Evolution, 58(10), 2004, pp. 2319–2331

HOW MUCH OF THE VARIATION IN ADAPTIVE DIVERGENCE CAN BE EXPLAINED BY GENE FLOW? AN EVALUATION USING LAKE-STREAM STICKLEBACK PAIRS

ANDREW P. HENDRY[1,2] AND ERIC B. TAYLOR[3,4]

[1]*Redpath Museum and Department of Biology, McGill University, 859 Sherbrooke Street West, Montréal, Québec H3A 2K6, Canada*

[2]*E-mail: andrew.hendry@mcgill.ca*

[3]*Department of Zoology and Native Fish Research Group, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia V6T 1Z4, Canada*

[4]*E-mail: etaylor@zoology.ubc.ca*

Abstract.—How much of the variation in adaptive divergence can be explained by gene flow? The answer to this question should objectively reveal whether gene flow generally places a substantial constraint on evolutionary diversification. We studied multiple independent lake-stream population pairs of threespine stickleback (*Gasterosteus aculeatus*). For each pair, we quantified adaptive divergence based on morphological traits that have a genetic basis and are subject to divergent selection. We then estimated gene flow based on variation at five unlinked microsatellite loci. We found a consistent and significant pattern for morphological divergence to be positively correlated with genetic divergence and negatively correlated with gene flow. Statistical significance and the amount of variation explained varied within and among traits: 36.1–74.1% for body depth and 11.8–51.7% for gill raker number. Variation within each trait was the result of differences among methods for estimating genetic divergence and gene flow. Variation among traits likely reflects different strengths of divergent selection. We conclude that gene flow has a substantial effect on adaptive divergence in nature but that the magnitude of this effect varies among traits. An alternative explanation is that cause and effect are reversed: adaptive divergence is instead constraining gene flow. This effect seems relatively unimportant for our system because genetic divergence and gene flow were not correlated with ecologically relevant habitat features of lakes (surface area) or streams (width, depth, flow, canopy openness).

Key words.—Adaptation, dispersal, ecological speciation, *Gasterosteus aculeatus*, microsatellites, migration, natural selection, population mixing.

Received June 15, 2004. Accepted July 14, 2004.

Theory predicts that gene flow can constrain the adaptive divergence of populations inhabiting different ecological environments (e.g., Haldane 1948; Slatkin 1973; Endler 1977; García-Ramos and Kirkpatrick 1997; Hendry et al. 2001; Lenormand 2002). Theory aside, however, the importance of gene flow in constraining adaptive divergence in nature remains debatable (Ehrlich and Raven 1969; Slatkin 1987; Blondel et al. 1999; Storfer 1999; Calsbeek and Smith 2003; Morjan and Rieseberg 2004). Although several studies have demonstrated that gene flow can indeed constrain adaptation for individual populations (e.g., Riechert 1993; King and Lawson 1995; Hendry et al. 2002; Saint-Laurent et al. 2003), these focused studies cannot be used to assess the overall extent to which adaptive divergence is influenced by gene flow. And yet such an assessment must be made if we are to determine whether gene flow is usually an important factor in adaptive divergence or whether it can be ignored in all but exceptional situations.

The amount of variation in a particular trait or phenomenon that can be explained by a particular ecological or evolutionary process is a topic of recent contention. Some authors argue that the amount of explained variation (regardless of trait, phenomenon, or process) is usually very small—typically about 2.5–5.4% of the total variation (Møller and Jennions 2002). Other authors argue that the amount of explained variation may be quite large, perhaps even 50% (Peek et al. 2003). Notwithstanding the conceptual and methodological differences between these two meta-analyses, the issue remains largely unresolved, particularly in regard to the effects of gene flow on adaptive divergence. Yet, resolving this uncertainty could be fundamental to our understanding of biological diversity. If gene flow usually explains only 2.5% of the variation in adaptive divergence, then perhaps gene flow can be ignored in most cases, allowing an exclusive focus on ecological environments (Schluter 2000). If, however, gene flow usually explains 50% of the variation in adaptive divergence, then studies seeking to understand evolutionary diversification will need to consider its potential effects (see also Calsbeek and Smith 2003). Discriminating between these alternatives, as well as intermediate or more extreme scenarios, also provides an objective way to appraise the opinion of some authors that natural selection usually overwhelms gene flow in nature (e.g., Ehrlich and Raven 1969; Levin 1979).

A profitable way to approach this question is to estimate the amount of adaptive divergence and gene flow between paired populations inhabiting different ecological environments and then test whether divergence and gene flow are negatively correlated across multiple evolutionarily independent pairs (e.g., Hendry et al. 2001). Specifically, adaptive divergence can be regressed on gene flow to reveal the amount of variation in the former that is explained by the latter. The principal advantage of this broad-brush approach is that it examines the importance of gene flow while retaining the potential effects of all other factors that might influence adaptation. Several studies have adopted a paired-sample correlative approach for testing associations between gene flow and adaptive divergence. Most of these, however, have not used independent pairs (e.g., Smith et al. 1997; Storfer and Sih 1998; Storfer et al. 1999; Ogden and Thorpe 2002; Calsbeek and Smith 2003), and those that have used few such pairs: $N = 4$ independent pairs for char morphs (Gíslason et

al. 1999), $N = 6$ for benthic and limnetic whitefish (Lu and Bernatchez 1999), $N = 4$ for benthic and limnetic stickleback (Hendry et al. 2001), and $N = 4$ for two Neotropical fish species (Langerhans et al. 2003). In the present study, we quantify the relationship between adaptive divergence and gene flow across eight independent population pairs of threespine stickleback (*Gasterosteus aculeatus*).

Lake and Stream Stickleback

Our analysis was based on parapatric lake and stream stickleback in British Columbia, Canada, for which three archetypal population pairs have been described: Mayer Lake (Moodie 1972a,b), Drizzle Lake (Reimchen et al. 1985), and Misty Lake (Lavin and McPhail 1993; Thompson et al. 1997; Hendry et al. 2002). Within each of these pairs, lake stickleback have shallower bodies and more gill rakers than stream stickleback, differences that have an additive genetic basis and a clear adaptive interpretation (Moodie 1972a,b; Reimchen et al. 1985; Lavin and McPhail 1993; McPhail 1994; Hendry et al. 2002). Specifically, shallower bodies are better suited for sustained swimming, which is typical in lakes, whereas deeper bodies are better suited for burst swimming and rapid maneuvering, which is typical in streams. Likewise, more gill rakers are better suited for feeding on zooplankton, which predominate in lakes, whereas fewer gill rakers are better suited for feeding on benthic macro-invertebrates, which predominate in streams (see references in Hendry et al. 2002). Lake and stream stickleback can also differ in several other traits, such as coloration and defensive armor, but the direction of divergence in these traits is not consistent. For example, relative to stream fish, lake fish have shorter pelvic spines and more lateral plates in the Mayer and Drizzle watersheds (Moodie 1972a,b; Reimchen et al. 1985) but longer pelvic spines and fewer lateral plates in the Misty watershed (Lavin and McPhail 1993; Hendry et al. 2002).

Divergent selection thus favors adaptive divergence between lakes and streams; but because the environments are parapatric, the amount of divergence may be constrained by gene flow. Hendry et al. (2002) confirmed this intuition for the Misty watershed: relative to the inlet, the outlet experienced higher gene flow from the lake and contained stickleback that were less divergent. Here we use multiple lake-stream pairs to determine the amount of variation in adaptive divergence that can be explained by gene flow. We chose this system in part because only three highly divergent lake-stream pairs have been described, even though stickleback have been sampled from many additional lakes and streams (McPhail 1994). This suggests that other watersheds are characterized by less adaptive divergence. If gene flow has played a constraining role in these other watersheds, the extent of adaptive divergence and the amount of gene flow should be negatively correlated when compared across lake-stream pairs.

An Alternative: Ecological Speciation

We have thus far discussed negative correlations between gene flow and adaptive divergence as though the former constrains the latter. It is also possible, however, that cause and effect are reversed and that adaptive divergence, which reflects divergent selection, instead constrains gene flow (i.e., ecological speciation, Schluter 2000). Furthermore, both causal pathways may act at the same time and their relative importance may vary among traits (Hendry et al. 2001; Nosil and Crespi 2004). For example, traits that influence mate choice might be particularly important in ecological speciation (Schluter 2000; McKinnon et al. 2004). Reflecting this uncertainty about the direction of causation, negative associations between gene flow and adaptive divergence have been variously interpreted as evidence that the former constrains the latter (Storfer and Sih 1998; Storfer et al. 1999; Hendry et al. 2001, 2002; Calsbeek and Smith 2003) or the latter constrains the former (Gíslason et al. 1999; Lu and Bernatchez 1999; Ogden and Thorpe 2002). Determining the direction of causality may be particularly germane for threespine stickleback because they have been advanced as providing evidence that gene flow constrains adaptive divergence (Bell and Richkind 1981; Bell 1982; Bourgeois et al. 1994; Hendry et al. 2002) and that adaptive divergence constrains gene flow (McPhail 1994; Schluter 2000; Reusch et al. 2001; McKinnon and Rundle 2002; McKinnon et al. 2004). We here attempt to distinguish between these alternatives by also testing whether the amount of gene flow is negatively correlated with an independent surrogate for the strength of divergent selection—the degree to which ecologically relevant habitat features differ between lakes and streams.

