doi:10.1111/evo.12917

Cline coupling and uncoupling in a stickleback hybrid zone

Timothy H. Vines,[1,2,3] Anne C. Dalziel,[4] Arianne Y. K. Albert,[5] Thor Veen,[6] Patricia M. Schulte,[1] and Dolph Schluter[1]

[1]*Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada*

[2]*Axios Review Editorial Office, Vancouver, British Columbia, V5V 3X3, Canada*

[3]*E-mail: vines@zoology.ubc.ca*

[4]*Department of Biology, Saint Mary's University, Nova Scotia, B3H 3C3, Canada*

[5]*Women's Health Research Institute, Vancouver, British Columbia, V6H 3N1, Canada*

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

Received March 21, 2015
Accepted March 18, 2016

Strong ecological selection on a genetic locus can maintain allele frequency differences between populations in different environments, even in the face of hybridization. When alleles at divergent loci come into tight linkage disequilibrium, selection acts on them as a unit and can significantly reduce gene flow. For populations interbreeding across a hybrid zone, linkage disequilibria between loci can force clines to share the same slopes and centers. However, strong ecological selection on a locus can also pull its cline away from the others, reducing linkage disequilibrium and weakening the barrier to gene flow. We looked for this "cline uncoupling" effect in a hybrid zone between stream resident and anadromous sticklebacks at two genes known to be under divergent natural selection (*Eda* and *ATP1a1*) and five morphological traits that repeatedly evolve in freshwater stickleback. These clines were all steep and located together at the top of the estuary, such that we found no evidence for cline uncoupling. However, we did not observe the stepped shape normally associated with steep concordant clines. It thus remains possible that these clines cluster together because their individual selection regimes are identical, but this would be very surprising given their diverse roles in osmoregulation, body armor, and swimming performance.

KEY WORDS: ***Ectodysplasin***, ***Gasterosteus aculeatus***, **hybridization, linkage disequilibrium, reproductive isolation.**

The process of speciation is hindered by hybridization because gene flow erodes allele frequency differences between the populations, and because recombination in hybrids breaks down linkage disequilibrium (LD). In general, only alleles under strong selection or tightly linked to strongly selected loci are expected to show large allele frequency differences between hybridizing populations (Yeaman and Otto 2011; Yeaman and Whitlock 2011). When strongly selected alleles at different loci come into LD (i.e., are found together more often than predicted by their respective frequencies), selection tends to act on them as unit, reducing hybrid or immigrant fitness by their combined effects (Bazykin 1969; Barton 1983; Kruuk et al. 1999; Barton and de Cara 2009), which in turn helps to maintain divergence between the hybridizing populations. The key processes in achieving further divergence between hybridizing populations are thus (a) a growing number of loci with large allele frequency differences between the populations, and (b) ever stronger LD between the alleles at the diverged loci (Felsenstein 1981; Barton and de Cara 2009; Smadja and Butlin 2011; Flaxman et al. 2013, 2014).

One situation that is particularly informative about the interplay between divergent selection and linkage disequilibrium during the process of speciation is a hybrid zone, where parapatric populations interbreed at a shared boundary (Barton and Hewitt 1985; Harrison 1990). Multiple generations of hybridization between the adjacent populations typically produce allele frequency clines at the differentiated loci. The slope and position of each cline is a product of the balance between dispersal and total selection at that locus. Total selection is given by direct selection on

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

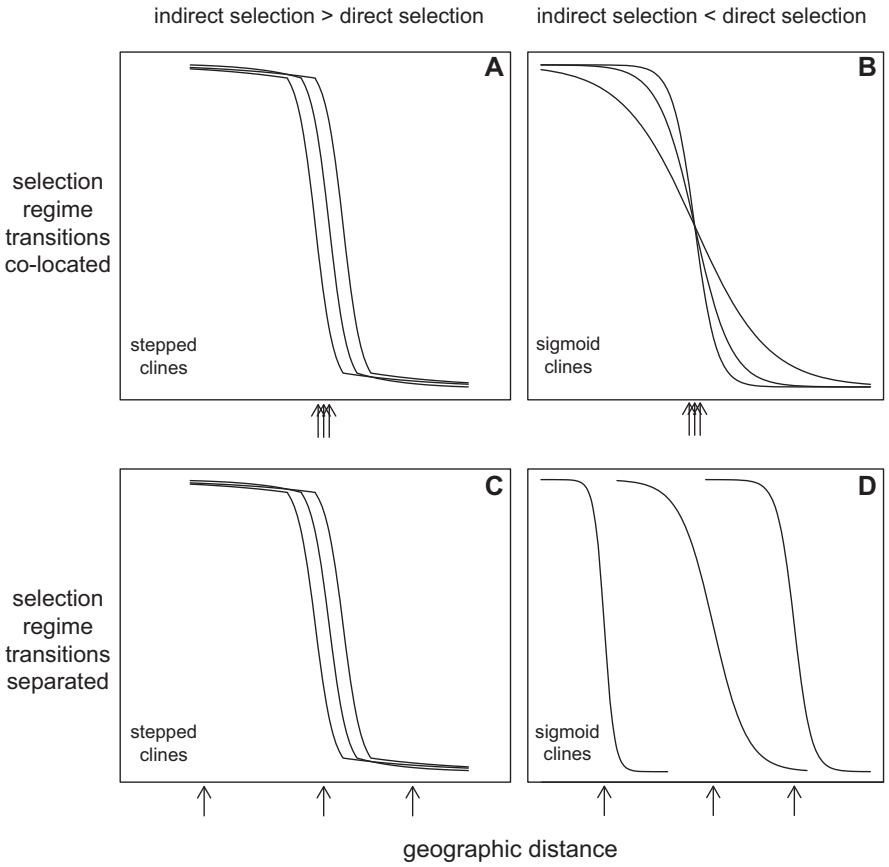

Figure 1. **Diagram showing how the spatial distribution of selection regime transitions combines with the relative strengths of direct and indirect selection to affect cline shape and location. Each panel shows three clines for loci under ecological selection. The arrows on the *x*-axis indicate the point at which the selection regime shifts from favoring the predominant allele from the right hand population to favoring the allele from the left population. (A and C) When there is strong linkage disequilibrium between the three loci, indirect selection forces the clines into concordance (i.e., having the same slope) and coincidence (having the same center), even when the selection regimes transition in different places. Since linkage disequilibria and hence indirect selection are strongest in the center of the zone, the clines are steeper than a sigmoid cline in the center, giving them a stepped shape. (B and D) When direct selection is stronger than indirect selection, the slope, and location of each cline reflects selection separately on that locus. If the selection regimes transition in the same place the clines will be coincident but not concordant, and since indirect selection is weak they are not expected to be stepped.**

the locus plus the sum of indirect selection on diverged loci elsewhere in the genome, weighted by the strength of LD between the focal locus and the other diverged loci (Barton 1983; Barton and Gale 1993). All loci for which direct selection is outweighed by indirect selection will experience approximately the same selection regime, and their clines will have similar slopes and centers (Fig. 1A). The clustering of clines in turn maintains strong LD, as "pure" individuals will have alleles characteristic of one or other population at all of the clustered loci (Slatkin 1975; Barton 1983). Clines held together by strong indirect selection thus tend to be steeper than expected in the zone center (where LD is strongest), and the clines are "stepped" (Szymura and Barton 1991).

