and pinned it at maximum extension onto the same grid (for measurements of fin area, perimeter, and the length of fin rays). We also inserted small pins into the fish to help indicate otherwise cryptic homologous anatomical points for subsequent geometric morphometrics (Berner et al. 2009). After photographing, we measured several morphological traits (described below) with a focus on body shape and gill raker traits because repeated patterns of phenotypic divergence between lake and stream fish often involve these traits (Berner et al. 2008, 2011; Sharpe et al. 2008; Hendry et al. 2011; Kaeuffer et al. 2012; Moser et al. 2016; Oke et al. 2016). In addition, we dissected each fish to determine sex by gonad inspection.

From the photographs, we measured 25 linear trait distances (Supplementary Table S2), as well as the area and perimeter of the pectoral fin using the image processing software *Fiji* (Schindelin et al. 2012). We standardized these 27 univariate measurements to a common standard body length with the allometric formula $M_S = M_0 (L_S/L_0)^b$, where M_S is the standardized trait measurement, M_0 is the unstandardized trait measurement, L_S is the overall mean body size of all fish in a given analysis, and L_0 is the standard length of the individual (Elliott et al. 1994; Lleonart et al. 2000). Standard length is defined as the distance from the anterior tip of the upper jaw to the posterior end of the hyperal plate. The exponent b was calculated as the common within-group slope (Reist 1986) from a linear mixed-effect model regressing $\log_{10}(M_0)$ on $\log_{10}(L_0)$ with pair as the random factor.

We counted the number of gill rakers on the first right gill arch in situ, then removed the right gill arch and photographed it under a dissecting scope using an ocular micrometer and a Canon EOS Rebel T5 digital camera. To quantify gill raker length, we measured the curved distance along the centre of the 3 longest gill rakers. We then calculated raker spacing as $5/l$, where l is the mm distance along the gill arch spanned by the 5 longest rakers. These 4 measurements (lengths of the 3 longest gill rakers, and gill raker spacing) were standardized to a common standard body length as described above.

To quantify overall geometric morphometric shape differences, we placed 19 homologous landmarks on the lateral photographs using the *TPSdig* software (Rohlf 2006) (Supplementary Figure S2). We then superimposed the 19 landmarks using a generalized Procrustes analysis (GPA) with *R* package *geomorph* (Adams and Otárola Castillo 2013). We performed this analysis for the global samples and the regional samples combined so that all fish would be positioned in the same shape space. These alignments resulted in 38 Procrustes residuals that described the shape differences between specimens unrelated to scale, rotation, or translation. The 38 principle components (PCs) derived from the 38 Procrustes residuals were then allometrically adjusted for centroid size and body depth using the above common within-slope approach (Reist 1986; Lleonart et al. 2000; Rolshausen et al. 2015). The first PC was found to describe shape differences due to vertical bending of the fish produced during storage. We therefore excluded this artifactual PC from further analyses.

DNA Extraction and Restriction Site-Associated DNA Sequencing Library Construction

DNA was extracted from stickleback fin clips using the Wizard® SV Genomic Purification kit (Promega® Corp., Madison, WI) and quantified using Picogreen® ds DNA assay (Thermo Fisher Scientific,

Waltham, MA) on an Infinite® 200 Nanoquant (Tecan Group Ltd, Männedorf, Switzerland). All samples were normalized to a dsDNA concentration of 15 ng/µL, requantified, and pooled according to sampling location. Thus, we created 24 pools of 40 individuals each, consisting of 12 pools for each geographic scale (at each scale, 6 lake and 6 stream pools).

ddRAD PoolSeq libraries were prepared following a modified version of the Peterson et al. (2012) protocol. Briefly, pooled samples were first digested using the NLAIII and MluCI enzymes (New England Biolabs Inc., Frankfurt, Germany), individually barcoded using unique P1 flex adapters (Peterson et al. 2012), and pooled together into a single library. Fragments of 371–416 bp were extracted using a Pippen Prep® 2% MarkerB 100–600 bp cassette (Sage Science Inc., Beverly, MA). In post-size selection, we isolated fragments including at least 1 biotin-tagged P2 adapter, using streptavidin-couple beads (Dynabeads® M-270, Thermo Fisher Scientific). The library was then PCR-amplified in 5 separate aliquots (12 cycles, Phusion High Fidelity® PCR kit, New England Biolabs Inc.). We used Agencourt® AMPure XP beads (Beckman Coulter, Indianapolis, IN) to clean samples after each enzymatic reaction.

Sequencing and ddRAD Data Processing

Barcoded ddRAD samples were sequenced on 4 lanes of an Illumina Hi-Seq 2500 for 100-bp paired end (~150 M reads total), and demultiplexed using *STACKS* (Catchen et al. 2013) v. 1.13 (*process_radtags* -P, -p, -r, -i, --inline_index, --disable_rad_check). Resulting FASTA files were trimmed using *PoPoolation* (Kofler et al. 2011a) to eliminate low quality regions (--min-length 50 --quality-threshold 20). Reads from all 24 pools were aligned to the threespine stickleback genome (release-084, Ensembl.org) separately using *BWA mem* v. 0.7.13 (Li and Durbin 2009). We then used *SAMtools* (Li et al. 2009) v. 1.3.1 to convert the resulting sam files to sorted bam format, keeping only reads with mapping quality above 20 (*samtools view* -q 20). This led to an average of ~53 M reads per pool. A pileup file was then generated using *SAMtools* v. 1.3.1 (*samtools mpileup* -B), and filtered for completeness by keeping reads with a minimum coverage of 5. We then converted the pileup file to a sync file using *PoPoolation2* (Kofler et al. 2011b) for further downstream analysis.

Statistical Analysis

All statistical analyses were performed in *R* 3.2.2 (R Core Team 2015).

Variation in Environmental Variables
To test whether variation in habitat characteristics was greater at the global than regional scale, we compared within-scale variation for each environmental parameter using Bartlett's tests (Bartlett 1937). In addition, we calculated levels of among-lake and among-stream variation for each environmental variable at each scale separately. To calculate among-lake and among-stream variation at each scale, we computed ANOVAs at each scale separately with each environmental variable as dependent variable and watershed as predictor. Among-lake or among-stream variation was calculated by dividing the pair term sum of squares by the total number of pairs in each scale respectively.

Univariate and Multivariate Phenotypic Analysis
To reduce the large number of univariate traits, we first performed a principal component analysis (PCA) on the size-standardized trait values for all samples (global and regional) combined. To consider

phenotypic differences between lake and stream stickleback across pairs for the global and regional sample sets separately, we then used multivariate analysis of covariance (MANCOVA) models, with Wilks's lambda (λ) as the test statistic. We tested for differences in specific trait sets for the regional and global samples separately using 3 different MANCOVA models, with dependent variables as: 1) all 27 (size-standardized) univariate measurements, 2) the 5 gill raker traits only, and 3) the PCs derived from the 38 Procrustes residuals.

To understand the contribution to parallelism of a subset of individual traits likely important to lake-stream divergence (body depth, gill raker number, gill raker spacing, and mean length of the 3 longest gill rakers), we analyzed each of these traits separately using an analysis of covariance (ANCOVA) at each regional and global scale. The independent variables in every model were the fixed effects of sex, pair, and habitat (lake vs. stream), along with all 2-way interactions. Centroid size was also included as a covariate (along with its 2-way interactions) to control for any remaining size effects.

Estimation of Genome-Wide Population Differentiation
We evaluated the genetic relationship within pairs by estimating the fixation index (F_{ST}) for each lake-stream pairwise comparison using *PoPoolation2* (Kofler et al. 2011b), applying a number of stringent criteria (see below) to define genomic sites for analysis across the entire genome. The accuracy of allele frequency estimation of pooled individuals is highly dependent on sequence coverage (Schlötterer et al. 2014). Therefore, to reduce stochastic error and increase the accuracy of F_{ST} estimates, for each pairwise combination we used a minimum minor allele count of 2 across all pools, a high sequence coverage filter (between 5 and 500 within each pool), and a large 100-kb nonoverlapping sliding window. This window size led to mean of 146.74 SNPs per window (±90.29, median = 123). We visualized population structure by building neighbor-joining trees for both scales separately based on average pairwise F_{ST} values.