Materials and Methods

Stickleback Collections

We collected freshwater-resident stickleback from lakes and streams in eight watersheds located in southwestern British Columbia, Canada. Lake-stream divergence within each watershed most likely occurred postglacially after independent colonization by marine ancestors, as appears generally the case for freshwater stickleback (Bell and Foster 1994; McPhail 1994; Thompson et al. 1997; Taylor and McPhail 2000). Three of the watersheds (Misty, Beaver, Stephen's) are located on northern Vancouver Island, three (Merrill/Boot, Pye, Amor de Cosmos) are on central Vancouver Island, one (Village Bay) is on Quadra Island, and one (Vanada) is on Texada Island. The geography of each watershed and the proximity of lake and stream collections varies in ways that might influence gene flow (Fig. 1).

From these eight watersheds, we obtained 16 paired lake-stream collections for morphological analysis, with 12 of these pairs then used for microsatellite analysis (see Results). Among watersheds, collections were always very genetically distinct, confirming their status as independent replicates (see Results). Even within some watersheds, different lake-stream pairs appeared essentially independent by the same criteria of physical separation and genetic divergence: that is the Mud and Mackie pairs in the Pye watershed and the pairs at either end of McCreight Lake in the Amor de Cosmos watershed (Fig. 1). Within other watersheds, however, different pairs could not be considered independent: that is in the Misty and Boot/Merrill watersheds (see Results).

All stickleback collections were made by deploying unbaited minnow traps in lakes and streams. Some lake-stream

pairs were sampled in each of two years and others in only one year. In one watershed (Village Bay), collections were made at two times (May and August) in the same year. Year- and/or month-specific samples were kept separate for consideration as temporal replicates. When fewer than 30 fish were collected at a given site, all were retained for analysis. When more than 30 fish were collected, a subset of 30 were haphazardly selected and retained. All retained fish were killed with an overdose of MS-222 and preserved immediately in 95% ethanol (as in Hendry et al. 2002).

Adaptive (Morphological) Divergence

Several months after fish collection, morphological traits were measured according to the procedures described in Hendry et al. (2002). Measured traits included body length (tip of the upper jaw to the end of the hypural plate), body depth (anterior insertion of the first dorsal spine to the bottom of the pelvic girdle, perpendicular to the lateral line), pelvic spine length (insertion to tip of the left spine), upper jaw length (tip of the upper jaw to the end of the left maxilla), number of gill rakers (on the left side of the outermost gill arch), number of lateral plates (on the left side), and pelvic girdle width (at its widest point). Measurements were performed by two people, one in each major collection year (1998 and 1999). One collection was processed by both individuals, which yielded highly correlated measurements (e.g., body depth, Pearson's $r = 0.993$, $P < 0.001$) but different mean values (e.g., 8.35% difference for body depth). To ensure that differences between measurers did not confound our interpretation, the unit of replication for all analyses was the difference between lake and stream fish within each pair, with the collections comprising a pair having been made at the same time and having been measured by the same person.

All traits were $\log_{10}$ transformed and those correlated with body size (i.e., body depth, pelvic spine length, upper jaw length, and pelvic girdle width) were then standardized to a common body length of 49.8 mm (the mean of all fish). Analysis of covariance (ANCOVA, performed in SPSS ver. 11.0.1) was first used to test for interactions between length and collection (i.e., heterogeneity of slopes). The interaction term for body depth was nonsignificant ($P = 0.189$) and was therefore removed from the ANCOVA to allow calculation of the common slope ($b = 1.083$). Allometric size-standardization must use this common-within-group slope, as opposed to collection-specific or pooled slopes (Reist 1986). Standardized body depth was then calculated for each fish as $M_{std} = M_{obs}(49.8/L_{obs})^b$, where M is the trait size (not log transformed), L is the body length (not log transformed), and the subscripts *std* and *obs* refer to standardized and observed (raw) measurements (Reist 1986).

The interaction terms from ANCOVA were significant for upper jaw length ($P = 0.011$), pelvic spine length ($P < 0.001$), and pelvic girdle width ($P < 0.001$). Serial removal of collections with the most divergent slopes revealed that heterogeneity was the result of just three collections for upper jaw length, eight collections for pelvic girdle width, and more than half of the collections for pelvic spine length. Regardless, common slope coefficients were similar when calculated with or without these divergent collections: upper jaw length (with, $b = 1.237$; without, $b = 1.217$), pelvic girdle width (with, $b = 1.943$; without, $b = 1.981$), and pelvic spine length (with, $b = 0.702$; without, $b = 0.823$). Moreover, size-standardized trait values were very similar regardless of which coefficient was used. For these reasons, we performed size-standardization for these traits using the common slope calculated with all collections in the ANCOVA. A similar approach was adopted by Hendry et al. (2002).

Four metrics were selected a priori to represent the amount of adaptive divergence between lake and stream fish within each pair. Two metrics were based on individual traits that are clearly subject to divergent selection: standardized body depth and gill raker number (see introduction). These traits are additionally appropriate because lake-stream divergence in them has an additive genetic basis (e.g., Hendry et al. 2002) and because neither trait is burdened by heterogeneous relationships with body size (see above). For each trait, the difference in the mean value of lake and stream fish for a pair was calculated so as to maintain a positive overall expectation: that is stream minus lake for standardized body depth but lake minus stream for gill raker number. The lake-stream difference for each pair was then divided by the mean trait value for the lake in that pair, generating relative differences for each pair. Armor traits (e.g., lateral plates and pelvic spine length) were additional candidates as measures of adaptive divergence (for a review see Reimchen 1994). However, we decided a priori against using them in this way because (1) they do not differ consistently for the archetypal lake/stream pairs (see introduction); (2) the genetic basis of their lake or stream divergence is less clear (Hendry et al. 2002); and (3) interpretations could be influenced by heterogeneous relationships with body size (see above).

The other two measures of adaptive divergence were multivariate composites of the traits, with body depth, upper jaw length, pelvic spine length, and pelvic girdle width standardized as above. These composites were calculated as eigenvalues from discriminant functions analysis performed separately for each pair (following Lu and Bernatchez 1999). One composite was the eigenvalue based on all seven traits, whereas the second was the eigenvalue based on all traits except body length. Eigenvalues without body length were calculated because this trait is considerably more plastic than the others (Hendry et al. 2002). Note that although armor traits were not considered individually, they were included in the multivariate composites of divergence. For each discriminant functions analysis, we also determined the overall success in classifying individuals to their specific collection within that pair.

Genetic Divergence and Gene Flow

We estimated genetic divergence and gene flow for 12 of the 16 paired collections. These 12 pairs were selected to encompass a wide range of adaptive divergence and to exclude pairs that were not truly parapatric. This last criterion excluded the Stephen's and Vanada watersheds because another lake was situated between our lake and stream collections (Fig. 1). The two collections comprising each of the 12 pairs were screened for allelic variation at five microsatellite

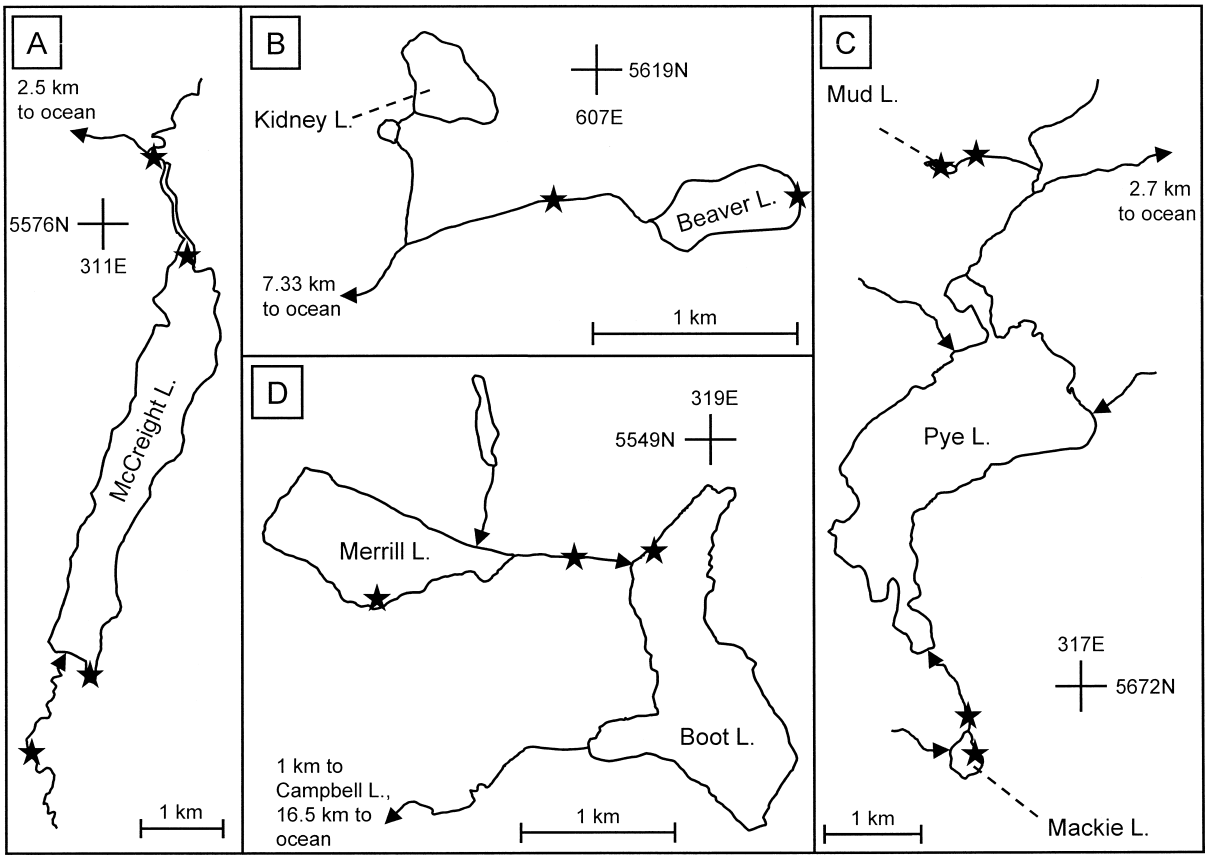

FIG. 1. Maps showing the locations (indicated by stars) from which lake and stream stickleback were collected. Each watershed is shown in a different panel: (A) Amor de Cosmos, (B) Beaver, (C) Pye, (D) Boot/Merrill, (E) Stephen's, (F) Village Bay, (G) Vanada, and (H) Misty. Arrows at the end of streams indicate the direction of water flow. Distances from the most downstream point shown in each watershed to the ocean are indicated. Selected UTM coordinates are shown as a cross in each panel.