By contrast, when direct selection on a locus is stronger than indirect selection, cline shape will reflect the regime imposed by selection on that locus alone (Fig. 1B, D). If direct selection has an ecological component (e.g., different alleles are favored on either side of an environmental transition), the cline may occur away from the clines at other loci. This scenario is especially likely when the hybrid zone straddles an ecotone (where there are step changes in many environmental variables), but the transition points for a few variables are close to but do not coincide with the ecotone (Fig. 1D).

Strong direct selection thus plays two potentially opposing roles: it ensures that allele frequency differences persist despite gene flow (Yeaman and Otto 2011), but it may also pull clines apart, reducing LD and thereby weakening the total indirect selection felt by the remaining loci (Nürnberger et al. 1995). Evidence for this "cline uncoupling" effect of strong direct selection is obtained when steep clines for genetic loci or phenotypic traits known to be under selection are close to one another but do not

share cline centers. Conversely, finding that clines for loci under varying types of selection are centered together suggests either that direct selection on each locus has been overwhelmed by indirect selection, or that the selection regimes for these loci all transition in the same location. Data on cline shape can help to distinguish between these latter two scenarios: concordant, stepped clines suggest that indirect selection predominates (Fig. 1A, C), while finding a range of cline slopes indicates the selection regime for each locus is the most important determinant of cline shape (Fig. 1B). One final possibility is that the clines are unstepped and have similar slopes and centers; this scenario is only expected if their selection regimes colocate and, despite their varying functions, they independently experience the same strength of direct selection.

Genetic divergence under gene flow has been well studied, both theoretically (Wright 1931; Bulmer 1972; Yeaman and Otto 2011; Yeaman and Whitlock 2011) and empirically (reviewed in Smadja and Butlin 2011; Feder et al. 2013). By contrast, the circumstances under which clines are situated together or apart are less well studied, either theoretically (with the exception of Slatkin 1975; Nürnberger et al. 1995; Bierne et al. 2011; Rosser et al. 2014) or empirically (Nürnberger et al. 1995; Rosser et al. 2014, see Abbott et al. 2013 for a review).

Here, we look for evidence of cline uncoupling in a natural hybrid zone by examining clines for two selected loci, five ecologically relevant morphological traits, and 13 anonymous SNPs in a zone between two differentially adapted threespine stickleback (*Gasterosteus aculeatus*) populations. In this hybrid zone, one population ("stream") inhabits the upper reaches of a small river, Bonsall Creek, throughout the year. The other ("anadromous") spends most of the year in the sea and migrates into the lower half of the river in the spring to breed (Hagen 1967). Because anadromous adults are only present in the estuary during the breeding season, the hybrid zone is transitory, forming and reforming each breeding season.

The hybrid zone between the two types is located close to the ecotone between the estuary and the freshwater environment. The estuary itself represents a gradient for many environmental variables, including salinity, tidal range, vegetation, and the community of predators, parasites, and competitors. These gradients end abruptly with the transition to freshwater. Sticklebacks can easily swim from one side of the zone to the other (T.H. Vines, A.Y.K. Albert, and A.C. Dalziel, unpub. data), so it is likely that the distribution of genotypes within the creek results from an active choice of breeding location along the salinity gradient. Moreover, lab crosses between anadromous and stream fish from Bonsall Creek show no evidence of intrinsic hybrid inviability or infertility (Dalziel et al. 2012; Dalziel and Schulte 2012), and it seems likely that the majority of reproductive isolation in the hybrid zone is either extrinsic or occurs prior to fertilization.

The two selected genes we examine are *Ectodysplasin* (hereafter *Eda*), and *Na^+, K^+ ATPase*'s catalytic α1 subunit (hereafter *ATP1a1*). *Eda* is located on chromosome IV and is responsible for one of the major phenotypic differences between these populations: stream fish have 4–8 lateral armor plates per side, whereas the anadromous fish have 30–35 plates (Hagen 1967). The low plated allele at *Eda* has risen to fixation in many freshwater environments (Colosimo et al. 2005), and shows evidence of repeated selection (Mäkinen et al. 2008; Shimada et al. 2011; DeFaveri et al. 2011; Jones et al. 2012; Raeymaekers et al. 2014). *ATP1a1* is located on chromosome I, and is part of a multisubunit, membrane bound enzyme, Na^+, K^+ ATPase, that maintains electrochemical gradients by moving K^+ ions into and Na^+ ions out of the cell (reviewed by Kaplan 2002). In fish, Na^+, K^+ ATPase plays a critical role in osmoregulation in both fresh and salt water (reviewed by Evans et al. 2005). Like *Eda*, the *ATP1a1* isoform has an allele that has repeatedly risen to high frequency in many freshwater stickleback populations (Jones et al. 2006; Shimada et al. 2011) and selection is related to salinity (Hohenlohe et al. 2010; DeFaveri et al. 2011, 2013; Shimada et al. 2011; Jones et al. 2012; Terekhanova et al. 2014).

All of the morphological traits we examined have evolved repeatedly in the same direction after sticklebacks colonized freshwater from the sea, and are predicted to be under strong selection. Anadromous populations have long pelvic and dorsal spines that likely defend against gape-limited predators, while freshwater populations typically have shorter spines, possibly to limit insect predation (Bell et al. 1993; Reimchen 1994; Marchinko 2009). Anadromous fish also have large, long fins that are capable of powering prolonged swimming during migration. Their smaller caudal peduncles are predicted to streamline the fish and reduce drag (Dalziel et al. 2012). Stream fish have repeatedly evolved smaller pectoral fins, which may help fish maneuver in smaller spaces, and deeper caudal peduncles that are predicted to increase burst swimming capacity (Taylor and McPhail 1986). The median fins (which includes the dorsal fin) are involved in maneuvering and force generation during steady swimming (Lauder et al. 2002), and have also repeatedly become reduced in lakes without predatory fish (Walker 1997; Walker and Bell 2000). A study by Dalziel et al. (2012) with anadromous and stream fish from Bonsall Creek found that the differences in fin morphology persist when both are raised in a common environment.

The selection regimes for alleles and traits advantageous in either anadromous or stream fish will not necessarily transition in the same part of the river. For example, selection on loci involved in osmoregulation, such as *ATP1a1*, may favor the stream allele where the stream bed salinity is less than isosmotic (~13 ppt) and the anadromous allele at higher salinities (Shimada et al. 2011; DeFaveri et al. 2011). By contrast, stream alleles at loci underlying reductions in lateral plates (*Eda*) and other body armor

traits (pelvic and dorsal spine length) may only be favored in full freshwater (< 1 ppt), where insects become a significant part of the predator regime. Strong direct selection on these loci and traits would force their clines to be centered in different parts of the river. Our goal in this paper was to thus test whether the clines at body armor loci and traits (*Eda*, pelvic, and dorsal spine length), osmoregulation loci (*ATP1a1*) and morphological traits related to swimming capacity (pectoral and dorsal fin length, caudal peduncle depth) have different slopes and centers, which would in turn suggest that strong direct selection has uncoupled the clines at these loci (Fig. 1D).