Evaluation of (Non)parallelism at Both Geographic Scales
We used 2 different approaches to estimate the extent of parallel evolution at the 2 geographic scales.

First, we compared the effect sizes of the habitat and habitat-by-pair interaction terms from the MANCOVA and ANCOVA analyses of phenotype (partial η^2 and η^2 values, respectively) (Oke et al. 2016; Stuart et al. 2017). A large effect size for the habitat term suggests consistent lake versus stream differences shared across pairs, that is, parallel evolution. A large effect size for the pair-by-habitat interaction term, on the other hand, suggests idiosyncratic lake-versus-stream differences across pairs, that is, nonparallel evolution. Thus, if the ratio of the habitat to the interaction effect size is greater in the pairs sampled at a regional scale as compared to the pairs sampled at a global scale, we infer that lake-stream divergence is more parallel at the regional scale.

Second, we calculated lake-to-stream divergence vectors for our phenotypic and genetic data. Each vector connects the multivariate lake mean to the multivariate stream mean and quantifies the magnitude and direction of lake-stream divergence for each population pair (Adams and Collyer 2009). Then, to describe the extent to which vectors diverged in parallel across pairs, we calculated: θ, the angle between any 2 vectors; L, the length of the vector for each pair; and ΔL, the difference in length between any 2 vectors. Strict parallel divergence between pairs would correspond to $\theta = 0$ and $\Delta L = 0$. Vector analysis for the phenotypic

and genetic data sets follows Stuart et al. (2017) and is briefly described below:

(i) Phenotypic: For each pair, for each trait (Supplementary Table S2), we ran a *t*-test comparing the distribution of a trait in a lake to that in its adjoining stream. Concatenating each *t*-statistic resulted in a $1 \times N$ vector of scale-standardized lake-stream differences, thereby constituting our lake-stream divergence vector. We then calculated angles between phenotypic divergence vectors (θ_p) as the arc-cosine of the Pearson correlation for each vector pair. L_p was the multi-variate Euclidian length of each vector, and ΔL_p was the difference in length between any 2 divergence vectors. We tested for differences in θ_p and ΔL_p between global and regional scales using a 2-tailed *t*-test.

(ii) Genetic: Using the sync file of allele frequencies extracted with *PoPoolation*, we first calculated the frequency of the minor allele at each SNP. We then used PCA (*R prcomp*) to reduce our large SNP data set into a smaller number of quantitative axes of genetic differentiation. We saved populations' scores of all the resulting 24 axes. We calculated the centroid for each lake and each stream, for these 24 axes, to obtain vectors in genotypic space. Thus, each row of the resulting matrix represents the estimate of genetic divergence between the 2 habitats of the same pair, thereby constituting the genetic lake-stream divergence vector for each pair. First, we performed this analysis on a set of "outlier" SNPs whose lake-stream F_{ST} values fell within the top 5% of estimated values for each of the 12 lake-stream pairs, thus representing markers putatively linked to selected loci (175 302 outlier loci). To test for differences in genetic divergence between scales, we performed the genetic vector analysis as described above and across both scales to obtain measures of $\theta_{OUTLIERS}$ and $\Delta L_{OUTLIERS}$. We tested for differences in $\theta_{OUTLIERS}$ and $\Delta L_{OUTLIERS}$ between global and regional scales using a 2-sided *t*-test. Second, we calculated θ, L, ΔL by excluding these outlier SNPs to obtain a set of markers likely to represent neutral population genetic processes (327 291 neutral loci—hereafter θ_G, L_G, and ΔL_G). We tested for differences in θ_G and ΔL_G between global and regional scales using a 2-sided *t*-test. Values of θ_G, L_G, and ΔL_G were used further to investigate potential constraints on divergence (see below). Finally, we considered whether neutral processes could be responsible for phenotypic divergence by testing for a linear relationship between θ_p values and averaged between-watershed neutral F_{ST} at each scale separately using linear models with F_{ST} as predictor and θ_p as response variable.

Third, we verified whether watersheds share the same genetic architecture underlying adaptive traits by calculating the proportion of shared outlier loci across population pairs. At both geographic scales, we counted the total number of outlier loci and calculated the proportion of shared outlier loci over one or multiple watersheds.

Finally, because it is possible that some environmental variables could be confounded with geographic scale (e.g., the lakes and streams that we sampled at the global scale were on average larger than those sampled at the regional scale), we tested the effect sizes of our measured environmental variables relative to geographic scale in models where both forms of information acted as predictors for phenotypic or genetic divergence.

Results

Variation in Environmental Variables

Overall, environmental variation is higher at the global scale. For all environmental variables (in both lake and stream habitats), variation among watersheds was significantly greater at the global than regional scale (Table 1, Supplementary Figure S3). Similarly, among-lake and among-stream variation was greater at the global than at the regional scale (Supplementary Figure S4).

Univariate and Multivariate Phenotypic Analysis

Consistent with previous work, we observed variable levels of (non) parallelism in univariate and multivariate traits. The first and second PCs for the 27 univariate measurements explained 24.7% and 17.5% of the total variance, respectively, and every trait except the length of the dorsal fin loaded negatively on the first axis. The within-pair difference between lake and stream fish on PC1 was always in the same direction, indicating that any habitat-by-pair interaction on PC1 was due to differences in the magnitude of divergence, rather than the direction of divergence (Figure 1). The MANCOVA models revealed significant effects of habitat, pair, and their interaction on all sets of traits at both regional and global scales (Supplementary Table S3).

ANCOVAs for individual traits revealed that the effects of habitat, pair, and their interaction were always significant at both the regional and global scales (Supplementary Table S4). Specifically, stream fish were generally deeper bodied than lake fish for the regional pairs, often with similar magnitudes of difference. However, in the global pairs, stream fish were sometimes not deeper bodied than their lake counterparts (Figure 2A,B). Lake fish also generally had more and longer gill rakers than stream fish (Figure 2C–F), although this difference was more pronounced at the regional scale. Gill raker spacing was greater in stream fish at the regional scale but not at the global scale (Figure 2G,H).

Table 1. Within-scale mean and standard deviation in environmental variables

	Mean		Standard deviation		Bartlett test		
	Regional	Global	Regional	Global	K^2	df	*P*-value
Lake area (m²)	8.50^{e+05}	1.13^{e+08}	7.16^{e+05}	2.25^{e+08}	45.98	1	**<0.001**
Lake perimeter (m)	6.19^{e+03}	7.19^{e+04}	5.84^{e+03}	1.00^{e+05}	19.56	1	**<0.001**
Lake-stream elevation (m)	106.33	100.17	72.80	159.75	2.56	1	0.109
Lake depth (cm)	14.02	33.25	12.93	35.96	3.84	1	0.004
Stream depth (cm)	73.14	75.37	738.62	4438.49	3.25	1	0.070
Stream width (cm)	505.62	2016.57	122.96	2883.69	22.39	1	**<0.001**
Stream flow (cm/s)	3.72	131.60	5.16	130.22	23.06	1	**<0.001**

Results from Bartlett's tests of homogeneity of variances are presented with *K*-squared (K^2), degrees of freedom (df), and *P*-values (*P*-value). *P*-values < 0.05 are in bold.
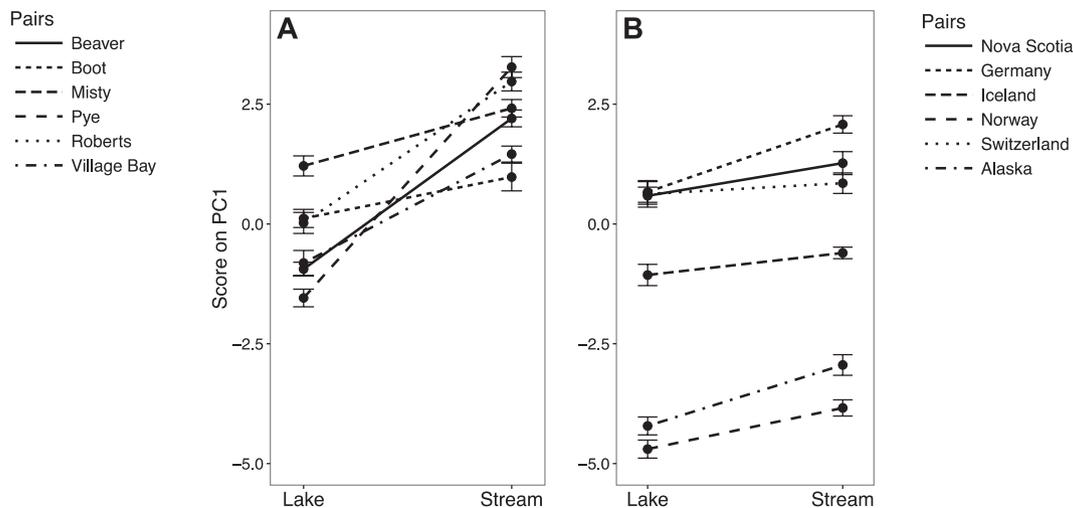
Figure 1. Mean and standard error of score on PC1 of the 27 univariate measurements for regional (**A**) and global (**B**) lake-stream pairs. The direction of phenotypic divergence is consistent but the magnitude of phenotypic divergence varies.