loci used in our previous work (Taylor and McPhail 2000; Hendry et al. 2001, 2002). These loci had been isolated from stickleback genomic libraries and were assayed using polymerase chain reaction and radiolabeled primers as described by Rico et al. (1993, *Cir*51) and Taylor (1998; *Gac*4, *Gac*7, *Gac*9, *Gac*14). Screening and genotyping procedures are detailed in Taylor (1998).

We first tested each locus-collection combination for deviations from Hardy-Weinberg and linkage equilibrium. We then used UPGMA with Nei's (1978) unbiased genetic distance to visualize genetic relationships among the collections. Clustering relationships were very similar when generated with other distance metrics and clustering algorithms (results not shown). We next examined lake-stream divergence within each pair by calculating a variety of metrics: F_{ST} (θ of Weir and Cockerham 1984), ρ_{ST} (Rousset 1996), genic and genotypic differentiation, various genetic distances, and classification success in assignment tests. These analyses were performed in GENEPOP (Raymond and Rousset 1995, ver. 3.3), FSTAT (Goudet 1995, ver. 2.9.3), and TFPGA (Miller 1997, ver. 1.3). Assignment tests employed the Rannala and Mountain (1997) Bayesian method with a threshold of 0.05, as implemented in GeneClass2 (Piry et al. 2004).

We next estimated gene flow between lake and stream populations within each pair. A variety of methods were used because no single method has unequivocal support (Slatkin and Barton 1989; Beerli and Felsenstein 1999; Gaggiotti et al. 1999; Whitlock and McCauley 1999; Abdo et al. 2004). As in Hendry et al. (2002), these methods included Wright's infinite island model ($F_{ST} = 1/[1 + 4N_em]$), Takahata's (1983) finite island model, rare/private alleles (Slatkin 1985), and the maximum-likelihood coalescent program MIGRATE (Beerli and Felsenstein 1999). For MIGRATE (ver. 1.7.3), we used the microsatellite model with a threshold of 30, except where noted. For each estimate of gene flow, lake and stream fish within a pair were treated as though they were the only two populations exchanging genes, which may or may not lead to biased estimates when unsampled populations are present (Beerli 2004). For each of the four methods, we first estimated the effective number of migrants (N_em) and then converted this to an estimate of the proportion of migrants (m).

For the Wright, Takahata, and Slatkin methods, m was estimated by dividing the total N_em for each pair (twice the obtained N_em estimate) by an estimate of the total effective population size (N_e) for that pair. Total N_e for a pair was the sum of the N_e estimates for the two collections, each obtained as $H/(1 - H)4\mu$ (Waples 1991), where H was the average expected heterozygosity from microsatellite data and μ (mutation rate) was assumed to be 10^{-4} (as in Hendry et al. 2002).

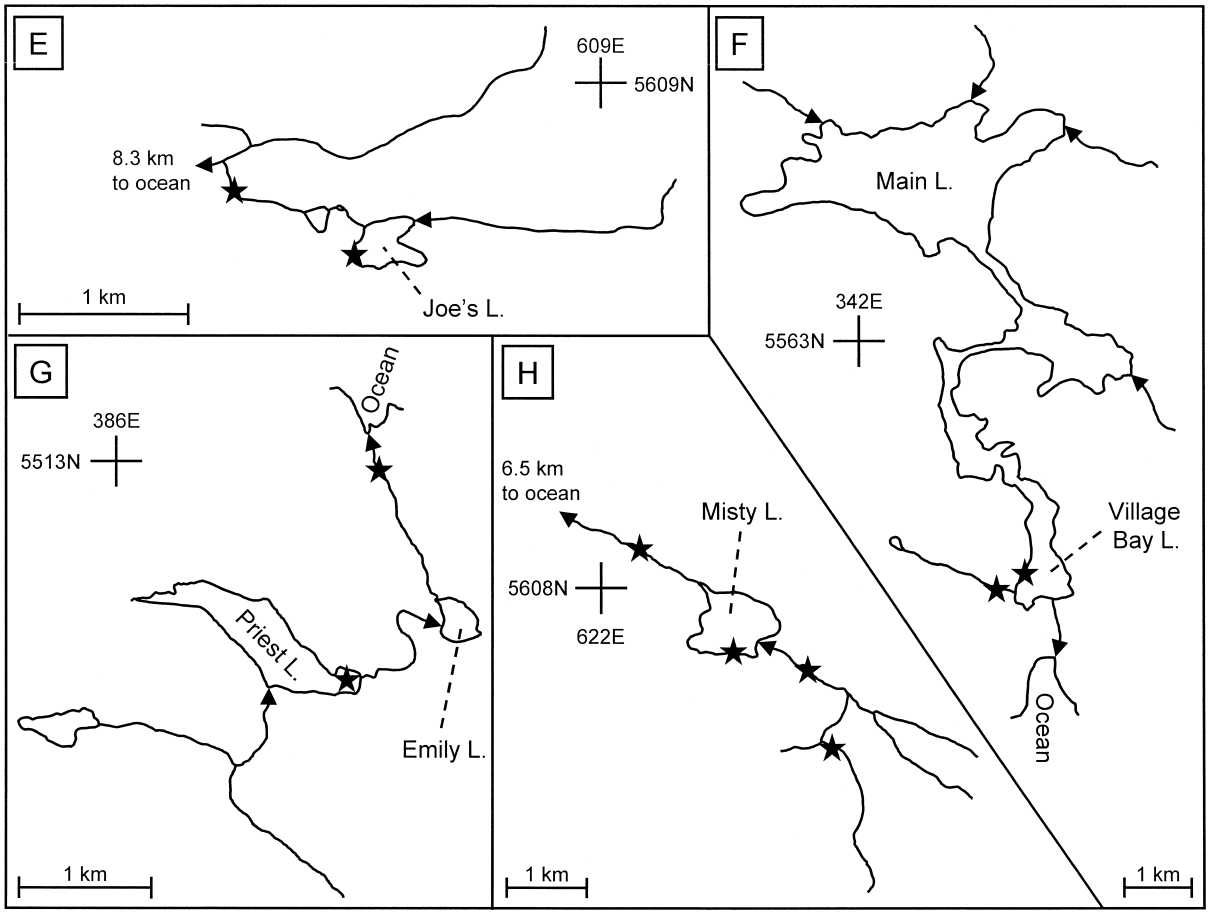

FIG. 1. Continued.

For the Beerli method, m was estimated by summing the two undirectional N_em values for each pair, and then dividing this sum by the sum of the two N_e estimates for each pair. Each of the above procedures requires many assumptions, some of which are likely violated by lake-stream stickleback pairs. However, we are here interested in *relative* differences among pairs in the amount of gene flow, which should be fairly robust to such violations.

Adaptive Divergence versus Gene Flow

We used simple linear regressions to assess the relationship between measures of adaptive divergence (dependent variable) and gene flow (independent variable). To statistically evaluate these relationships, we needed independent datapoints which necessitated the exclusion of several of the 12 paired collections. Specifically, we could only use one pair in the Misty watershed and therefore chose the best-studied contrast: Misty Lake versus Misty Lower Inlet (Lavin and McPhail 1993; Thompson et al. 1997; Hendry et al. 2002). Of the two paired collections available for this contrast, we chose the one (1998) with the larger sample size. For the Boot/Merrill watershed, the stream collection was identical for both lake-stream pairs (Fig. 1), and so we chose the pair for which gene flow should be higher (Merrill Lake vs. Merrill Outlet). Note, however, that we retained both Amor de Cosmos (McCreight Lake) pairs because each lake collection was adjacent to its paired stream collection at either end of a large lake (Fig. 1), and each lake collection was genetically more similar to its adjacent stream collection than to the other lake collection (see Results). Based on the same criteria of geographic separation (Fig. 1) and genetic distinctiveness (see Results), we retained both pairs in the Pye watershed (Mud and Mackie).