By contrast, finding that these loci and traits have coincident clines suggests either that their selection regimes transition in the same part of the river (Fig. 1B), or that strong indirect selection has forced the clines to share the same slope and center, irrespective of where the selection regimes transition (Fig. 1A and C). In the latter case, indirect selection is expected to make the clines stepped. Distinguishing between these scenarios in this hybrid zone will shed light on the relative roles of direct and indirect selection in promoting (or inhibiting) the speciation process in stickleback.

Methods

DATA COLLECTION

We collected sticklebacks from Bonsall Creek (approximately 48°52′47.5"N 123°40′26.8"W) between 15th May and 8th June 2006. The creek is a relatively short (15 km) and narrow (~3 m wide) tidal river on the South Eastern end of Vancouver Island, British Columbia (see Hagen 1967). Anadromous stickleback migrate into Bonsall Creek in late spring. The majority remain in the estuary (below 2.35 km in Fig. 2), but a few move further up into freshwater. Fish that resemble phenotypically pure anadromous sticklebacks are rare beyond 2.7 km upstream. The stream resident population reaches its highest densities upstream of 3.3 km, but individuals with stream resident phenotypes are common to around 2.3 km and can occasionally be found as far downstream as 1.65 km from the sea. Hybrid fish with intermediate phenotypes can be found where the stream resident and anadromous populations overlap, but are most common between 2.2 km and 2.6 km. A similar distribution of genotypes was reported by Hagen (1967): "A large breeding congregation of *trachurus* [anadromous stickleback] was found 1.5 miles [2.4 km] from the estuary ... within 200 ft upstream ... *leiurus* [stream stickleback] alone was collected ... [in] the intervening section hybrids were common." We can thus assume that a narrow hybrid zone has persisted in Bonsall creek for at least 40 years.

We captured fish at 100–150 m intervals through the hybrid zone, using 5 minnow traps per site (Fig. 2). Site location is indicated by distance from the sea, in km. Fish over 40 mm were killed with an overdose of MS222 anesthetic and frozen on dry ice or preserved in ethanol. We used a YSI 85 Handheld Conductivity and Oxygen meter (YSI Inc., Ohio, USA) to measure the surface and creek bed salinity at each site at low tide and at the peak of the highest tide during sampling period (11 pm on 14th June 2006), and found the upstream limit of saltwater 2.35 km from the sea (at 48°52′48.46"N, 123°40′26.93"W).

MORPHOLOGICAL DATA

We selected ten morphological traits that have been found to diverge repeatedly between anadromous and freshwater sticklebacks (Hagen and Gilbertson 1972). Ten marine and ten stream individuals from each side of the hybrid zone were measured for standard length, body depth, head depth, eye diameter, dorsal fin length, caudal peduncle depth, pectoral fin length, caudal peduncle width, left pelvic spine length, and second dorsal spine length. Eye diameter, head depth, and body depth showed no consistent differences between marine and stream fish, even when each trait was regressed onto standard length to remove the effects of size, and these traits were discarded. The remaining seven traits were measured on the 428 ethanol preserved fish. We measured standard length and dorsal and pelvic fin length with calipers, the length of the second dorsal and pelvic spine and the caudal peduncle depth and width with an ocular micrometer fitted to a binocular microscope. As standard length showed a strong relationship with all traits, we took the residuals from a regression of each trait on standard length.

We discovered that variation in caudal peduncle width resulted from two confounding sources: anadromous fish had a bony "keel" that adds 1–2 mm to the total width of the peduncle, but this feature is not present in stream individuals. However, when the keel is ignored, stream fish had a wider caudal peduncle than anadromous fish. As the total width of the peduncle could not be expected to fit a simple cline model it was discarded at this stage.

GENETIC DATA

We extracted DNA from tail fin tissue from all 428 fish using the protocol in Peichel et al. (2001); the DNA was resuspended in 50 μl of double distilled water and stored at –20°C.

Ectodysplasin (Eda)

The *Eda* locus is located on linkage group IV. We chose a T/C SNP at position 421 of an amplicon spanning the 7th and 8th exon of the *Eda* gene (Colosimo et al. 2005) for genotyping. The primers are given in Table S1.

Na^+, K^+ ATPase subunit α1a (ATP1a1)

There are two *ATP1a1* paralogs in the stickleback genome, which are found in tandem within approximately 9000 bp of each other

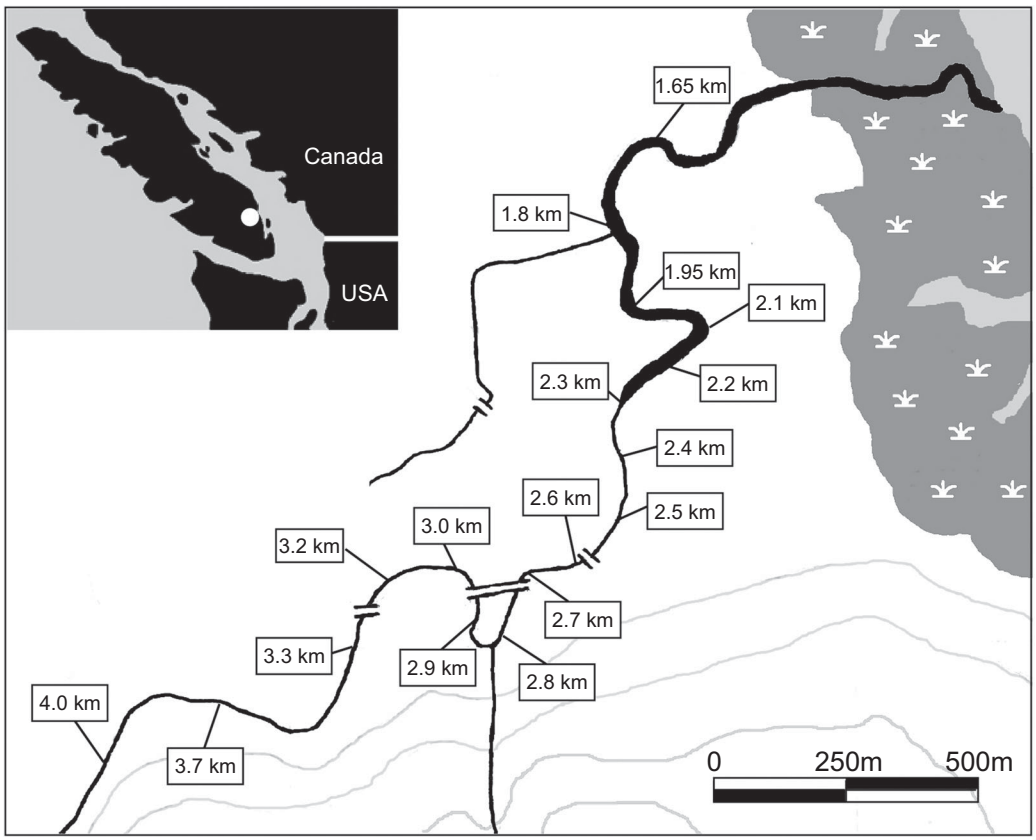

Figure 2. **Map of the Bonsall Creek study site on Vancouver Island (white dot on inset) and of the sampling locations along the river itself (main map), indicated by distance from the sea in kilometers. The gray shaded area represents salt marsh; salt water at high tide reaches 2.35 km into the creek, although there is some tidal fluctuation in water level at 2.4 km.**