Evaluation of (Non)parallelism at Both Geographic Scales

Differences in Effect Sizes of Multivariate Models

Significant differences in effect sizes from the MANCOVAs and ANCOVAs revealed that phenotypic parallelism was greater at the regional than at the global scale. Indeed, the ratio of the habitat term to the habitat by pair interaction term was always greater at the regional than at the global scale—across all models. In the MANCOVA models, the averaged ratio was 3.2 at the regional scale and 1.5 at the global scale. In the ANCOVA models, the averaged ratio was 6.7 at the regional scale and 0.3 at the global scale (Table 2, Fig. 3; see Supplementary Tables S3 and S4 for more detailed results about these models). This finding suggests that lake-stream divergence in phenotypic traits was more parallel at the regional than global scale.

Vector Analysis

F_{ST} extracted from the sliding window analysis revealed that lake-stream populations in each pair were each other's sister taxon, regardless of the geographic scale of comparison (Supplementary Table S5, Figure S5). This pattern is consistent with evolutionarily independent lake-stream divergence in each pair, although ongoing gene flow could also contribute to this pattern.

Overall, we found phenotypic divergence to be more parallel at the regional than global scale. Multivariate vector analysis of the phenotypic data revealed a continuum of lake-stream parallelism at both scales (Supplementary Tables S6 and S7). The mean angle between any 2 phenotypic vectors (θ_P) was 61.63° ± 12.24 SD at the regional scale and 90.17° ± 15.98 SD at the global scale (Figure 4A; $t = 5.49$, degrees of freedom [df] = 26.23, $P = 8.97e^{-06}$). In addition, phenotypic vector lengths, which measure the magnitude (but not direction) of lake-stream divergence, also exhibit greater similarity among regional pairs than among global pairs. Pairwise comparisons of vector lengths (ΔL_P) averaged –2.85 ± 10.57 SD locally, and 6.52 ± 7.98 SD globally (Figure 4C; $t = 3.16$, df = 14.00, $P = 0.007$).

At the genetic level, divergence was significantly more parallel at the regional than global scale, but only for outlier loci: The mean angle between any 2 outlier genetic vectors ($\theta_{OUTLIERS}$) was 85.51°

± 5.81 SD at the regional scale and at 89.15° ± 3.83 SD the global scale (Figure 4B; $t = 2.02$, df = 24.21, $P = 0.04$; Supplementary Table S8). The mean angle between any 2 neutral genetic vectors (θ_G) was 89.32° ± 2.72 SD at the regional scale and 90.03° ± 1.81 SD at the global scale ($t = 0.85$, df = 24.36, $P = 0.41$). Genetic vector lengths were significantly smaller at the regional than global scale at both outlier and neutral loci. However, at outlier loci, genetic vector length differences were primarily driven by 1 population (Norway, Supplementary Table S9D). Outlier vector lengths ($\Delta L_{OUTLIERS}$) differed by a mean of 1.49 ± 4.12 SD and 10.29 ± 13.05 SD at the regional and global scales, respectively (Figure 4D; $t = 2.49$, df = 16.75, $P = 0.02$). Neutral genetic vector lengths (ΔL_G) differed by a mean of 0.46 ± 1.54 SD and 6.12 ± 6.77 SD at the regional and global scales, respectively ($t = 3.16$, df = 15.44, $P = 0.006$). Finally, θ_P and averaged between-watershed neutral F_{ST} were not significantly associated at either scale (regional: $t = –0.88$, df = 13, $P = 0.39$, $R^2 = –0.24$; global: $t = 1.12$, df = 13, $P = 0.28$, $R^2 = 0.29$).

If different watersheds share the same genetic architecture underlying adaptive traits, we would expect outlier loci under selection to overlap across population pairs when lake-stream divergence proceeds in parallel. However, 63% of the outlier loci were unique to a specific watershed across all watersheds, whereas 61% and 77% of outliers were unique to a specific watershed at the regional and global scales, respectively (21% and 4% were shared across a minimum of 2 watersheds at the regional scale and global scale, respectively). As an overall summary, comparisons of vector pairs showed less parallelism at the global scale than at the regional scale, at both the phenotypic and genetic outlier levels; for neutral genetic loci, (non)parallelism was similar at both scales. Importantly, perfect parallel divergence is nonexistent at both phenotypic and genetic levels and at both geographic scales.

Finally, and with only a few exceptions (lake depth and stream depth with ΔL_P, stream flow with θ_P), environmental variables were not associated with divergence metrics (Supplementary Table S10). Moreover, in only 1 case (lake depth and ΔL_P) was the effect size of an environmental variable greater than that of geographic scale. Thus, the effect of geographic scale on divergence between lake and stream pairs appears to be mostly independent of any consistent environmental differences between the scales.
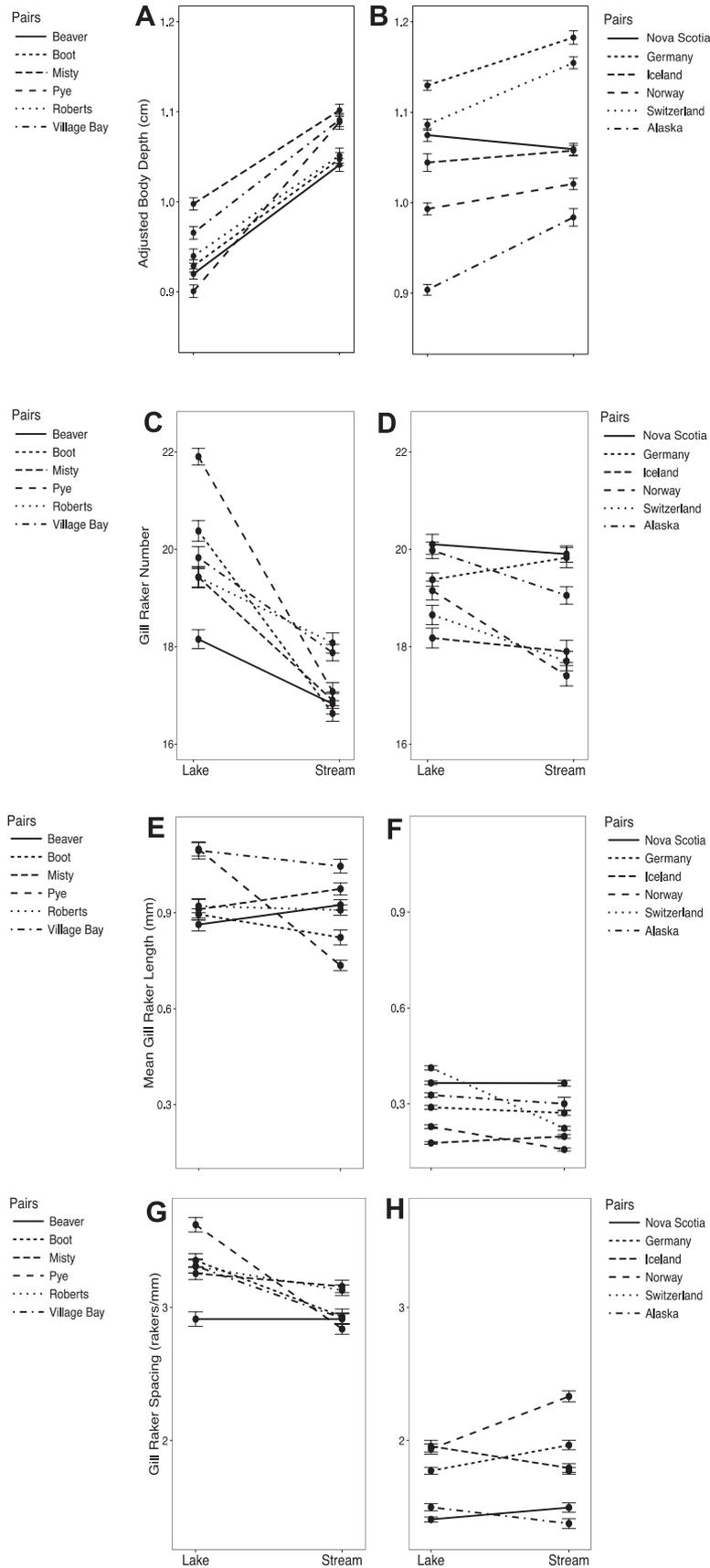
Figure 2. Mean and standard error of body depth (**A**, **B**), number of gill rakers (**C**, **D**), mean gill raker length (**E**, **F**), and gill raker spacing (**G**, **H**) for regional (left panels) and global (right panels) lake-stream pairs.