We regressed each of the four measures of adaptive divergence against four different measures of genetic divergence and four different measures of gene flow (all chosen a priori). For genetic divergence, we used one distance measure based on mutation (unbiased distance of Nei 1978), one distance measure based on drift (coancestry coefficient of Reynolds et al. 1983), one correlation measure based on allele frequencies (F_{ST}, based on θ of Weir and Cockerham 1984) and one correlation measure based on allele sizes (ρ_{ST} of Rousset 1996). For gene flow, we first used Wright's standard estimates of N_em, which were perfectly correlated with estimates using Takahata's method (the latter simply divides the former by four). The two methods are therefore grouped henceforth as Wright-Takahata. We then used each of the three estimates of m: Wright-Takahata, Slatkin, and Beerli. Regressions were performed using $\log_{10}$ transformed genetic divergence or gene flow, except for F_{ST} and ρ_{ST}, which provided better fits without transformation.

Divergent Selection versus Gene Flow

If adaptive divergence constrains gene flow, as in ecological speciation, an increase in the strength of divergent selection should cause a decrease in gene flow. Directly measuring selection at all of our study sites was prohibitive, given the difficulties associated with accurately estimating selection at even a single location (Schluter 2000; Hersch and Phillips 2004). Fortunately, a surrogate for divergent selection can be obtained by measuring ecologically relevant habitat features, which can then be examined for associations with gene flow (e.g., Smith et al. 1997; Reusch et al. 2001; Ogden and Thorpe 2002). Selection on foraging-related traits (including gill rakers and body depth) will be a function of the overall size of lakes and streams. Specifically, large lakes have more limnetic (plankton) foraging opportunities, whereas small lakes have more benthic (macro-invertebrate) foraging opportunities (Lavin and McPhail 1985, 1986). The same contrast should hold for large versus small streams. For our study lakes, we obtained data on lake size (surface area and maximum depth) from the online British Columbia government database FishWizard (available via pisces.env.gov.bc.ca). Depth data were lacking for two lakes (Mackie and Mud) but surface area and depth were highly correlated for the other lakes ($r^2 = 0.77$, $P = 0.022$). We therefore used surface area as our sole measure of lake size.

For streams, we directly measured habitat features at each of our stickleback collection sites. All data were obtained between July 20 and 31, 2003, a period during which rainfall was light and water levels remained reasonably constant. At each site, we established seven to 11 transects across the stream. Spacing between transects was equidistant at each site but varied from 2.5 m to 10 m among sites to match the locations from which stickleback were collected. At each transect, we measured the wetted width of the stream and then established three equidistant points across the stream. At each of these points, we measured water depth and water flow, the latter with a Swoffer (Seattle, WA) model 2100 flow meter (impeller positioned 60% of the distance from the substrate to the surface). At five equidistant points between our lowest and highest transects at each site, we used a spherical densiometer (Lemmon 1957) to measure canopy openness (the proportion of the sky not covered by forest canopy). This method yields openness estimates that are highly correlated with those obtained using hemispherical photography (Englund et al. 2000). Site averages were calculated for each habitat variable and used in subsequent analyses. We also used principal components (PCs) to combine stream width and depth into an overall measure of stream size (PC1).

To determine whether divergent selection might influence gene flow, we used simple linear regressions with genetic divergence or gene flow as the dependent variable and each habitat feature as an independent variable. These regressions were directly equivalent to those used for testing whether adaptive divergence was correlated with gene flow: the lake-stream pairs and the measures of genetic divergence and gene flow were identical. Habitat variables used in the regressions included lake surface area, stream size (PC1), water flow, canopy openness, and lake surface area (*Z*-scores) minus stream size (PC1). This last variable generated values that were positively correlated with the difference in size between lakes and streams (weighted equally).

Results

Adaptive (Morphological) Divergence

All four measures of adaptive divergence were positively correlated with each other: Pearson's $r = 0.527$–0.971 and all $P < 0.05$ (based on all 16 paired collections in Table 1). Of particular interest was the strong correlation between divergence in standardized body depth and divergence in gill raker number ($r = 0.775$, $P < 0.001$; Fig. 2), two very different traits that both have an additive genetic basis and are subject to divergent selection between lakes and streams (see introduction). For lake-stream pairs in which divergence was substantial, it was usually in the expected direction: shallower bodies and more gill rakers for lake fish. Exceptions included Beaver Lake–Outlet, for which lake fish had more gill rakers but not shallower bodies, and McCreight Lake–Outlet, for which lake fish had shallower bodies but not more gill rakers.

Despite this generally consistent direction of divergence between lakes and streams, the magnitude of divergence varied greatly among pairs. Specifically, divergence was greatest for Misty Lake–Inlet and Village Bay Lake–Inlet; least for Merrill Lake–Outlet, Priest Lake–Outlet, and McCreight Lake–Outlet; and intermediate for the other pairs. Moreover, the relative magnitude of divergence remained consistent for the pairs with temporal replication: greatest for both Misty Lake–Lower Inlet replicates, least for both McCreight Lake–Inlet replicates, and intermediate (but relatively high) for both Village Bay Lake–Inlet replicates (Table 1). Our estimates of divergence were therefore repeatable and consistent through time.

Genetic Divergence and Gene Flow

The five microsatellite loci were independent because only 13 of 190 tests for linkage disequilibrium were individually significant at $\alpha = 0.05$ (9.5 significant tests would be expected by chance). Moreover, the 13 deviations were not consistently associated with any particular pair of loci. Four of the loci appeared to be in approximate Hardy-Weinberg equilibrium, but *Gac*9 showed significant heterozygote deficits in about half of the collections. Owing to the possibility of null alleles at this locus, we performed all subsequent analyses both with and without *Gac*9, which had almost no effect on our results. For example, genetic divergence and gene flow estimates including *Gac*9 were highly correlated with those excluding *Gac*9: for all 12 pairs; F_{ST}, $r = 0.970$, $P < 0.001$; ρ_{ST}, $r = 0.997$, $P < 0.001$; N_em from private alleles, $r = 0.802$, $P = 0.002$; and m from MIGRATE (microsatellite model with a threshold of 10), $r = 0.937$, $P < 0.001$. Given this equivalency, we present results based on all five loci.

Stickleback collections were genetically more similar within than among watersheds (Fig. 3). The only ambiguous case was Village Bay, where lake and stream collections were so different from each other that they did not cluster together. Both collections nevertheless remained quite distinct from all other watersheds (Fig. 3). These patterns support our ex-

TABLE 1. Morphological divergence between lake and stream fish for each pair of collections. Size-standardized body depth (body depth) and the number of gill rakers (gill rakers) are relative differences calculated so as to maintain a positive expected difference (i.e., stream minus lake for body depth and lake minus stream for gill rakers). Eigenvalues are from discriminant function analysis for each pair individually and were based on all traits (eigenvalue) or all traits except body length (eigenvalue w/o length). For each discriminant functions analysis, we also report the percentage of individuals correctly assigned to their collection location (assignment success). Sample sizes (*n*) are the number of individuals for which all traits could be measured.

Lake	Stream	Year	*n* (lake/stream)	Body depth	Gill rakers	Eigenvalue	Eigenvalue w/o length	Assignment success (%)	Assignment success w/o length (%)
Misty	Upper inlet	1999[1]	21–30/19	0.12	0.14	3.550	3.101	94.9	95.8
Misty	Lower inlet	1997	21/30	0.16	0.14	7.340	6.039	100	100
Misty	Lower inlet	1998	30/30	0.19	0.15	5.820	4.992	98.3	98.3
Misty	Outlet	1998	30/29	0.09	0.11	1.840	1.638	96.6	96.6
Boot	Inlet	1998	30/30	0.01	0.05	0.635	0.554	76.7	81.7
Merrill	Outlet	1998	30/22	−0.03	−0.01	2.600	0.977	98.1	84.6
Beaver	Outlet	1999	27/29	−0.02	0.12	1.405	1.395	85.7	85.7
McCreight (south)	Inlet	1999	30/30	0.02	0.04	0.989	0.986	83.3	81.7
McCreight (south)	Inlet	1998	30/14	0.06	0.07	0.818	0.661	86.4	77.3
McCreight (north)	Outlet	1999	30/30	0.05	−0.01	0.406	0.369	68.3	78.3
Mackie	Outlet	1999	27/30	−0.03	−0.05	1.739	0.586	91.2	82.5
Village Bay (May)	Inlet	1998	30/22	0.16	0.11	2.253	1.604	94.2	88.5
Village Bay (Aug)	Inlet	1998	10/28	0.13	0.12	1.369	0.971	92.1	84.2
Mud	Outlet	1999	30/30	0.03	0.07	1.945	1.830	95.0	90.0
Joe's	Outlet	1999	28/27	0.06	0.05	3.243	2.915	92.7	94.5
Priest	Outlet	1998	30/30	−0.06	0.00	0.639	0.591	75.0	73.3

[1] This collection could not be compared to a Misty Lake collection from the same year; values shown here are the average of comparisons between this collection and the two Misty Lake collections from other years (1997, 1998).