on contigs 7064/7065 and 7066 within linkage group I. We studied the isoform on contig 7066, which is the same isoform studied by Jones et al. (2006) and Barrett et al. (2008). Our initial genotyping of the anadromous and stream alleles at the 16th to 18th intron of this gene followed the procedure in Jones et al. (2006), but as these sites are not variable in our populations we could not achieve allele-specific amplification. We therefore sequenced this genomic region in 10 pure stream (above 4.0 km) and 10 pure anadromous fish (from 1.65 km). We found a number of other diagnostic SNPs within this region, and selected an A/G SNP at position 446 of the amplicon (see Table S1 for sequences and primers).

Anonymous SNPs

We began with the list of 25 loci used to construct the phylogenetic tree in Colosimo et al. (Fig. 3C in Colosimo et al. 2005). We excluded loci on the same linkage group as *Eda* (LG IV) or *ATP1a1* (LG I), although loci with an unknown location were retained. After updates to the stickleback genome assembly (Jones et al. 2012), we found that one marker, P7E08, was also on LG I. We then tested the primers pairs from Colosimo et al. (2005) on 10 stream (4.0 km) and 10 anadromous (1.65 km) fish from Bonsall Creek. Of the 16 loci that amplified successfully in both populations, we retained 13 loci that had a SNP minor allele frequency greater than 0.1 in at least one of the samples. These loci are located on LGXI (P3A06 and P9D09), LGIII (P3D05), LGX (P4G01), LGVII (P6A10), LGXII (P6B12), LGIX (P6D05), LGXVIII (P7A07), and LGI (P7E08). Four others (P7A07, P7C08, P7G05, and P7H05) are on LGXVII (Glazer et al. 2015), although this was only identified as "Scaffold 27" when the genotyping was conducted (Colosimo et al. 2005).

These sequences of these 13 loci (Table S1) were used by the McGill University and Génome Québec Innovation center to design a custom genotyping assay using Sequenom® iPLEX®Gold Genotyping Technology. Genomic DNA from all 428 fish was quantified and diluted to 20 ng/μl and at least 30 μl of each sample were sent on dry ice to the McGill University and Génome Québec Innovation center for genotyping.

STATISTICAL METHODS

Cline fitting: Genetic loci

The clines at *Eda* and *ATP1a1* and the anonymous SNP loci were characterized with the program CFit version 7 (Gay et al. 2008). This program uses a Metropolis algorithm to find the cline

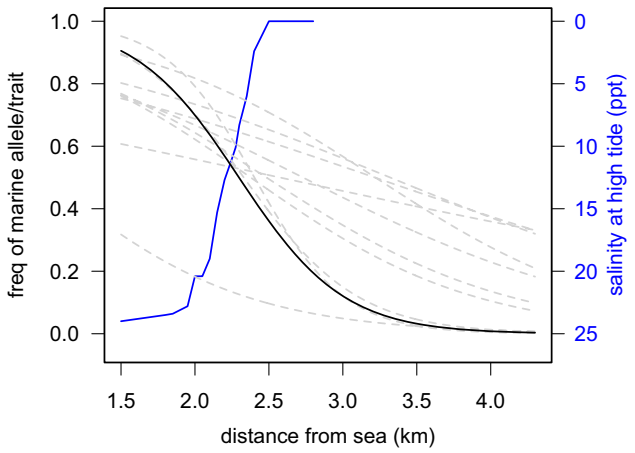

Figure 3. The best-fit clines in the hybrid zone. Salinity at high tide (blue line) is plotted on an inverse scale on the right hand axis. The black line shows the joint cline *Eda*, *ATP1a1*, and the five morphological traits, and the dashed gray lines the clines for the 10 anonymous SNPs.

shape that best fits the available data. The program fits two basic parameters that describe a sigmoid cline, cline slope l, and cline center c, and also allows for deviations from Hardy–Weinberg equilibrium (fitted as F_{IS}) within each site. Initial tests found no support for including F_{IS} at any of the genetic loci and we do not consider it further. The slope output parameter from CFit 7 (denoted l here) must be divided by 4 to give the slope parameter used in other hybrid zone work (e.g., Barton and Gale 1993). We use s ($= l/4$) to denote this more usual slope parameter, and present s wherever possible.

We also fitted a more complex stepped cline model, where the center part of the cline remains sigmoid, but exponential curves are permitted for each of the tails. This model has six parameters: slope l, center c, two locations where the exponential tails begin (*xpos1*, *xpos2*) and the two slopes of the exponential curves (*tslope1* and *tslope2*); the latter two range between 0 and 1 and are the proportion by which the slope of the sigmoid curve is reduced. Note that *xpos2* and *tslope2* are the parameters for the anadromous side of the hybrid zone.

CLINE FITTING: MORPHOLOGICAL DATA

CFit 7 requires morphological traits to be at their maximum at the left hand side (i.e., anadromous) side of the hybrid zone. This is true for the residuals of dorsal fin length, left pectoral fin length, pelvic spine length, and second dorsal spine length. However, the residual of caudal peduncle depth is greatest in stream fish, and so we used the negative of the residuals in the CFit 7 analysis for this trait.

We used a simple clinal model to characterize the morphological data across the hybrid zone (Gay et al. 2008). Observed measurements were assumed to be drawn from a normal distribution $N(\mu_x, \sigma_x)$, where μ_x and σ_x are functions of location x across the hybrid zone. More specifically, μ_x was modeled as $\mu_{min}+(\mu_{max} - \mu_{min})\, p_x$, where μ_{min} and μ_{max} measure the mean in the stream and anadromous population, respectively, and p_x is the sigmoid function of location (with maximum slope l and center c). The variance σ_x was modeled as $p_x{}^2\sigma_1 + 2\,p_x(1 - p_x)\sigma_2 + (1 - p_x)^2\sigma_3$ to account for differences among individuals from either the anadromous or the stream populations (σ_1 and σ_3, respectively) and between either pure population and hybrids from the zone center (σ_2). This model assumes that in each location, the distribution of each trait is unimodal (drawn from a simple normal distribution). Since the focus of this paper is comparing cline slopes and centers, we only fitted models where σ_1, σ_2, and σ_3 were allowed to vary independently. As with the genetic data, parameters were estimated by maximum likelihood using CFit 7.