Table 2. Results from MANCOVAs on the multivariate traits (A) and ANCOVAs on single traits (B) at both regional and global scales

(A) MANCOVAs

	Regional							Global						
	Hab.	Lo. CI	Up. CI	Hab. × Pair	Lo. CI	Up. CI	Ratio	Hab.	Lo. CI	Up. CI	Hab. × Pair	Lo. CI	Up. CI	Ratio
Univariate	0.81	0.79	0.85	0.35	0.34	0.46	2.30	0.41	0.30	0.89	0.23	0.22	0.35	1.75
Gill rakers	0.45	0.41	0.65	0.09	0.06	0.19	5.04	0.03	0.18	0.05	0.03	0.02	0.11	0.95
Body shape	0.85	0.85	0.88	0.38	0.34	0.46	2.25	0.34	0.03	0.41	0.19	0.18	0.25	1.82

(B) ANCOVAs

	Regional							Global						
	Hab.	Lo. CI	Up. CI	Hab. × Pair	Lo. CI	Up. CI	Ratio	Hab.	Lo. CI	Up. CI	Hab. × Pair	Lo. CI	Up. CI	Ratio
Body depth	0.48	0.39	0.72	0.02	0.02	0.11	21.20	0.02	0.01	0.21	0.03	0.02	0.18	0.54
Gill raker number	0.29	0.12	0.58	0.09	0.05	0.14	3.39	0.01	0.01	0.11	0.03	0.02	0.10	0.54
Mean gill raker length	0.03	0.02	0.11	0.10	0.10	0.34	0.22	0.00	0.00	0.29	0.04	0.04	0.14	0.03
Gill raker spacing	0.15	0.10	0.27	0.08	0.07	0.20	1.35	0.00	0.00	0.25	0.02	0.02	0.12	0.00

Effect sizes (partial η^2) and their lower and upper 95% confidence limits (Lo. CI; Up. CI) of the habitat (Hab.) and habitat by pair interaction (Hab. × Pair) as well as their ratios (Ratio: Hab./Hab. × Pair) are reported.

Discussion

We assessed the influence of geographic scale on the repeatability of adaptive radiation by quantifying levels of phenotypic and genetic parallelism in lake-stream population pairs of threespine stickleback. In agreement with previous work conducted over restricted spatial scales (Lavin and McPhail 1993; Hendry and Taylor 2004; Berner et al. 2008, 2009; Kaeuffer et al. 2012; Stuart et al. 2017), we found a continuum of phenotypic and genetic parallelism, with lake-stream divergence being more pronounced for some traits, and in some population pairs, than in others. At the phenotypic level, parallelism was especially notable in body depth and gill raker number—also consistent with previous work. When comparing between geographic scales, parallelism in phenotypes and genetic outliers was greater at the regional scale (Vancouver Island) than at the global scale (Europe and both sides of North America). In contrast, parallelism was similar across scales for neutral genetic loci. We suggest that the observed disparities in phenotypic and genetic parallelism between scales are likely the result of differences in environmental variation, selection pressures, and genetic architecture of traits between scales. We develop these ideas further in the text below.

Phenotypic (Non)parallelism

Univariate, multivariate, and vector analyses all revealed that phenotypic parallelism was greater at the regional than the global scale (Figure 4A,C). This difference between scales was especially pronounced in specific morphological traits, such as body depth and gill raker length, that are known to have a strong heritable basis and that have been previously implicated in adaptive parallel evolution (Berner et al. 2008, 2011; Sharpe et al. 2008; Hendry et al. 2011; Kaeuffer et al. 2012; Miller et al. 2014; Moser et al. 2016; Oke et al. 2016).

Lake-stream body depth divergence was highly parallel in that stream fish always had deeper bodies than lake fish; yet the level of this parallelism was almost always greater at the regional than the global scale (Figures 1–4). A similar pattern emerged for gill raker number, with strong parallelism evident in lake-stream divergence at

the regional scale, consistent with previous studies (Berner et al. 2008, 2009; Kaeuffer et al. 2012; Stuart et al. 2017). Although lake-stream parallelism in these traits was also present at the global scale, it was less pronounced than at the regional scale (Table 2, Figure 3). For instance, also in accordance with previous studies (Berner et al. 2008; Kaeuffer et al. 2012), gill raker spacing and length exhibited low levels of parallelism (Figure 2E–H), yet the parallelism that was present in these traits was greater at the regional than at the global scale (Figure 3B).

With environmental variation across watersheds being greater at global than regional scales (Supplementary Figures S3 and S4), our data are consistent with the hypothesis that reduced environmental variation among watersheds at smaller spatial scales enhances (or enables) stronger parallel evolution. For instance, divergent selection related to diet, foraging mode, and swimming performance has been implicated in variation in body shape and gill rakers between lake and stream stickleback. In lakes, sticklebacks generally feed on small limnetic prey (Berner et al. 2008, 2009; Kaeuffer et al. 2012), leading to selection for an elongated body shape (Webb 1984; Walker 1997) and longer/more numerous gill rakers (Bentzen and McPhail 1984). In streams, stickleback commonly feed on benthic prey, leading to selection for deeper bodies and shorter/less numerous gill rakers (Bentzen and McPhail 1984; Berner et al. 2008, 2009; Hendry et al. 2011; Kaeuffer et al. 2012). Body depth and gill raker number are both known to be strongly heritable (Hermida et al. 2002; Aguirre et al. 2004; Oke et al. 2016; McPhail 1984), and as such, the stronger parallelism of these traits at regional than at global scales presumably reflects greater determinism of selection owing to more consistent ecological differences between these habitat types across replicate watersheds on smaller spatial scales.