FIG. 2. Divergence between lake and stream stickleback is highly correlated for standardized body depth versus the number of gill rakers. Squares are lake-outlet pairs and circles are lake-inlet pairs, all based on values from Table 2. Open symbols (squares and circles) are pairs for which microsatellite data were not obtained. Note that the axes do not cross at the origin.

pectation that lake-stream pairs from different watersheds have a long history of evolutionary independence. This independence might take the form of: (1) separate colonizations of each watershed by marine ancestors, followed by de novo divergence into lake and stream forms within each watershed and then low gene flow within and among watersheds; (2) a single origin of lake and stream forms from marine ancestors, followed by their spread to all watersheds and then moderate gene flow within watersheds but low gene flow among watersheds; or (3) some intermediate between these extremes. Separate origins from marine ancestors seems most likely (Bell and Foster 1994; McPhail 1994; Thompson et al. 1997; Taylor and McPhail 2000), but any of these scenarios are consistent with thousands of generations of effective independence among watersheds. Based on physical separation (Fig. 1) and genetic differences (Fig. 3), we further infer the effective independence of two pairs within the Pye watershed (Mackie and Mud) and two pairs within the Amor de Cosmos watershed (southern McCreight Lake vs. McCreight Inlet and northern McCreight Lake vs. McCreight Outlet).

Lake and stream fish within each pair of collections differed significantly in genotype and allele frequencies. The magnitude of divergence, however, varied dramatically among pairs (Table 2) and was consistent across the four measures of genetic divergence (based on all data in Table 2, Pearson's $r = 0.656$–0.998, all $P < 0.05$). This consistency was remarkable given the fundamental theoretical differences among these metrics. In general, genetic divergence was greatest for Misty Lake–Inlet and Village Bay Lake–Inlet, least for Misty Lake–Outlet and Merrill Lake–Outlet, and intermediate for the other pairs (Table 3). Assignment tests using multilocus genotypes were consistent with these distinctions, except that Beaver Lake–Outlet had the lowest assignment success (Table 3).

Substantial variation among pairs was also evident for gene flow (Table 3), with some of the methods yielding estimates that were strongly correlated (Slatkin vs. Wright-Takahata: N_em, $r = 0.759$, $P = 0.004$; m, $r = 0.737$, $P = 0.006$), some that were moderately correlated (Slatkin vs. Beerli: N_em, r

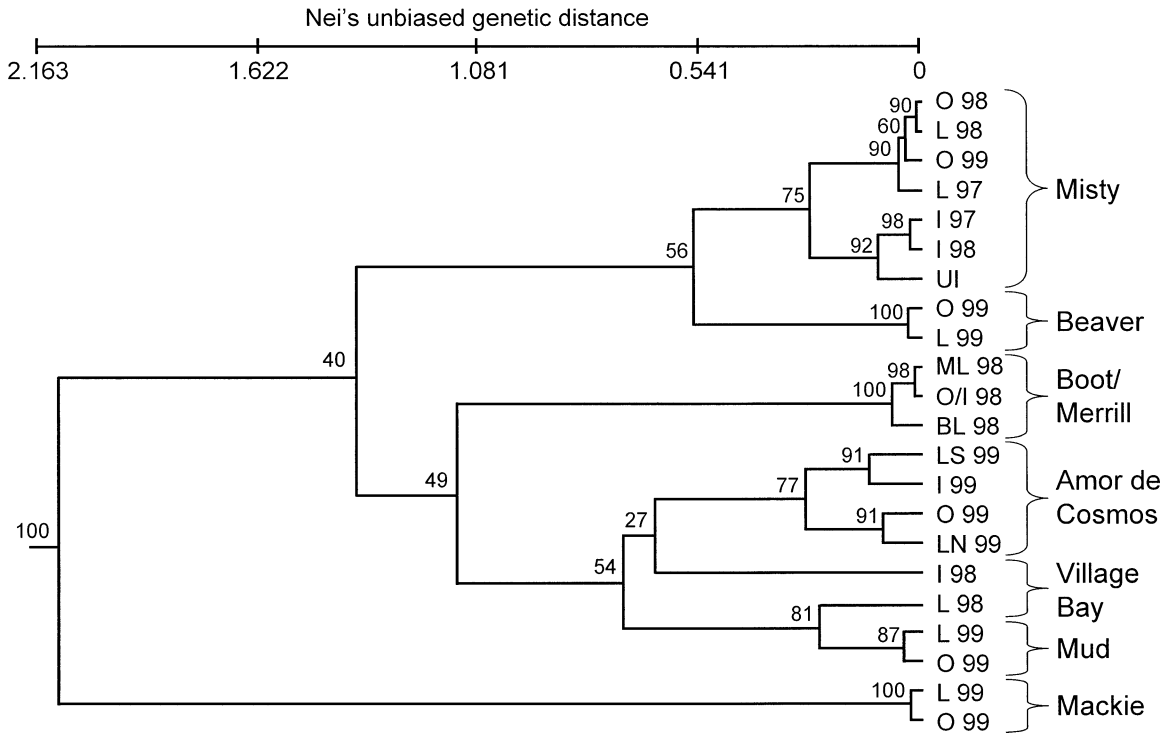

FIG. 3. A UPGMA dendogram of Nei's (1978) unbiased genetic distance (as implemented in TFPGA, Miller 1997) showing that collections are more closely related within than among watersheds. The figure also shows that different pairs within some systems (Pye and Amor de Cosmos) are effectively genetically independent. Labels at the end of terminal branches refer to specific collections, with a location abbreviation followed by the collection year. Standard location abbreviations are O, outlet; I, inlet; and L, lake. Additional abbreviations appear in the Misty watershed (UI, upper inlet), the Boot/Merrill watershed (ML, Merrill Lake; BL, Boot Lake; O/I, the stream connecting these two lakes), and the Amor de Cosmos watershed (LS, the southern McCreight Lake collection, which was adjacent to the inlet; LN, the northern McCreight Lake collection, which was adjacent to the outlet). One collection included here (Misty outlet 1999) was excluded from other analyses owing to a lack of morphological data. Bootstrap values are shown for all nodes.

= 0.697, $P = 0.012$; m, $r = 0.645$, $P = 0.024$), and some that were not correlated (Wright–Takahata vs. Beerli: N_em, $r = 0.482$, $P = 0.112$; m, $r = 0.220$, $P = 0.492$). In general, gene flow was least for Misty Lake–Inlet and Village Bay Lake–Inlet, greatest for Misty Lake–Outlet and Merrill Lake–Outlet, and intermediate for the other pairs (Table 3).

Adaptive Divergence versus Gene Flow

All measures of adaptive divergence were positively associated with all measures of genetic divergence and negatively associated with all but one measure of gene flow (Table 4). This strong correspondence between the observed and predicted direction of associations (31 of 32 times) was highly significant ($\chi^2 = 28.13$, $P < 0.001$), and the number of significant regressions at $\alpha = 0.05$ (seven) was greater than expected by chance (binomial likelihood of seven significant P-values by chance = 0.001). Other results corroborate these findings. First, all measures of adaptive divergence showed a negative trend with other measures of gene flow that we did not select a priori: that is, N_em based on the Slatkin and Beerli methods. Second, the regressions were similar, with some being stronger and some weaker, if we included all 12 paired collections, rather than just the eight independent pairs. Only one collection (Misty Outlet) was an obvious outlier, with greater adaptive divergence than expected for some measures of gene flow (Fig. 4). This consistent agreement between observed and predicted trends, despite dramatic differences in estimation methods, strongly suggests that adaptive divergence is indeed negatively correlated with gene flow.

Although the direction of observed trends was consistent, their strength and significance varied greatly depending on the measure of adaptive divergence, genetic divergence, and gene flow (Table 4). For adaptive divergence, body depth consistently generated the strongest ($r^2 = 0.361$–0.741) and most significant regressions (Table 4, Fig. 4). Eigenvalues based on discriminant functions that included body length generated the weakest regressions. For genetic divergence and gene flow, ρ_{ST} usually generated the strongest ($r^2 = 0.388$–0.645) and most significant regressions (Table 4), whereas Slatkin's method always generated the weakest regressions.

Most gene flow between parapatric lake and stream populations is probably from lakes into streams, simply because lake stickleback are much more numerous than stream stickleback. If gene flow constrains adaptive divergence, we might therefore expect less divergence for lake-outlet pairs than for lake-inlet pairs, simply because lake stickleback should be more likely to move downstream than upstream. Indeed, lake (but not stream) stickleback move downstream (but not upstream) when placed in a stream environment (Hendry et al. 2002). Consistent with this expectation, adaptive divergence

TABLE 2. Genetic divergence between lake and stream fish for each pair of collections. Divergence measures include Nei's (1978) unbiased genetic distance, Reynolds et al.'s (1983) coancestry coefficient, Weir and Cockerham's (1984) θ estimate of F_{ST}, and Rousset's (1996) ρ_{ST}. Assignment success within each pair (percent correctly classified to their collection location) was based on the Rannala and Mountain (1997) Bayesian approach with a threshold of 0.05, as implemented in GeneClass 2 (Piry et al. 2004).