Comparing cline shapes

We tested for coincidence and concordance between clines by forcing them to have the same center and/or slope, respectively, and comparing the Akaike Information Criteria (AIC; Akaike 1973) values of these models to models where each cline was fitted independently. We considered a difference of 2 AIC units between models an indication of a "significant" difference, with the caveat that a difference of 7–10 AIC units is needed before concluding that the worse model has essentially no support (Burnham and Anderson 2002).

Estimating LD and trait covariance

We used CubeX (Gaunt et al. 2007; http://www.oege.org/software/cubex/) to estimate gametic linkage disequilibrium. We obtained an LD estimate between *Eda* and *ATP1a1* for the center of the hybrid zone by pooling the genotypes from the 105 individuals sampled at sites 2.2, 2.3, and 2.4 km.

We calculated the covariance between all pairings of our five morphological traits. We created three groups corresponding to the edges and center of the zone. For the anadromous side we pooled individuals from sites 1.65 and 1.8 km ($N = 50$), for the zone center we pooled sites 2.2, 2.3, and 2.4 km ($N = 105$), and the stream side sites 3.3 and 4.0 km ($N = 38$). We used bootstrapped confidence intervals based on 10,000 replicates to assess whether a covariance estimate was significantly greater than zero, as implemented in the R package boot (R Core Team 2014; Canty and Ripley 2014). An estimate was judged significant if the 99.8% confidence intervals did not contain zero; this interval approximates a Bonferroni correction for these tests ($0.05/30 = 0.0017$).

Cline analysis with the hybrid index

When genetic loci are more or less fixed on either side of the zone they can be treated as diagnostic, and can thus be used to

calculate a Hybrid Index (HI) for each individual. The Hybrid Index was calculated by counting the number of anadromous alleles across the four loci with large differences in allele frequency between the two populations (*Eda*, *ATP1a1*, P3D05, and P6A10, see results).

We can use the variance in HI in the center of the zone to estimate linkage disequilibria ($\bar{D}$) and hence the dispersal rate σ (p23 of Barton and Gale 1993). Since cline width is a product of the relationship between dispersal and selection, we can use this information to estimate the strength of selection acting on the genetic loci and the morphological traits.

Results

SAMPLE SIZES

We genotyped 428 adult (> 40 mm) fish from 16 sites through the hybrid zone at *Eda*, *ATP1a1*, and 13 unlinked anonymous SNP loci. These data are available on Dryad (http://dx.doi.org/10.5061/dryad.v7r0b). Locus P7A09 failed to amplify in most anadromous individuals and was excluded from the subsequent analyses. Amplification success was otherwise very good, with an average of 9.7 (2.2%) missing genotypes per locus. Two of the anonymous loci (P3A06 and P6B12) showed no allele frequency differences across the hybrid zone, and were also excluded at this stage. We were therefore left with genotype data for 10 anonymous SNPs. We also obtained morphological data for these 428 fish (Table 1 and Dryad: http://dx.doi.org/10.5061/dryad.v7r0b).

INDIVIDUAL CLINE FITS

The parameters for the individually fitted genetic and morphological clines are given in Tables 2 and 3, respectively, and the clines are shown in Figs. 4 and 5. The simple sigmoid clines generally fall into two categories. First, with the possible exception of pectoral fin length (which is steeper and further downstream), the traits or loci thought to be under strong selection have steep clines centered near the top of the estuary. Two of the anonymous SNPs (P3D05 and P6A10) also had steep clines centered near the selected traits and loci. The remaining anonymous clines had shallower slopes and a wide range of centers (Table 2).

FITTING STEPPED CLINES

Stepped clines fit the data worse than the basic sigmoid clines in the case of both *Eda* and *ATP1a1* (Table 2). Similarly, a stepped cline model was a significantly worse fit for dorsal fin, peduncle depth, left pelvic spine, and 2nd dorsal spine (Table 3). Fitting a stepped cline model to the pectoral fin data proved very difficult, as the runs did not converge on a single set of best-fit parameters or they gave biologically implausible parameter estimates. Removing three outlier fish with very small (likely damaged) pectoral fins from site 2.1 km (relative fin length = –3.76), site 2.3 km

Table 1. Sample sizes, mean allele frequencies for *Eda* and *ATP1a1*, mean, and variance for the Hybrid Index (HI) as calculated from *Eda*, *ATP1a1*, P3D05, and P6A10. Also shown are mean body length and the mean residuals from the morphological traits, by site. The limit of salt water for the highest tide during the study period was 2.35 km from the sea.

Location (distance from sea in km)	*N*	*Eda* anadromous allele freq	*ATP1a1* anadromous allele freq	Mean Hybrid Index	Var in Hybrid Index	Body length (mm)	Dorsal fin	Peduncle depth	Pectoral fin	L Pelvic spine	2nd dorsal spine
1.65	28	0.80	0.83	0.81	0.04	51.0	0.58	–0.10	0.31	0.79	0.60
1.80	22	0.81	0.76	0.81	0.04	48.3	0.26	0.07	0.79	0.54	0.27
1.95	20	0.75	0.68	0.74	0.03	47.7	0.59	–0.07	0.55	0.76	0.37
2.1	43	0.71	0.69	0.71	0.06	46.9	0.33	–0.06	0.20	0.45	0.25
2.2	41	0.56	0.48	0.60	0.07	47.1	0.18	–0.02	–0.17	0.25	0.22
2.3	35	0.57	0.45	0.53	0.11	45.4	0.14	0.00	0.00	0.25	0.12
2.4	24	0.37	0.48	0.50	0.06	46.4	0.11	0.00	–0.36	0.11	0.03
2.5	23	0.42	0.37	0.40	0.07	44.1	–0.12	–0.01	0.08	–0.14	0.02
2.6	31	0.23	0.28	0.28	0.06	46.9	–0.20	0.05	–0.02	–0.12	–0.07
2.7	35	0.24	0.26	0.24	0.06	47.5	–0.17	0.04	0.00	–0.39	–0.27
2.8	11	0.38	0.15	0.23	0.04	43.1	–0.29	0.04	0.11	–0.40	–0.34
2.9	27	0.19	0.15	0.14	0.03	45.1	–0.34	0.06	–0.01	–0.36	–0.26
3.0	25	0.06	0.08	0.08	0.01	49.1	–0.45	0.10	–0.10	–0.81	–0.46
3.2	9	0.00	0.17	0.04	0.00	47.3	–0.60	0.02	–0.25	–0.98	–0.53
3.3	17	0.00	0.09	0.08	0.02	46.0	–0.42	0.02	–0.19	–0.71	–0.29
4.0	21	0.00	0.00	0.02	0.00	46.6	–0.49	0.00	–1.10	–0.62	–0.42

Table 2. Parameters for the individually fitted two parameter and stepped clines for all genetic loci. The variables xpos1 and xpos2 give the number of meters upstream and downstream (respectively) from the center where the exponential tails begin. The better fit cline for each locus is given in bold typeface.