Genetic (Non)parallelism

Like phenotypic parallelism, genetic parallelism at outlier loci was more pronounced at the regional than the global scale, especially in angles ($\theta_{OUTLIERS}$). In contrast, for neutral loci, genetic parallelism was similar at the 2 geographic scales. Why might we observe
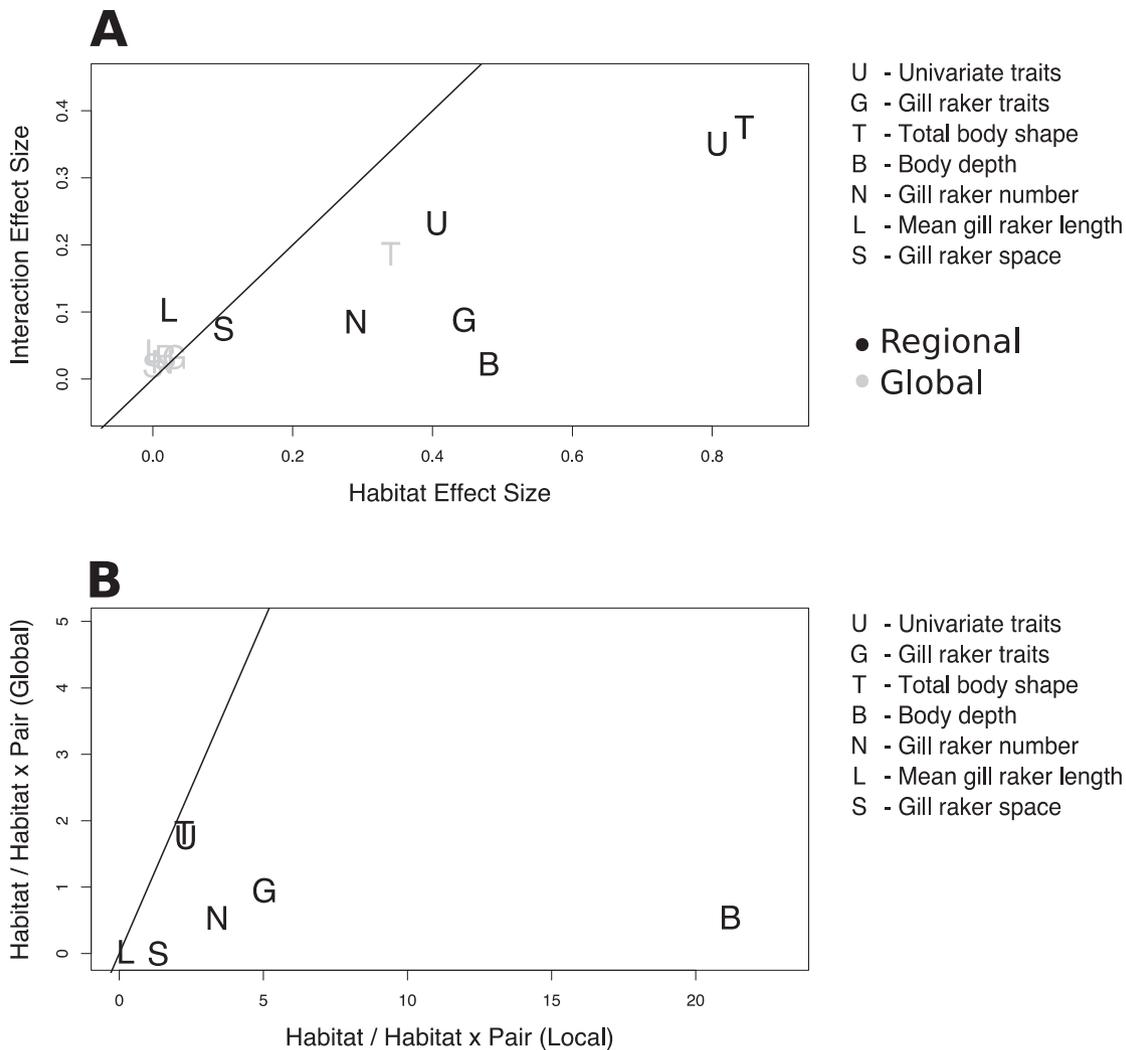
Figure 3. Phenotypic parallel divergence is higher in regional versus global comparisons. (**A**) Effect size of the habitat-by-pair interaction term versus effect size of the habitat term for regional and global pairs. (**B**) Ratio of the habitat effect size to the habitat-by-pair interaction effect size for regional pairs (*x* axis) versus global pairs (*y* axis) for all traits examined. In (A) and (B), effect size measured as η^2 for single traits and partial η^2 for multivariate traits. In (A), traits found below the 1:1 line show more parallelism (large, consistent effect of habitat) than nonparallelism (large, but inconsistent effect of habitat). Traits above the line show the opposite pattern. In (B), traits found below the 1:1 line show greater parallelism in regional than in the global pairs.

greater parallelism at the regional versus global scales for outlier genotypes? One explanation is that these outliers are associated with the underlying genetic basis of adaptive phenotypes, which—as described above—show greater parallelism at the regional than the global scale. However, phenotypic parallelism does not necessarily lead to genetic parallelism if different genetic changes are capable of attaining similar phenotypic outcomes (Conte et al. 2012; Kautt et al. 2012; Westram et al. 2014; Le Moan et al. 2016). The greater parallelism of outlier loci at regional than global scales might thus reflect more similar genetic architectures and higher levels of shared standing genetic variation available to selection at smaller geographical scales (Barrett and Schluter 2008; Agrawal and Stinchcombe 2009; Flint and Mackay 2009; Terekhanova et al. 2014; Llaurens et al. 2017; Saltz et al. 2017; Bassham et al. 2018). In support of this idea, we found that Vancouver Island lake-stream pairs shared more outlier loci than did lake-stream pairs at the global scale (21% of outliers were shared across a minimum of 2 pairs at the regional scale but only 4% of outliers were shared across a minimum of 2 pairs at the global scale). Note however that

the interpretation of the above results might be more specific to the direction of divergence rather than the magnitude. Indeed, outlier genetic vector length differences were mostly driven by 1 population (Norway, Supplementary Table S9D). Thus, rather than the magnitude of genetic divergence, differences between scale might mostly be driven by the relative weight of specific outlier loci in contributing to lake-stream divergence.

In contrast to the outlier loci, we detected no differences between geographic scales in the levels of (non)parallelism at neutral loci—at least when quantified using the *direction* (θ_G) of genetic vectors. However, neutral genetic vectors were somewhat more similar in *magnitude* (ΔL_G) at the regional than the global scale. This last result might suggest a role for gene flow in reducing divergence between watersheds over smaller spatial scales—that is, they are less genetically "independent" not only in their colonization but also in their contemporary genetic connections (Berner and Roesti 2017; Stuart et al. 2017; Haenel et al. 2018). It is also possible that these patterns could be driven by a difference in the effect of drift in shaping the extent of lake-stream neutral genetic divergence between scales.
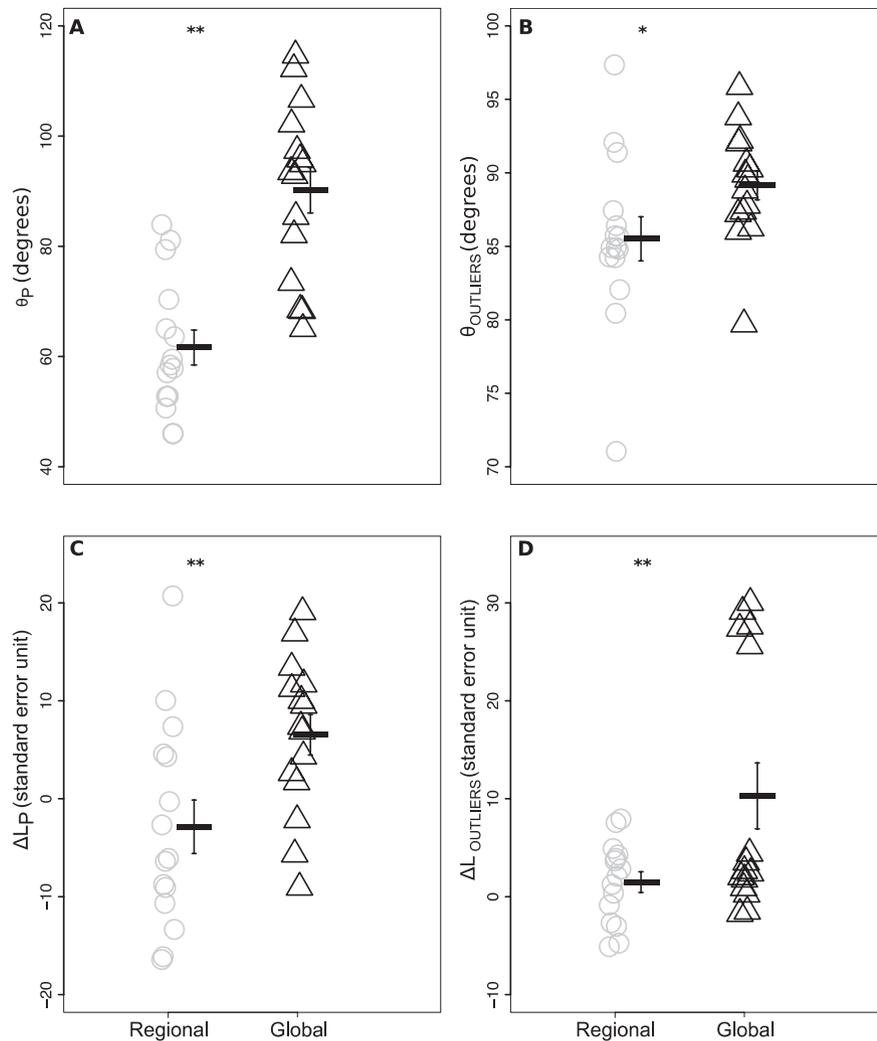
Figure 4. Pairwise angles between phenotypic vectors (**A**; θ_P), outlier genetic vectors (**B**; $\theta_{OUTLIERS}$), as well as differences in vector length between phenotypic vectors (**C**; ΔL_P) and outlier genetic vectors (**D**; $\Delta L_{OUTLIERS}$) at regional (gray circles) and global (black triangles) scales. Overall mean values are represented by a straight line (±SE). Significance in mean differences is represented by 2 stars ($P < 0.01$) or 1 star ($P < 0.05$).