Lake	Stream	Year	Nei's *D*	Reynold's coancestry	F_{ST} (θ)	ρ_{ST}	Assignment success (%)
Misty	Upper inlet	1999[1]	0.349	0.366	0.304[2]	0.304	95.7
Misty	Lower inlet	1997	0.180	0.142	0.134[2]	0.007	96.0
Misty	Lower inlet	1998	0.196	0.151	0.137[2]	0.230	93.3
Misty	Outlet	1998	0.008	0.007	0.005	0.005	62.7
Boot	Inlet	1998	0.059	0.046	0.040[2]	0.039	73.3
Merrill	Outlet	1998	0.009	0.013	0.009	−0.017	67.3
Beaver	Outlet	1999	0.026	0.033	0.028	0.067	59.6
McCreight (south)	Inlet	1999	0.121	0.051	0.044[2]	0.123	86.7
McCreight (north)	Outlet	1999	0.083	0.026	0.022	0.002	70.0
Mackie	Outlet	1999	0.027	0.083	0.078[2]	0.031	76.3
Village Bay (May)	Inlet	1998	0.618	0.210	0.187[2]	0.169	94.4
Mud	Outlet	1999	0.035	0.015	0.014[2]	−0.007	76.7

[1] Misty upper inlet data from 1997 ($N = 10$) and 1999 ($N = 30$) were combined and compared to Misty Lake data for 1997 and 1998. Averages of these two comparisons are shown.
[2] Bootstrapped 95% confidence intervals do not include zero.

appeared qualitatively lower for outlet fish (Table 1, Fig. 2), but the difference from lake-inlet pairs was not significant ($P > 0.05$ in *t*-tests based on all pairs in Table 1). The lack of significance in this comparison was due to substantial overlap between lake-inlet and lake-outlet pairs in the region of moderate divergence (Fig. 2).

As noted above, armor traits (e.g., lateral plates and pelvic spines) are not appropriate for testing whether gene flow constrains lake-stream divergence. In fact, we would not expect a significant correlation between morphological divergence and gene flow for these traits. An exploratory analysis confirmed this assertion. Specifically, lateral plates and pelvic spines did not differ in a consistent direction between lakes and streams. Moreover, lake-stream armor divergence was not significantly correlated with genetic divergence or gene flow ($P > 0.27$ for all regressions; except for Slatkin's *m* vs. gill rakers, $P = 0.032$).

Divergent Selection versus Gene Flow

Variation among lake-stream pairs in the strength of divergent selection does not appear to influence gene flow. Specifically, none of the four measures of genetic divergence or the four measures of gene flow was significantly correlated with lake surface area (all $r^2 < 0.125$ and all $P > 0.388$), stream size PC1 (all $r^2 < 0.279$ and all $P > 0.177$), water flow (all $r^2 < 0.275$ and all $P > 0.164$), canopy openness (all $r^2 < 0.202$ and all $P > 0.264$), or lake surface area *Z* scores minus stream size PC1 (all $r^2 < 0.339$ and all $P > 0.130$). Given the reasonable success of gene flow in explaining adaptive divergence (above), but the failure of divergent selection in explaining gene flow (here), we conclude that adaptive divergence in our system is likely a function of gene flow and not vice versa.

DISCUSSION

How much of the variation in adaptive divergence can be explained by gene flow? Our study suggests a wide range of possible answers (0.3–74.1%), with the specific amount depending on the measure of adaptive divergence and of gene flow. Despite this wide range, observed correlations were nearly always (31 of 32 times) in the expected direction: adaptive divergence was positively correlated with genetic

TABLE 3. Gene flow estimates between lake and stream fish for each pair of collections. Details on the various estimation methods are provided in the text. Takahata's N_em and *m* can be obtained by dividing Wright's N_em and *m* by a factor of four.

Lake	Stream	Year	Wright's N_em	Slatkin's N_em	Beerli's N_em	Wright's *m*	Slatkin's *m*	Beerli's *m*
Misty	Upper inlet	1999[1]	0.579	1.838	1.842	0.00027	0.00086	0.00034
Misty	Lower inlet	1997	1.621	1.430	2.337	0.00060	0.00053	0.00030
Misty	Lower inlet	1998	1.579	2.127	1.510	0.00059	0.00079	0.00018
Misty	Outlet	1998	49.750	4.554	2.257	0.01762	0.00161	0.00026
Boot	Inlet	1998	5.938	2.508	1.038	0.00197	0.00083	0.00019
Merrill	Outlet	1998	28.159	3.638	2.837	0.01218	0.00157	0.00039
Beaver	Outlet	1999	8.841	3.039	1.891	0.00429	0.00148	0.00041
McCreight (south)	Inlet	1999	5.393	2.172	2.128	0.00100	0.00040	0.00034
McCreight (north)	Outlet	1999	10.911	4.070	2.622	0.00151	0.00056	0.00022
Mackie	Outlet	1999	2.968	1.110	0.888	0.00380	0.00142	0.00054
Village Bay (May)	Inlet	1998	1.086	0.588	0.350	0.00023	0.00013	0.00006
Mud	Outlet	1999	17.607	1.945	1.958	0.00334	0.00037	0.00018

[1] Misty upper inlet data from 1997 ($N = 10$) and 1999 ($N = 30$) were combined and compared to Misty Lake data for 1997 and 1998. Averages of these two comparisons are shown.

TABLE 4. Linear regressions between various measures of genetic divergence or gene flow (x-axis, taken from Tables 2, 3) and various measures of adaptive divergence (y-axis, taken from Table 1). Values shown are coefficients of determination (r^2). The sign before each r^2 value indicates the direction of the correlation. All regressions were based on the eight independent lake/stream pairs.

	Body depth	Gill rakers	Eigenvalue	Eigenvalue (w/o length)
Nei's unbiased distance	+0.486*	+0.176	+0.168	+0.199
Reynold's coancestry	+0.489*	+0.179	+0.185	+0.202
F_{ST} (θ)	+0.638**	+0.235	+0.261	+0.242
ρ_{ST}	+0.645**	+0.517**	+0.388*	+0.523**
N_em (Wright-Takahata)	−0.498**	−0.181	−0.175	−0.203
m (Wright-Takahata)	−0.741***	−0.283	−0.060	−0.163
m (Slatkin)	−0.361	−0.118	+0.009	−0.003
m (Beerli)	−0.681**	−0.300	−0.060	−0.120

* $P < 0.10$, ** $P < 0.05$, *** $P < 0.01$.

divergence and negatively correlated with gene flow (Table 4). This consistency was impressive given the diverse suite of estimation methods, large genetic differences among the lake-stream pairs (Fig. 3), and a wide variety of parapatric contexts (Fig. 1). We conclude that adaptive divergence is indeed negatively correlated with gene flow, but that the strength of this association is uncertain and perhaps inherently variable. Although gene flow thus appears important in constraining adaptive divergence, the large amount of unexplained variation implies that other factors are also important. The most obvious of these factors is potential variation among pairs in the strength of divergent selection.

To attempt more precise statements about the general importance of gene flow, we must narrow the range of estimates for the amount of variation explained. To this end, we might first argue for the exclusion of the eigenvalues that included body length—because body size is plastic in stickleback. We might next argue that the most appropriate way to express gene flow when examining its effects on adaptive divergence is to use the proportion of migrants exchanged, m (Hendry et al. 2001). We might further argue that the least reliable estimate of m is that based on Slatkin's (1985) rare alleles method because: (1) it is difficult to precisely estimate the frequency of rare alleles (Slatkin and Barton 1989); (2) it

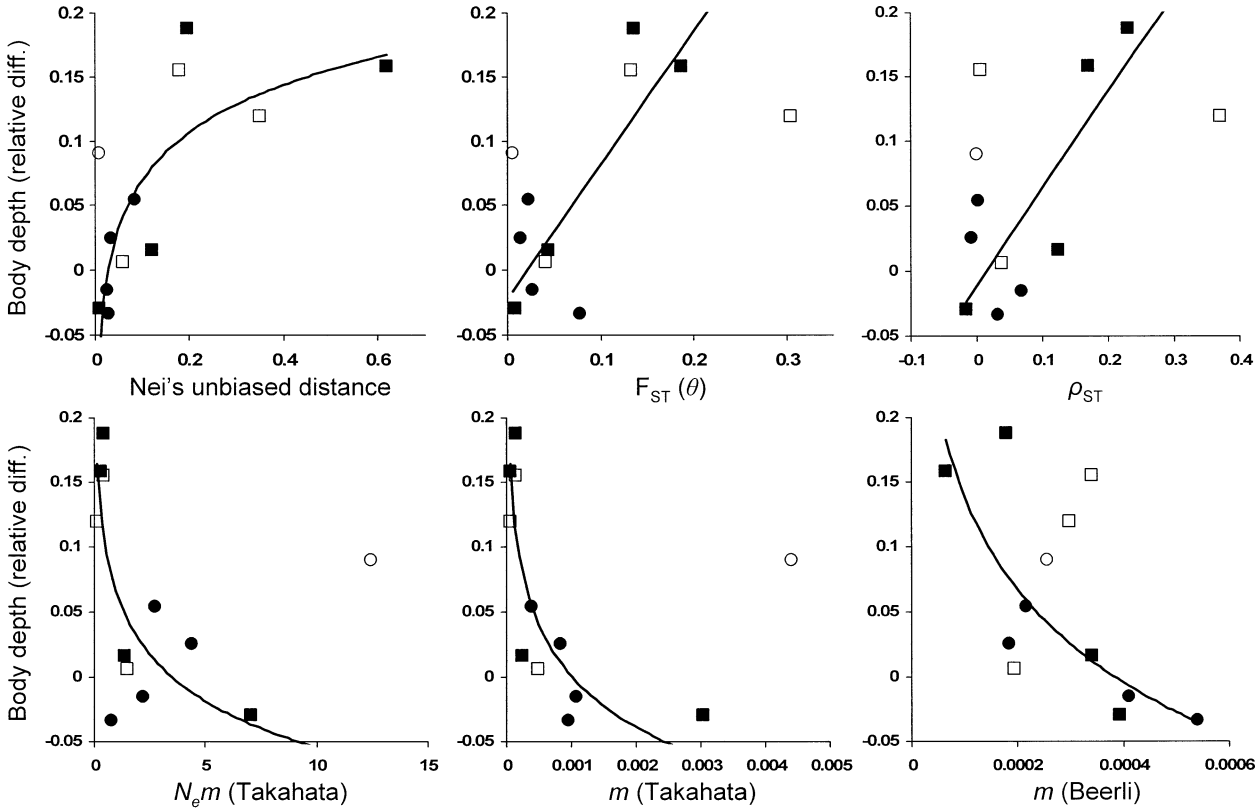

FIG. 4. Lake-stream divergence for size-standardized body depth is positively correlated with genetic divergence and negatively correlated with gene flow. Data are taken from Tables 2 and 3 and regression statistics are shown in Table 4. Squares are lake-outlet pairs and circles are lake-inlet pairs. Open symbols (squares and circles) are pairs excluded from analyses, including the regression lines shown here, owing to a lack of independence from other pairs. Note that the axes do not cross at the origin.