	Sigmoid cline parameters		Stepped cline parameters				Model fit output variables		
Locus	Centre (km)	Slope (*s*)	*xpos1* (m)	*xpos2* (m)	*tslope1*	*tslope2*	d.f.	LogL	AIC
Eda	**2.308**	**–0.75**					**2**	**–338.8**	**681.7**
Eda (stepped, *tslope1* ≠ *tslope2*)	2.334	–0.84	1.66	0.335	0.999	0.267	6	–336.8	685.6
ATP1a1	**2.270**	**–0.66**					**2**	**–354.7**	**713.4**
ATP1a1 (stepped, *tslope1* ≠ *tslope2*)	2.270	–0.69	2.34	1.514	0.811	0.855	6	–354.7	721.4
P3D05	2.330	–0.65					2	–354.2	712.4
P3D05 (stepped, *tslope1* = *tslope2*)	**2.465**	**–2.25**	**0.104**	**0.0001**	**0.246**	**0.246**	**5**	**–350.6**	**711.3**
P6A10	2.402	–0.83					2	–339.6	683.2
P6A10 (stepped, *tslope1* = *tslope2*)	**2.479**	**–1.35**	**0.418**	**0.102**	**0.348**	**0.348**	**5**	**–332.2**	**674.3**
P4G01	**3.320**	**–0.19**					**2**	**–409.2**	**822.5**
P4G01 (stepped, *tslope1* ≠ *tslope2*)	2.374	–2.18	0.0001	0.015	0.012	0.142	6	–408.6	829.3
P6D05	**2.574**	**–0.10**					**2**	**–438.5**	**881.0**
P6D05 (stepped, *tslope1* ≠ *tslope2*)	2.920	–2.48	0.03	0.009	0.056	0.0077	6	–436.6	885.2
P7A07	**3.216**	**–0.16**					**2**	**–421.6**	**847.3**
P7A07 (stepped, *tslope1* ≠ *tslope2*)	2.909	–2.01	0.0001	0.029	0.042	0.092	6	–421.2	854.5
P7C08	3.220	–0.31					2	–379.0	762.0
P7C08 (stepped, *tslope1* ≠ *tslope2*)	**2.910**	**–0.51**	**0.0001**	**0.96**	**0.14**	**0**	**6**	**–374.8**	**761.6**
P7E08	**2.730**	**–0.24**					**2**	**–413.4**	**830.8**
P7E08 (stepped, *tslope1* ≠ *tslope2*)	2.570	–0.48	0.05	0.47	0.31	0	6	–409.6	831.1
P7G05	2.376	–0.33					2	–405.7	815.4
P7G05 (stepped, *tslope1* ≠ *tslope2*)	**2.352**	**–2.35**	**0.032**	**0.066**	**0.11**	**0**	**6**	**–398.7**	**809.5**
P7H05	**0.980**	**–0.37**					**2**	**–229.6**	**463.3**
P7H05 (stepped, *tslope1* ≠ *tslope2*)	1.570	–2.22	0.12	1.3	0.17	0.44	6	–229.2	470.4
P9D09	2.483	–0.31					2	–414.7	833.3
P9D09 (stepped, *tslope1* ≠ *tslope2*)	**2.357**	**–0.60**	**0.266**	**0.71**	**0.045**	**0.888**	**6**	**–402.3**	**816.6**

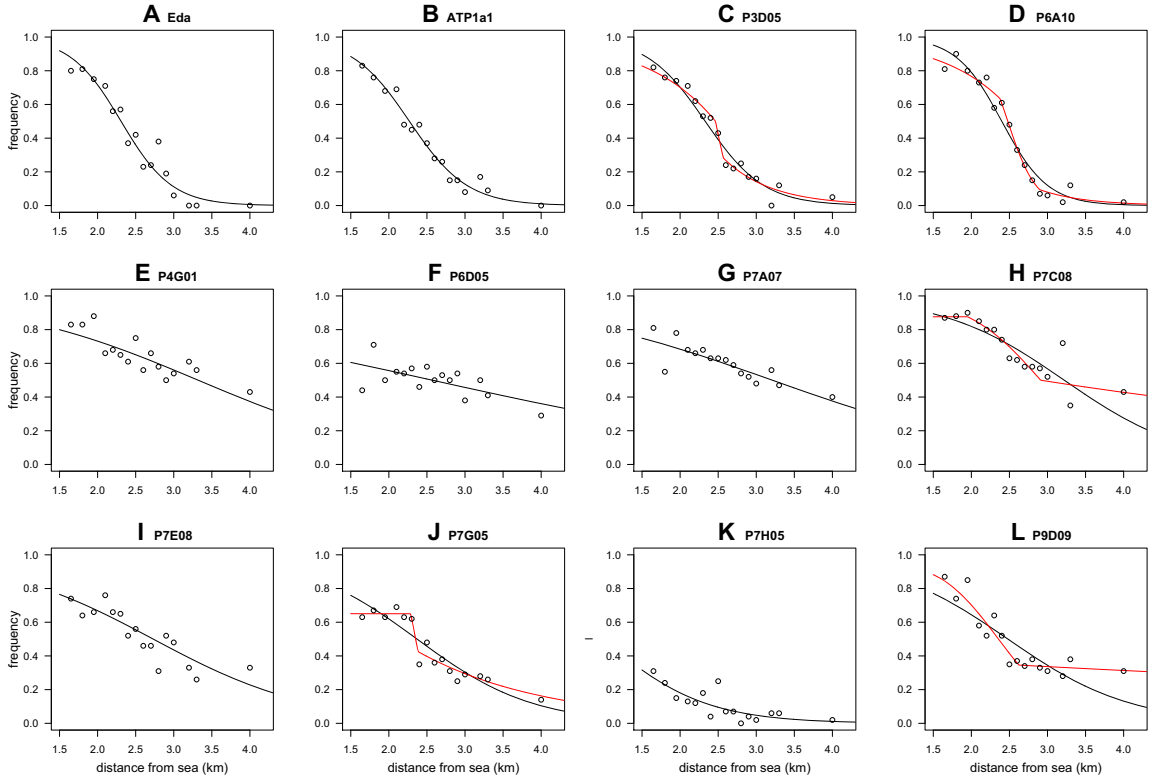

Figure 4. Individually fitted clines for the 12 genetic loci. The sigmoid cline is shown in black; the stepped cline is shown in red for the loci where this was a better fit.

(–2.7), and site 2.4 km (–3.2) and all of the fish from the 4.0 km site allowed us to find a meaningful stepped cline fit; this model fit was not an improvement over the simpler sigmoid cline model (Table 3).