Nonetheless, the lack of differences in θ_G between scales, and the lack of a relationship between θ_P and the average between-watershed neutral F_{ST} at either scale, imply that neutral processes are unlikely to be responsible for phenotypic divergence at either scale.

Our conclusion that parallelism is stronger at the regional than the global scale is based on only a single region (Vancouver Island). Thus, the alternative conclusion might be that Vancouver Island lake-stream stickleback are simply more parallel than lake-stream stickleback in other locations. Unfortunately, few of the regions we examined contain numerous lake-stream pairs, thus precluding a formal comparison. However, several studies have compared multiple lake-stream pairs within some regions—frequently revealing strong parallelism. For instance, strong within-region lake-stream parallelism has been argued for Ireland (Ravinet et al. 2013), Iceland (Lucek et al. 2014b), Switzerland (Lucek et al. 2014b), and Haida Gwaii (Deagle et al. 2012). Thus, while it is possible that parallelism is stronger on Vancouver Island than in other regions (Berner et al. 2010), it is clear that global lake-stream parallelism is even lower—as revealed here.

Constraints on Parallel Evolution at Regional Versus Global Scales

Lake-stream pairs at both scales showed relatively low levels of parallel evolution, especially for multivariate traits. Indeed, the mean angles between any 2 phenotypic vectors (θ_P) were 61.63° ± 12.24 SD and 90.17° ± 15.98 SD at the regional and global scales, respectively. Similarly, the mean angles between any 2 genetic outlier vectors ($\theta_{OUTLIERS}$) were 85.51° ± 5.81 SD at the regional scale and 89.15° ± 3.83 SD at the global scale. Multiple factors can restrict the level of parallel evolution at both scales (Bolnick et al. 2018). For instance, a number of ecological variables that differ between lakes and streams are known to affect the degree of stickleback divergence, including diet (Berner et al. 2008; Kaeuffer et al. 2012; Ravinet et al. 2013), parasite load (Eizaguirre et al. 2011), flow regimes (generally higher in streams: Jiang et al. 2015), and predators (predatory fish and birds likely more common in lakes: Reimchen et al. 1985; Reimchen 1994). However, the direction and magnitude of lake-stream divergence in these factors is not always consistent (Berner et al. 2008; Kaeuffer et al. 2012; Stuart et al. 2017). Parasite

communities are not universally more diverse in lakes than streams (Feulner et al. 2015), flow can be low in many stream reaches (Moore et al. 2007), and fish and bird predators vary dramatically among different lakes—and among different streams (Moodie and Reimchen 1976). Thus, variation across watersheds in the consistency of the ecological differences between lake and stream habitats could disrupt patterns of parallel evolution. We found this type of among-watershed environmental variation to be highest at the global geographic scale, and suggest that this nonparallel ecological divergence is likely to be a strong contributing factor to the nonparallel phenotypic and genetic divergence. However, it is noteworthy that univariate traits, such as body depth and gill raker number, were highly parallel in comparison to multivariate traits. One possible explanation is that "simple" univariate traits generally experience more parallel selection than "complex" multivariate traits for which multiple evolutionary solutions are possible to the same ecological challenge (Thompson et al. 2017).

At the genetic level, several factors can constrain the level of parallel evolution. Parallel evolution at loci associated with ecologically relevant phenotypes is likely shaped by environmental heterogeneity for the same reasons as those mentioned for phenotypes above (Orr 2005). Physically proximate populations are also be expected to be more closely related genetically (both because of colonization and gene flow), and therefore could be expected to display greater levels of parallelism due to sharing a more similar molecular basis of phenotypes (Barrett and Schluter 2008; Conte et al. 2012). For instance, with partially shared pools of standing genetic variation, closely related populations might have more similar genetic solutions to similar adaptive problems (Barrett and Schluter 2008). The Vancouver Island populations in our study likely diverged from a marine ancestor at a similar time ~5000 years ago (Stuart et al. 2017). In contrast, the global scale populations are estimated to have diverged at quite different times: as early as 5000–15 000 years ago in Alaska (Reger and Pinney 1996; Cresko et al. 2004) to hundreds to thousands of years in Europe and Iceland (Moser et al. 2012; Lucek et al. 2014b; Roesti et al. 2015). This discrepancy between scales in divergence times might also explain the observed differences in parallel evolution; for the reasons outlined above, populations with similar divergence times might be more parallel than populations with diverse divergence times.

Gene flow is another factor known to affect adaptive divergence (Slatkin 1987; Lenormand 2002; Garant et al. 2007) and therefore also evolutionary (non)parallelism. On the one hand, populations exchanging genes will become more genetically similar, which can promote parallelism. On the other hand, gene flow between populations in distinct environments can impede regional adaptation (Garant et al. 2007). Thus, gene flow within watersheds containing populations inhabiting distinct lake and stream habitat types should generally constrain their adaptive differentiation—as has frequently been inferred for lake-stream stickleback (Hendry et al. 2002; Moore et al. 2007; Berner et al. 2009; Stuart et al. 2017). If levels of gene flow are high within some watersheds and low in others, parallelism could be restricted, in particular as reflected by ΔL (Stuart et al. 2017). We detected lower ΔL_P and ΔL_G at regional than global scales, suggesting that the role of within-watershed gene flow in constraining adaptive differentiation could be more consistent at smaller spatial scales. However, ΔL_OUTLIERS was not different between the regional and global scales, which could imply similar levels of within-watershed gene flow across scales for outlier loci. The differences in ΔL_P may then reflect either differing levels of

phenotypic plasticity between scales (Oke et al. 2016), or that our genetic markers have not included the loci associated with the traits experiencing parallel adaptive divergence between habitats, which might show distinct patterns of admixture to those described here (Nosil et al. 2009).

Implications for Adaptive Radiation

Threespine stickleback are a remarkable example of adaptive radiation: They show dramatic phenotypic and evolutionary divergence from very small to very large geographical scales (Bell and Foster 1994). Moreover, most of this divergence has played out over just a few thousand generations; and, in some cases, impressive freshwater divergence from a marine ancestor has evolved in less than 50 years (Kimmel et al. 2012; Lescak et al. 2015; Bassham et al. 2018). At the same time, different stickleback populations in the adaptive radiation can still interbreed in nearly all cases (Hendry 2009); and they show at most 2 forms in a single location (but see: Hippel von and Weigner 2004). Hence, the applicability of the stickleback radiation to inferences about other classic adaptive radiations should be limited to the early stages of those radiations. Yet this potential limitation is also a strength in some respects; by capturing multiple replicates of the very early stages of adaptive radiation, we can more easily draw inferences about the processes that promote and constrain adaptive radiation—because we are removed from concern over confounding effects of noncausal factors that accumulate after the radiation is largely complete.

From this perspective, we found that, when replicate adaptive radiations are farther apart in space, and therefore are more likely to be independent, they show greater dissimilarities as reflected in lower phenotypic and genetic parallelism. We suggest that this scale-dependent repeatability is primarily driven by ecological rather than genetic factors. That is, the set of divergent environments in a given location will differ more from the same set of divergent environments in another location when those locations are situated farther apart in space. Indeed, the largest differences between scales appears to derive from phenotypes and genetic outliers, as opposed to neutral genetic markers. Further effort should be directed toward exploring both the ecological and genetic causes of scale-dependent determinism.