was the most sensitive method to the inclusion or exclusion of locus *Gac*9 (which seems to have null alleles); and (3) it always showed the weakest correlations with adaptive divergence (Table 4). All of these exclusions only narrow the range of explained variation to a still high 6.0–74.1% (Table 4). A substantial fraction of the remaining range was the result of differences among methods for estimating gene flow. This was not surprising because the two remaining methods (Wright-Takahata and Beerli) produced gene flow estimates that were not significantly correlated (see Results). Unfortunately, a theoretical and empirical consensus as to the best method for estimating gene flow is currently lacking (Slatkin and Barton 1989; Gaggiotti et al. 1999; Whitlock and McCauley 1999; Abdo et al. 2004).

In our foregoing attempt to make a general statement about the importance of gene flow, we did not separate the different measures of adaptive divergence. Different traits, however, will be differentially affected by natural selection, perhaps making it more informative to examine individual adaptive traits. In particular, correlations between adaptive divergence and gene flow should depend on the strength of divergent selection acting on a trait, even if this strength is consistent across all population pairs. We confirmed this expectation by exploring various combinations of divergent selection and gene flow in Hendry et al.'s (2001) theoretical model (results not shown). In the present study, correlations between morphological divergence and either genetic divergence or gene flow were considerably higher for body depth (36.1–74.1%) than for gill raker number (11.8–51.7%). If we again use only the Wright/Takahata and Beerli *m* estimates (see above), the ranges become 68.1–74.1% for body depth and 28.3–30.0% for gill raker number. Moreover, this difference between body depth and gill raker number is consistent for each individual measure of genetic divergence or gene flow (Table 4). Relationships of the strength we found for body depth (up to 74.1%), and perhaps also gill raker number (up to 51.7%), should have considerable biological relevance because lake-stream divergence in these traits is adaptive and heritable (see introduction).

Our study thus generated two principal conclusions. First, gene flow can appreciably constrain adaptive divergence across a broad range of natural populations. Previous studies have sometimes invoked gene flow in an all-or-nothing sense: it either does or does not constrain divergence. This narrow view obscures the reality that adaptive divergence represents a balance between gene flow and selection, with different populations arrayed along a continuum between the extremes of complete adaptive divergence versus no adaptive divergence. Our results confirm that gene flow can hold many populations short of complete adaptation, but to varying degrees (see also King and Lawson 1995; Hendry et al. 2002; Calsbeek and Smith 2003; Saint-Laurent et al. 2003). Second, different traits do not respond in the same way to the same amount of gene flow, probably because they are subject to different strengths of divergent selection. For lake and stream stickleback, gene flow has a much stronger effect on body depth than on gill raker number. Future work would profit from additional information on how divergent selection varies among traits.

Evaluating an Alternative: Ecological Speciation

An alternative explanation for negative correlations between gene flow and adaptive divergence is that the latter, which reflects divergent selection, constrains the former. This is the process of ecological speciation, which has considerable support from many taxa (for a review see Schluter 2000) including threespine stickleback (McPhail 1994; Schluter 2000; Reusch et al. 2001; McKinnon and Rundle 2002; McKinnon et al. 2004). Although we do not deny that this process is important for some traits in some taxa and populations, we have yet to find evidence for it in lake and stream stickleback. In our most direct appraisal, none of our measures of genetic divergence or gene flow was even remotely correlated with any of our measures of ecologically relevant habitat features, which should have been the case if adaptive divergence constrains gene flow (e.g., Smith et al. 1997; Schluter 2000; Ogden and Thorpe 2002). This negative result is robust because it contrasts with the much stronger observed correlations between adaptive divergence and gene flow, which were obtained using the same lake-stream pairs and the same measures of genetic divergence and gene flow. We conclude that adaptive divergence does not appreciably constrain gene flow in our system. However, we also acknowledge that our surrogates for divergent selection are quite crude and that the final say on ecological speciation in this system will require more refined surrogates (e.g., types and availability of various prey items) or direct measures of selection.

Gene flow thus appears to constrain adaptive divergence for lake and stream stickleback and not vice versa. Several lines of indirect evidence also support this conclusion. First, the only other correlative study to explicitly distinguish between the two causal alternatives (in *Timema* walking-sticks, Nosil and Crespi 2004) concluded that the arrow of causality flew from gene flow to morphological divergence and not vice versa, at least for the traits examined. Second, gene flow appears to constrain adaptive divergence in other stickleback populations (e.g., Bell and Richkind 1981; Bell 1982; Bourgeois et al. 1994). Third, our previous work in the Misty watershed revealed that gene flow plays an important causal role, even if ecological speciation makes some additional contribution. For example, the potential for physical dispersal from the lake (much higher into the outlet than into the inlet) was positively correlated with the amount of gene flow and negatively correlated with morphological divergence (Hendry et al. 2002).

In truth, the most realistic biological scenario is that the two causal pathways act at the same time and in a self-reinforcing feedback loop. That is, limited gene flow allows some adaptive divergence, which then causes a reduction in gene flow, which allows further adaptive divergence, which further reduces gene flow, and so on until some quasi-equilibrium is reached. These feedback effects may differ among traits depending on how closely each is tied to prezygotic and postzygotic reproductive isolation. It will be important for future work on the evolution of biological diversity to disentangle and quantify the interacting effects of natural selection, gene flow, and adaptive divergence.

If adaptive divergence does not influence gene flow, what

then are the causes of the dramatic differences among pairs in the amount of gene flow? The primary factors are probably the relative sizes of lake and stream populations, the direction of water flow, and the presence of any partial barriers to dispersal. In general, we expect that most gene flow will be from lakes into streams because lake populations are much larger than stream populations. In this case, gene flow into streams should increase as the size of the lake population increases relative to the stream population. We cannot directly assess this prediction because the sizes of our lake and stream populations are unknown. Given constraints on swimming ability, gene flow should also be higher in the downstream direction (into outlets) than in the upstream direction (into inlets). Indeed, this was the case in the present study because one-tailed *t*-tests based on the eight independent pairs showed that lake-inlet pairs had significantly higher genetic divergence ($P < 0.05$ for all four measures) and marginally lower gene flow ($P < 0.10$ for all four measures) than lake-outlet pairs. We also expect that gene flow into streams will be influenced by partial barriers to gene flow (such as beaver dams) but these too have yet to be quantified.

Implications

The potential role of gene flow should be evaluated when interpreting adaptive divergence in nature. Previous work has validated this assertion for specific populations (see above), and our work extends it more generally (see also Calsbeek and Smith 2003). We selected independent population pairs essentially at random from a larger suite of possible pairs, with the only proviso being a potential for dispersal within a pair (i.e., no major physical barriers). The two environments in each pair were very different (lake vs. stream), suggesting that divergent selection on foraging-related traits should be ubiquitous. And yet lake and stream stickleback varied dramatically in the extent of their adaptive divergence, which was least (sometimes nonexistent) for the very pairs that experienced the most gene flow. Owing to the conservative nature of our sampling and analyses, these results should generalize to hundreds of other situations where lake and stream stickleback live in parapatry. Our overall conclusions may also extend to many other natural systems in which gene flow is possible. In short, the perception of some authors that selection routinely overwhelms gene flow in nature (e.g., Ehrlich and Raven 1969; Levin 1979) is demonstrably untenable (see also Calsbeek and Smith 2003; Morjan and Rieseberg 2004).

Acknowledgments

Advice on study design and implementation was provided by J. Boughman, S. Latham, H. Rundle, D. Schluter, and S. Vamosi. Fieldwork was assisted by J. Akins, S. Anderson, M. Hendry, N. Millar, J.-S. Moore, K. Räsänen, and H. Roffey. Preserved fish were measured by C. Logtenberg and S. Shih. Additional help with the manuscript was provided by O. Bargelletti, C. Benkman, E. Crispo, N. Millar, J.-S. Moore, K. Räsänen, H. Roffey, and two anonymous reviewers. Special thanks go to E. Crispo and J.-S. Moore, who performed much of the initial statistical analysis. We also thank B. Robinson, whose question during a presentation by A. Hendry increased our interest in the arrow of causality. Financial support was provided by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant to EBT and NSERC Postdoctoral and Discovery Grants to APH.

Literature Cited

Abdo, Z., K. A. Crandall, and P. Joyce. 2004. Evaluating the performance of likelihood methods for detecting population structure and migration. Mol. Ecol. 13:837–851.

Beerli, P. 2004. Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. Mol. Ecol. 13:827–836.

Beerli, P., and J. Felsenstein. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics 152:763–773.