Fitting a stepped cline led to an improvement in model fit for five anonymous SNP loci (Table 2). A stepped cline with *tslope1* = *tslope2* was the best fit for P6A10 (5 parameters, LogL = –332.2, AIC = 674.3) compared to the simpler sigmoid fit (2 parameters, LogL = –339.6, AIC = 683.2). The improvement in fit of a stepped cline over a simple sigmoid cline is slight for P3D05 (stepped: 5 parameters, LogL = –350.6, AIC = 711.3; sigmoid: 2 parameters, LogL = –354.2, AIC = 712.4). Of the SNPs with shallower clines, a stepped cline with *tslope1* ≠ *tslope2* was supported for P7G05 (stepped: 6 parameters, LogL = –398.7, AIC = 809.4; sigmoid: 2 parameters, LogL = –405.7, AIC = 815.4) and P9D09 (stepped: 6 parameters, LogL = –402.2, AIC = 816.5; sigmoid: 2 parameters, LogL = –414.6, AIC = 833.3), but only weakly supported for P7C08 (stepped: 6 parameters, LogL = –374.8, AIC = 761.6; sigmoid: 2 parameters, LogL = –379.01, AIC = 762.02).

COMPARING CLINE CENTERS AND SLOPES

Our central question is whether the clines for loci and traits predicted to experience strong direct selection, that is *Eda*, *ATP1a1*, and the five morphological traits, have the same center and slope. We therefore fit a model in which these seven clines were constrained to share the same slope and center, and compared it to (a) a model in which only center was constrained, and (b) a model in which both slope and center were allowed to vary independently.

Constraining all seven clines to have the same center and slope gave the best fit (27 parameters, LogL = –2078.7, AIC = 4211.5). These clines were jointly centered at 2.298 km (about 60 m downstream from the upstream limit of saltwater), and the jointly fitted slope was s = –0.7. The fit of other two models was about 6 AIC units worse: when centers were constrained but slopes were allowed to vary, we found LogL = –2075.9, AIC = 4217.8 (33 parameters); and where both slope and center varied freely, we obtained LogL = –2069.7, AIC = 4217.4 (39 parameters). In a likelihood ratio test, a model in which cline centers were the same but slopes were free to vary was not a significantly better fit to the data than the simplest model in which centers and slopes were constrained to be equal (LRT = 5.6, df = 6, P = 0.45). Similarly, the model in which slopes and centers were free to vary was not a better fit than the simplest model (LRT = 18.1, df = 12, P = 0.11). Since most of these seven clines had very similar slopes and centers (Table 2 and 3), we preferred the simplest model in which all seven share a common slope and center.

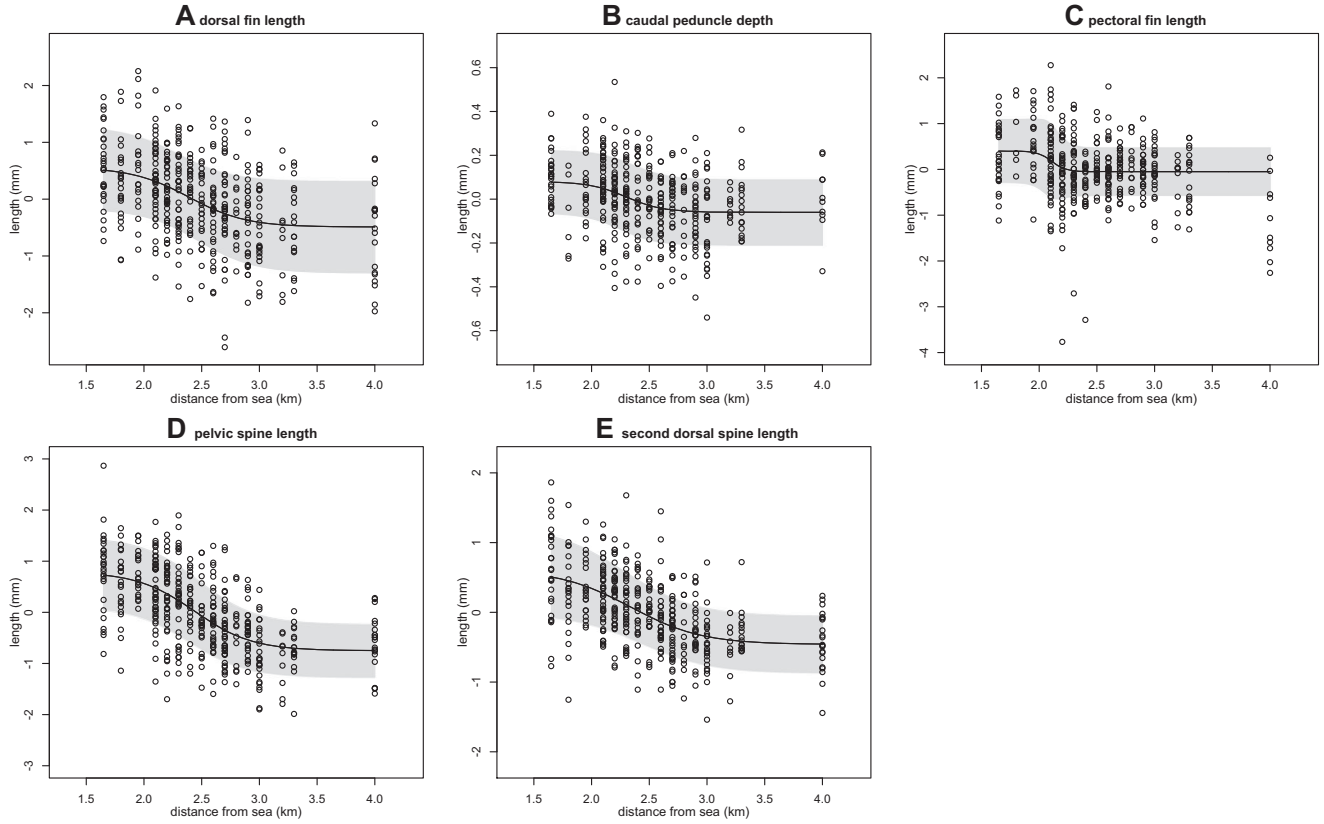

Figure 5. **(A–E) Clines in morphological traits through the hybrid zone. The width of the shaded band shows the trait variance through the zone, as calculated from σ_1, σ_2, and σ_3. The cline for pectoral fin length (5c) was fitted with the data from site 4.0 km and the three outliers omitted.**

The same pattern was found when the pectoral fin data are excluded. A model in which the remaining six clines have the same center and slope (LogL = –1669.8, AIC = 3383.7, 22 parameters) fit better than a model in which the slopes were free to vary (LogL = –1667.9, AIC = 3389.4, 27 parameters) and a model in which all slopes and centers were free to vary (LogL = –1665.6, AIC = 3395.2, 32 parameters).

We also fit the same three models for the 10 anonymous SNPs. In this case, the model in which each cline had its own slope and center was very strongly supported (LogL = –3805.5, AIC = 7651.1, 20 parameters), compared to a model in which both slopes and centers were constrained to be equal (LogL = –4369.6, AIC = 8743.2, 2 parameters) or only slope was allowed to vary independently (LogL = –4235.6, AIC = 8493.2, 11 parameters). The independent slopes and centers model is very strongly supported by a likelihood ratio test when compared to either the same slope and center model (LRT = 564.0, df = 18, $P < 0.0001$) or the model with a common slope and different centers (LRT = 430.0, df = 9, $P < 0.0001$).