Author Contributions

A.P., D.H., A.P.H., and R.D.H.B. designed the study. Y.E.S. and D.I.B. provided sequence and environmental data from the regional pairs. F.A.v.H., M.K., T.K., S.S., and B.K.K. provided fish samples and collected environmental data from the global pairs. A.P. and D.H. extracted sequence data from the global pairs and phenotypic data from both scales. A.P. and D.H. analyzed the data. A.P. and D.H. wrote the manuscript with input from all coauthors.

Data Availability

We have deposited the primary data underlying these analyses as follows:

Raw phenotypic data and sync file of allele frequencies: Dryad (doi:10.5061/dryad.sq0tp3s).

References

Adams DC, Collyer ML. 2009. A general framework for the analysis of phenotypic trajectories in evolutionary studies. *Evolution*. 63:1143–1154.

Adams DC, Otárola Castillo E. 2013. geomorph: an r package for the collection and analysis of geometric morphometric shape data. *Methods Ecol Evol*. 4:393–399.

Agrawal AF, Stinchcombe JR. 2009. How much do genetic covariances alter the rate of adaptation? *Proc Biol Sci*. 276:1183–1191.

Aguirre WE. 2009. Microgeographical diversification of threespine stickleback: body shape–habitat correlations in a small, ecologically diverse Alaskan drainage. *Biol J Linn Soc*. 98:139–151.

Aguirre WE, Doherty PK, Bell MA. 2004. Genetics of lateral plate and gillraker phenotypes in a rapidly evolving population of threespine stickleback. *Behaviour*. 141:1465–1483.

Barrett RD, Schluter D. 2008. Adaptation from standing genetic variation. *Trends Ecol Evol*. 23:38–44.

Bartlett MS. 1937. Properties of sufficiency and statistical tests. *Proc R Soc A*. 160:268–282.

Bassham S, Catchen J, Lescak E, von Hippel FA, Cresko WA. 2018. Repeated selection of alternatively adapted haplotypes creates sweeping genomic remodeling in stickleback. *Genetics*. 209:921–939.

Bell MA, Foster SA. 1994. *The evolutionary biology of the threespine stickleback*. Oxford: Oxford University Press.

Bentzen P, McPhail JD. 1984. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): spatial segregation and seasonal habitat shifts in the Enos Lake species pair. *Can J Zool*. 62:2280–2286.

Berner D, Adams DC, Grandchamp AC, Hendry AP. 2008. Natural selection drives patterns of lake-stream divergence in stickleback foraging morphology. *J Evol Biol*. 21:1653–1665.

Berner D, Grandchamp AC, Hendry AP. 2009. Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution*. 63:1740–1753.

Berner D, Kaeuffer R, Grandchamp AC, Raeymaekers JA, Räsänen K, Hendry AP. 2011. Quantitative genetic inheritance of morphological divergence in a lake-stream stickleback ecotype pair: implications for reproductive isolation. *J Evol Biol*. 24:1975–1983.

Berner D, Roesti M. 2017. Genomics of adaptive divergence with chromosome-scale heterogeneity in crossover rate. *Mol Ecol*. 26:6351–6369.

Berner D, Roesti M, Hendry AP, Salzburger W. 2010. Constraints on speciation suggested by comparing lake-stream stickleback divergence across two continents. *Mol Ecol*. 19:4963–4978.

Blount ZD, Lenski RE, Losos JB. 2018. Contingency and determinism in evolution: replaying life's tape. *Science*. 362:eaam5979.

Bolnick DI, Barrett RDH, Oke KB, Rennison DJ, Stuart YE. 2018. (Non)parallel evolution. *Ann Rev Ecol Evol Syst*. 49:303–330.

Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan S, Simakov O, Ng AY, Lim ZW, Bezault E, *et al*. 2014. The genomic substrate for adaptive radiation in African cichlid fish. *Nature*. 513:375–381.

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

Collyer ML, Adams DC. 2007. Analysis of two-state multivariate phenotypic change in ecological studies. *Ecology*. 88:683–692.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G Jr, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science*. 307:1928–1933.

Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proc Biol Sci*. 279:5039–5047.

Craig MacLean R. 2005. Adaptive radiation in microbial microcosms. *J Evol Biol*. 18:1376–1386.

Cresko WA, Amores A, Wilson C, Murphy J, Currey M, Phillips P, Bell MA, Kimmel CB, Postlethwait JH. 2004. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc Natl Acad Sci USA*. 101:6050–6055.

Deagle BE, Jones FC, Chan YF, Absher DM, Kingsley DM, Reimchen TE. 2012. Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. *Proc Biol Sci*. 279:1277–1286.

Deagle BE, Reimchen TE, Levin DB. 1996. Origins of endemic stickleback from the Queen Charlotte Islands: mitochondrial and morphological evidence. *Can J Zool*. 74:1045–1056.

Eizaguirre C, Lenz TL, Sommerfeld RD, Harrod C, Milinski M. 2011. Parasite diversity, patterns of MHC II variation and olfactory based mate choice in diverging three-spined stickleback ecotypes. *Evol Ecol*. 25:605–622.

Elliott NG, Haskard K, Koslow JA. 1994. Morphometric analysis of orange roughy (*Hoplostethus atlanticus*) off the continental slope of Southern Australia. *J Invest Med*. 46:202–220.

Feulner PG, Chain FJ, Panchal M, Huang Y, Eizaguirre C, Kalbe M, Lenz TL, Samonte IE, Stoll M, Bornberg-Bauer E, *et al*. 2015. Genomics of divergence along a continuum of parapatric population differentiation. *PLoS Genet*. 11:e1004966.

Flint J, Mackay TF. 2009. Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res*. 19:723–733.

Garant D, Forde SE, Hendry AP. 2007. The multifarious effects of dispersal and gene flow on contemporary adaptation. *Funct Ecol*. 21:434–443.

Gillespie R. 2004. Community assembly through adaptive radiation in Hawaiian spiders. *Science*. 303:356–359.

Gillespie RG, Benjamin SP, Brewer MS, Rivera MAJ, Roderick GK. 2018. Repeated diversification of ecomorphs in Hawaiian stick spiders. *Curr Biol*. 28:941–947.e3.

Gould SJ. 1990. *Wonderful life: the burgess shale and the nature of history*. New York: WW Norton & Company.

Grant PR. 1981. Speciation and the adaptive radiation of Darwin's Finches. *Am Sci*. 69:653–663.

Haenel Q, Laurentino TG, Roesti M, Berner D. 2018. Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. *Mol Ecol*. 27:2477–2497.

Hendry AP. 2009. Ecological speciation! Or the lack thereof? *Can. J. Fish. Aquat. Sci*. 66:1383–1398.

Hendry AP, Hendry AS, Hendry CA. 2013. Hendry Vineyard stickleback: testing for contemporary lake–stream divergence. *Evol Ecol Res*. 15:343–359.

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

Hendry AP, Taylor EB. 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 AP, Taylor EB, McPhail JD. 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.

Hippel von FA, Weigner H. 2004. Sympatric anadromous-resident pairs of threespine stickleback species in young lakes and streams at Bering Glacier, Alaska. *Behaviour* 141:1441–1464.

James HF. 2004. The osteology and phylogeny of the Hawaiian finch radiation (Fringillidae: Drepanidini), including extinct taxa. *Zool J Linn Soc*. 141:207–255.

Jiang Y, Torrance L, Peichel CL, Bolnick DI. 2015. Differences in rheotactic responses contribute to divergent habitat use between parapatric lake and stream threespine stickleback. *Evolution*. 69:2517–2524.

Kaeuffer R, Peichel CL, Bolnick DI, Hendry AP. 2012. Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution*. 66:402–418.

Kautt AF, Elmer KR, Meyer A. 2012. Genomic signatures of divergent selection and speciation patterns in a 'natural experiment', the young parallel radiations of Nicaraguan crater lake cichlid fishes. *Mol Ecol*. 21:4770–4786.