Bell, M. A. 1982. Differentiation of adjacent stream populations of threespine sticklebacks. Evolution 36:189–199.

Bell, M. A., and S. A. Foster. 1994. Introduction to the evolutionary biology of the threespine stickleback. Pp. 1–27 *in* M. A. Bell and S. A Foster, eds. The evolutionary biology of the threespine stickleback. Oxford Univ. Press, Oxford, U.K.

Bell, M. A., and K. E. Richkind. 1981. Clinal variation of lateral plates in threespine stickleback fish. Am. Nat. 117:113–132.

Blondel, J., P. C. Dias, P. Perret, M. Maistre, and M. M. Lambrechts. 1999. Selection-based biodiversity at a small spatial scale in a low-dispersing insular bird. Science 285:1399–1402.

Bourgeois, J. F., D. M. Blouw, J. P. Koenings, and M. A. Bell. 1994. Multivariate analysis of geographic covariance between phenotypes and environments in the threespine stickleback, *Gasterosteus aculeatus*, from the Cook Inlet area, Alaska. Can. J. Zool. 72:1497–1509.

Calsbeek, R., and T. B. Smith. 2003. Ocean currents mediate evolution in island lizards. Nature 426:552–555.

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

Endler, J. A. 1977. Geographical variation, speciation, and clines. Princeton Univ. Press, Princeton, NJ.

Englund, S. R., J. J. O'Brien, and D. B. Clark. 2000. Evaluation of digital and film hemispherical photography and spherical densiometry for measuring forest light environments. Can. J. Forest Res. 30:1999–2005.

Gaggiotti, O. E., O. Lange, K. Rassmann, and C. Gliddon. 1999. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. Mol. Ecol. 8: 1513–1520.

García-Ramos, G., and M. Kirkpatrick. 1997. Genetic models of adaptation and gene flow in peripheral populations. Evolution 51:21–28.

Gíslason, D., M. M. Ferguson, S. Skúlason, and S. S. Snorrason. 1999. Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (*Salvelinus alpinus*). Can. J. Fish. Aquat. Sci. 56:2229–2234.

Goudet, J. 1995. Fstat version 1.2: a computer program to calculate *F*-statistics. J. Hered. 86:485–486.

Haldane, J. B. S. 1948. The theory of a cline. J. Genet. 48:277–284.

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

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

Hersch, E. I., and P. C. Phillips. 2004. Power and potential bias in field studies of natural selection. Evolution 58:479–485.

King, R. B., and R. Lawson. 1995. Color-pattern variation in Lake Erie water snakes: the role of gene flow. Evolution 49:885–896.

Langerhans, R. B., C. A. Layman, A. K. Langerhans, and T. J. Dewitt. 2003. Habitat-associated morphological divergence in two Neotropical fish species. Biol. J. Linn. Soc. 80:689–698.

Lavin, P. A., and J. D. McPhail. 1985. The evolution of freshwater

diversity in the threespine stickleback (*Gasterosteus aculeatus*): site-specific differentiation of trophic morphology. Can. J. Zool. 63:2632–2638.
———. 1986. Adaptive divergence of tropic phenotype among freshwater populations of the threespine stickleback (*Gasterosteus aculeatus*). Can. J. Fish. Aquat. Sci. 43:2455–2463.
———. 1993. Parapatric lake and stream sticklebacks on northern Vancouver Island: disjunct distribution or parallel evolution? Can. J. Zool. 71:11–17.
Lemmon, P. E. 1957. A new instrument for measuring forest overstory density. J. Forest. 55:667–668.
Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends Ecol. Evol. 17:183–189.
Levin, D. A. 1979. The nature of plant species. Science 204: 381–384.
Lu, G., and L. Bernatchez. 1999. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. Evolution 53:1491–1505.
McKinnon, J. S., and H. D. Rundle. 2002. Speciation in nature: the threespine stickleback model systems. Trends Ecol. Evol. 17: 480–488.
McKinnon, J. S., S. Mori, B. K. Blackman, L. David, D. M. Kingsley, L. Jamieson, J. Chou, and D. Schluter. 2004. Evidence for ecology's role in speciation. Nature 429:294–298.
McPhail, J. D. 1994. Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of south-western British Columbia. Pp. 399–437 *in* M. A. Bell and S. A Foster, eds. The evolutionary biology of the threespine stickleback. Oxford Univ. Press, Oxford, U.K.
Miller, M. P. 1997. Tools for population genetic analyses (TFPGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author via http://bioweb.usu.edu/mpmbio/index.htm.
Møller, A. P., and M. D. Jennions. 2002. How much variance can be explained by ecologists and evolutionary biologists? Oecologia 132:492–500.
Moodie, G. E. E. 1972a. Morphology, life history, and ecology of an unusual stickleback (*Gasterosteus aculeatus*) in the Queen Charlotte Islands, Canada. Can. J. Zool. 50:721–732.
———. 1972b. Predation, natural selection and adaptation in an unusual threespine stickleback. Heredity 28:155–167.
Morjan, C. L., and L. H. Rieseberg. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Mol. Ecol. 13:1341–1356.
Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.
Nosil, P., and B. J. Crespi. 2004. Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and sexual isolation in *Timema cristinae* walking-sticks. Evolution 58: 102–112.
Ogden, R., and R. S. Thorpe. 2002. Molecular evidence for ecological speciation in tropical habitats. Proc. Natl. Acad. Sci. USA 99:13612–13615.
Peek, M. S., A. J. Leffler, S. D. Flint, and R. J. Ryel. 2003. How much variance is explained by ecologists? Additional perspectives. Oecologia 137:161–170.
Piry, S., A. Alpetite, J.-M. Cornuet, D. Paetkau, L. Baudoin, and A. Estoup. 2004. GeneClass2: a software for genetic assignment and first-generation migrant detection. J. Hered. 95: *in press*.
Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA 94: 9197–9201.
Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenism. J. Hered. 86:248–249.
Reimchen, T. E. 1994. Predators and morphological evolution in the threespine stickleback. Pp. 240–276 *in* M. A. Bell and S. A Foster, eds. The evolutionary biology of the threespine stickleback. Oxford Univ. Press, Oxford, U.K.
Reimchen, T. E., E. M. Stinson, and J. S. Nelson. 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, J. D. 1986. An empirical evaluation of coefficients used in residual and allometric adjustment of size covariation. Can. J. Zool. 64:1363–1368.
Reusch, T. B. H., K. M. Wegner, and M. Kalbe. 2001. Rapid genetic divergence in postglacial populations of threespine stickleback (*Gasterosteus aculeatus*): the role of habitat type, drainage and geographical proximity. Mol. Ecol. 10:2435–2445.
Reynolds, J., B. S. Weir, and C. C. Cockerham. 1983. Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767–779.
Rico, C., D. Zadworny, U. Kuhnlein, and G. J. Fitzgerald. 1993. Characterization of hypervariable microsatellite loci in the threespine stickleback *Gasterosteus aculeatus*. Mol. Ecol. 2: 271–272.
Riechert, S. E. 1993. Investigation of potential gene flow limitation of behavioral adaptation in an aridlands spider. Behav. Ecol. Sociobiol. 32:355–363.
Rousset, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics 142: 1357–1362.
Saint-Laurent, R., M. Legault, and L. Bernatchez. 2003. Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchill). Mol. Ecol. 12:315–330.
Schluter, D. 2000. The ecology of adaptive radiation. Oxford Univ. Press, Oxford, U.K.
Slatkin, M. 1973. Gene flow and selection in a cline. Genetics 75: 733–756.
———. 1985. Rare alleles as indicators of gene flow. Evolution 39:53–65.
———. 1987. Gene flow and the geographic structure of natural populations. Science 236:787–792.
Slatkin, M., and N. H. Barton. 1989. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349–1368.
Smith, T. B., R. K. Wayne, D. J. Girman, and M. W. Bruford. 1997. A role for ecotones in generating rainforest biodiversity. Science 276:1855–1857.
Storfer, A. 1999. Gene flow and endangered species translocations: a topic revisited. Biol. Cons. 87:173–180.
Storfer, A., and A. Sih. 1998. Gene flow and ineffective antipredator behavior in a stream-breeding salamander. Evolution 52: 558–565.
Storfer, A., J. Cross, V. Rush, and J. Caruso. 1999. Adaptive coloration and gene flow as a constraint to local adaptation in the streamside salamander, *Ambystoma barbouri*. Evolution 53: 889–898.
Takahata, N. 1983. Gene identity and genetic differentiation of populations in the finite island model. Genetics 104:497–512.
Taylor, E. B. 1998. Microsatellites isolated from the threespine stickleback *Gasterosteus aculeatus*. Mol. Ecol. 7:930–931.
Taylor, E. B., and J. D. McPhail. 2000. Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus*. Proc. R. Soc. Lond. B. Biol. Sci. 267: 2375–2384.
Thompson, C. E., E. B. Taylor, and J. D. McPhail. 1997. Parallel evolution of lake-stream pairs of threespine sticklebacks (*Gasterosteus*) inferred from mitochondrial DNA variation. Evolution 51:1955–1965.
Waples, R. S. 1991. Genetic methods for estimating the effective size of cetacean populations. Rep. Int. Whaling Comm. Spec. Issue 13:279–300.
Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358–1370.
Whitlock, M. C., and D. E. McCauley. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. Heredity 82: 117–125.

Corresponding Editor: C. Benkman