Kimmel CB, Cresko WA, Phillips PC, Ullmann B, Currey M, von Hippel F, Kristjánsson BK, Gelmond O, McGuigan K. 2012. Independent axes of genetic variation and parallel evolutionary divergence of opercle bone shape in threespine stickleback. *Evolution*. 66:419–434.

Kocher TD. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. *Nat Rev Genet*. 5:288–298.

Kofler R, Orozco-terWengel P, De Maio N, Pandey RV, Nolte V, Futschik A, Kosiol C, Schlötterer C. 2011a. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS One*. 6:e15925.

Kofler R, Pandey RV, Schlötterer C. 2011b. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*. 27:3435–3436.

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

Le Moan A, Gagnaire PA, Bonhomme F. 2016. Parallel genetic divergence among coastal-marine ecotype pairs of European anchovy explained by differential introgression after secondary contact. *Mol Ecol*. 25:3187–3202.

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

Lescak EA, Bassham SL, Catchen J, Gelmond O, Sherbick ML, von Hippel FA, Cresko WA. 2015. Evolution of stickleback in 50 years on earthquake-uplifted islands. *Proc Natl Acad Sci USA*. 112:E7204–E7212.

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

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. 25:2078–2079.

Llaurens V, Whibley A, Joron M. 2017. Genetic architecture and balancing selection: the life and death of differentiated variants. *Mol Ecol*. 26:2430–2448.

Lleonart J, Salat J, Torres GJ. 2000. Removing allometric effects of body size in morphological analysis. *J Theor Biol*. 205:85–93.

Loh YH, Bezault E, Muenzel FM, Roberts RB, Swofford R, Barluenga M, Kidd CE, Howe AE, Di Palma F, Lindblad-Toh K, *et al.* 2012. Origins of shared genetic variation in African cichlids. *Mol Biol Evol*. 30:906–917.

Losos JB. 2009. *Lizards in an evolutionary tree*. Berkley (CA): University of California Press.

Lucek K, Lemoine M, Seehausen O. 2014a. Contemporary ecotypic divergence during a recent range expansion was facilitated by adaptive introgression. *J Evol Biol*. 27:2233–2248.

Lucek K, Roy D, Bezault E, Sivasundar A, Seehausen O. 2010. Hybridization between distant lineages increases adaptive variation during a biological invasion: stickleback in Switzerland. *Mol Ecol*. 19:3995–4011.

Lucek K, Sivasundar A, Kristjánsson BK, Skúlason S, Seehausen O. 2014b. Quick divergence but slow convergence during ecotype formation in lake and stream stickleback pairs of variable age. *J Evol Biol*. 27:1878–1892.

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

McPhail JD. 1984. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): morphological and genetic evidence for a species pair in Enos Lake, British Columbia. *Can J Zool*. 62:1402–1408.

Miller CT, Glazer AM, Summers BR, Blackman BK, Norman AR, Shapiro MD, Cole BL, Peichel CL, Schluter D, Kingsley DM. 2014. Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. *Genetics*. 197:405–420.

Moodie GEE, Reimchen TE. 1976. Phenetic variation and habitat differences in *Gasterosteus* populations of the Queen Charlotte Islands. *Syst Zool*. 25:49–61.

Moore JS, Gow JL, Taylor EB, Hendry AP. 2007. Quantifying the constraining influence of gene flow on adaptive divergence in the lake-stream threespine stickleback system. *Evolution*. 61:2015–2026.

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

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

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

Oke KB, Bukhari M, Kaeuffer R, Rolshausen G, Räsänen K, Bolnick DI, Peichel CL, Hendry AP. 2016. Does plasticity enhance or dampen phenotypic parallelism? A test with three lake-stream stickleback pairs. *J Evol Biol*. 29:126–143.

Orr HA. 2005. The probability of parallel evolution. *Evolution*. 59:216–220.

Peichel CL, Marques DA. 2017. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philos Trans R Soc Lond B Biol Sci*. 372:20150486.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*. 7:e37135.

Ravinet M, Prodöhl PA, Harrod C. 2013. Parallel and nonparallel ecological, morphological and genetic divergence in lake-stream stickleback from a single catchment. *J Evol Biol*. 26:186–204.

R Core Team. 2015. *R: A language and environment for statistical computing*. Vienna (Austria): R Foundation for Statistical Computing. Available from: https://www.R-project.org/

Reger RD, Pinney DS. 1996. Late Wisconsin glaciation of the Cook Inlet region with emphasis on Kenai lowland and implications for early peopling. In: Davis NY, Davis WE, editors. *Adventures through time readings in the anthropology of Cook Inlet, Alaska*. Anchorage (AK): Cook Inlet Historical Society. p. 15–35.

Reimchen TE. 1994. Predators and morphological evolution in threespine stickleback. In: Bell M, Foster SA, editors. *The evolutionary biology of the threespine stickleback*. Oxford: Oxford University Press. p. 240–276.

Reimchen TE, Stinson EM, Nelson JS. 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.

Reist JD. 1986. An empirical evaluation of coefficients used in residual and allometric adjustment of size covariation. *Can J Zool*. 64:1363–1368.

Reusch TB, Wegner KM, Kalbe M. 2001. Rapid genetic divergence in post-glacial populations of threespine stickleback (*Gasterosteus aculeatus*): the role of habitat type, drainage and geographical proximity. *Mol Ecol*. 10:2435–2445.

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

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

Rohlf FJ. 2006. *tpsDig, digitize landmarks and outlines, version 2.10*.

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

Saltz JB, Hessel FC, Kelly MW. 2017. Trait correlations in the genomics era. *Trends Ecol Evol*. 32:279–290.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, *et al*. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 9:676–682.

Schlötterer C, Tobler R, Kofler R, Nolte V. 2014. Sequencing pools of individuals - mining genome-wide polymorphism data without big funding. *Nat Rev Genet*. 15:749–763.

Schluter D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution*. 50:1766–1774.

Schluter D, Marchinko KB, Barrett RD, Rogers SM. 2010. Natural selection and the genetics of adaptation in threespine stickleback. *Philos Trans R Soc Lond B Biol Sci*. 365:2479–2486.

Schluter D, McPhail JD. 1992. Ecological character displacement and speciation in sticklebacks. *Am Nat*. 140:85–108.

Sharpe DMT, Räsänen K, Berner D, Hendry AP. 2008. Genetic and environmental contributions to the morphology of lake and stream stickleback: implications for gene flow and reproductive isolation. *Evol Ecol Res*. 10:849–866.

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science*. 236:787–792.

Stuart YE, Veen T, Weber JN, Hanson D, Ravinet M, Lohman BK, Thompson CJ, Tasneem T, Doggett A, Izen R, *et al*. 2017. Contrasting effects of environment and genetics generate a continuum of parallel evolution. *Nat Ecol Evol*. 1:158.

Taylor EB, McPhail JD. 1999. Evolutionary history of an adaptive radiation in species pairs of threespine sticklebacks (Gasterosteus): insights from mitochondrial DNA. *Biol J Linn Soc Lond* 66:271–291.

Terekhanova NV, Logacheva MD, Penin AA, Neretina TV, Barmintseva AE, Bazykin GA, Kondrashov AS, Mugue NS. 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. *PLoS Genet*. 10:e1004696.

Thompson CJ, Ahmed NI, Veen T, Peichel CL, Hendry AP, Bolnick DI, Stuart YE. 2017. Many-to-one form-to-function mapping weakens parallel morphological evolution. *Evolution*. 71:2738–2749.

Thompson CE, Taylor EB, McPhail JD. 1997. Parallel evolution of lake-stream pairs of threespine sticklebacks (*Gasterosteus*) inferred from mitochondrial DNA variation. *Evolution*. 51:1955–1965.

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

Webb PW. 1984. Body form, locomotion and foraging in aquatic vertebrates. *Am Zool*. 24:107–120.

Weber JN, Kalbe M, Shim KC, Erin NI, Steinel NC, Ma L, Bolnick DI. 2017. Resist globally, infect locally: a transcontinental test of adaptation by stickleback and their tapeworm parasite. *Am Nat*. 189:43–57.

Westram AM, Galindo J, Alm Rosenblad M, Grahame JW, Panova M, Butlin RK. 2014. Do the same genes underlie parallel phenotypic divergence in different *Littorina saxatilis* populations? *Mol Ecol*. 23:4603–4616.