ESTIMATING LD AND TRAIT COVARIANCE

CubeX estimated linkage disequilibrium between *Eda* and *ATP1a1* in the zone center (sites 2.2, 2.3, and 2.4 km) as D' = 0.31, indicating that LD in these sites is about 30% of its maximum value. A similar estimate of $D' = 0.38$ was obtained using only the fish from site 2.3 km.

The covariance between all pairs of morphological traits is shown in Table 4. The covariances between pelvic spine and second dorsal spine are significantly greater than zero in the anadromous, zone center, and stream groups, perhaps indicating a shared genetic basis for these two traits. The only other significant covariances were in the zone center, where five of the remaining nine trait combinations were significantly greater than zero.

ESTIMATING DISPERSAL AND SELECTION

The four genetic loci with steep clines (*Eda*, *ATP1a1*, P3D05, and P6A10) are more or less fixed on either side of the zone and can be treated as diagnostic. We used these four loci to calculate a Hybrid Index (HI), as described in Barton and Gale (1993). The distribution of the HI through the hybrid zone is shown in Fig. 6, while the mean and variance for each sampling location are given in Table 1.

The variance in HI in the center of the zone was used to estimate mean linkage disequilibrium ($\bar{D}$); this parameter can then be used to estimate the dispersal rate. A portion of the variance in HI in the zone center arises from the variance in allele frequency

Table 3. Parameters of the individually fitted clines for all morphological traits. The variables xpos1 and xpos2 give the number of meters upstream and downstream (respectively) from the center where the exponential tails begin. The better fit cline for each locus is given in bold typeface. For pectoral fin, the fish sampled at 4.0 km and three outliers were removed from the dataset to facilitate model fitting (see Results for details).

	Sigmoid cline parameters		Trait means		Trait variances			Stepped cline parameters				Model fit parameters		
	Centre (km)	Slope (s)	μ_{min}	μ_{max}	σ_1	σ_2	σ_3	*xpos1* (m)	*xpos2* (m)	*tslope1*	*tslope2*	d.f.	LogL	AIC
Dorsal fin	**2.39**	**–1.01**	**–0.49**	**1.05**	**0.73**	**0.57**	**0.81**					**7**	**–466.8**	**947.6**
Dorsal fin (stepped, *tslope1* ≠ *tslope2*)	2.45	–1.01	–0.61	0.93	0.71	0.62	0.78	0.83	0.42	0.70	0.36	11	–466.7	955.4
Peduncle depth	**2.30**	**–1.56**	**–0.06**	**0.14**	**0.14**	**0.17**	**0.15**					**7**	**176.1**	**–338.3**
Peduncle depth (stepped, *tslope1* ≠ *tslope2*)	2.32	–2.34	–0.19	0.38	0.02	0.32	0.00	0.06	0.00	0.00	0.16	11	177.8	–333.6
Pectoral fin	**2.11**	**–4.09**	**–0.05**	**0.49**	**0.69**	**0.99**	**0.52**					**7**	**–337.9**	**689.9**
Pectoral fin (stepped, *tslope1* ≠ *tslope2*)	2.10	–395.2	–0.05	0.51	0.70	1.02	0.52	0.00	1.46	0.01	0.93	11	–337.5	697.1
Pelvic spine	**2.44**	**–1.00**	**–0.75**	**1.54**	**0.68**	**0.72**	**0.52**					**7**	**–412.9**	**839.8**
Pelvic spine (stepped, *tslope1* ≠ *tslope2*)	2.45	–0.47	–1.41	2.83	0.58	0.88	0.29	0.62	0.50	0.00	0.30	11	–411.4	844.8
2nd dorsal spine	**2.34**	**–0.81**	**–0.46**	**1.07**	**0.63**	**0.39**	**0.41**					**7**	**–268.5**	**551.1**
2nd dorsal spine (stepped, *tslope1* ≠ *tslope2*)	2.70	–0.40	–2.03	3.56	0.94	0.09	0.53	0.52	0.10	0.08	0.59	11	–267.3	556.6

Table 4. Estimates of variance and covariance among the five morphological traits. For each trait combination, the estimates are given in the order: anadromous (top, sites 1.65 and 1.8 km), hybrid (middle, sites 2.2, 2.3, and 2.4 km), and stream (bottom, sites 3.3 and 4.0 km). Variances are listed on the diagonal (italics), covariances are listed below the diagonal. Covariances for which the 99.8% bootstrap confidence intervals do not contain zero are in bold.

	Dorsal fin	Peduncle depth	Pectoral fin	Pelvic spine	2nd dorsal spine
Dorsal fin	*0.51* *0.43* *0.65*				
Peduncle depth	0.02 0.02 –0.01	*0.02* *0.03* *0.02*			
Pectoral fin	0.24 0.09 0.19	–0.02 –0.01 –0.06	*0.53* *0.66* *0.70*		
Pelvic spine	0.20 **0.22** 0.08	0.02 **0.03** –0.01	0.21 **0.23** 0.15	*0.55* *0.51* *0.26*	
2nd dorsal spine	0.13 **0.10** 0.08	0.04 0.01 0.00	0.08 **0.13** 0.09	**0.42** **0.26** **0.14**	*0.45* *0.23* *0.15*

at the individual loci (var(p) in Barton and Gale 1993), and the remainder is due to migration bringing in purer genotypes. Taking the 64 individuals sampled at 2.3 km and 2.4 km (between which the clines at these four loci are roughly centered), we estimate the variance in allele frequency to be var(p) = 0.0025. From the first term of equation 2b of Barton and Gale (1993), the proportion of $\bar{D}$ due to var(p) is 0.0309. The mean HI in this sample is $\bar{z} = 0.515$, and the variance in HI is var(z) = 0.087. The remaining variance (0.0875–0.0309 = 0.0566) leads to the calculation $\bar{D} = 0.151$, with 95% confidence intervals (using $F_{67,Infinity}$) of 0.112–0.228.

Next, we convert the jointly fitted slope s of these four clines (–0.71) to cline width w via $1/|-0.71| = 1.39$ km (see Gay et al. 2008). We can use the estimates of $\bar{D}$ and cline width to estimate effective dispersal within the hybrid zone, as measured by the standard deviation σ of the distance between where a fish is hatched and where it reproduces, using the equation $\sigma = \sqrt{r\bar{D}w^2/(1+r)}$ (page 24 of Barton and Gale 1993). Our samples were restricted to fish over 40 mm in length, which were very likely sampled after they had dispersed. None of these four loci are physically linked, so $r = 0.5$, and thus $\sigma = 0.25$ km generation$^{-1/2}$.

Finally, we can use dispersal and cline width to estimate s^*, the difference in mean fitness at that locus or trait between

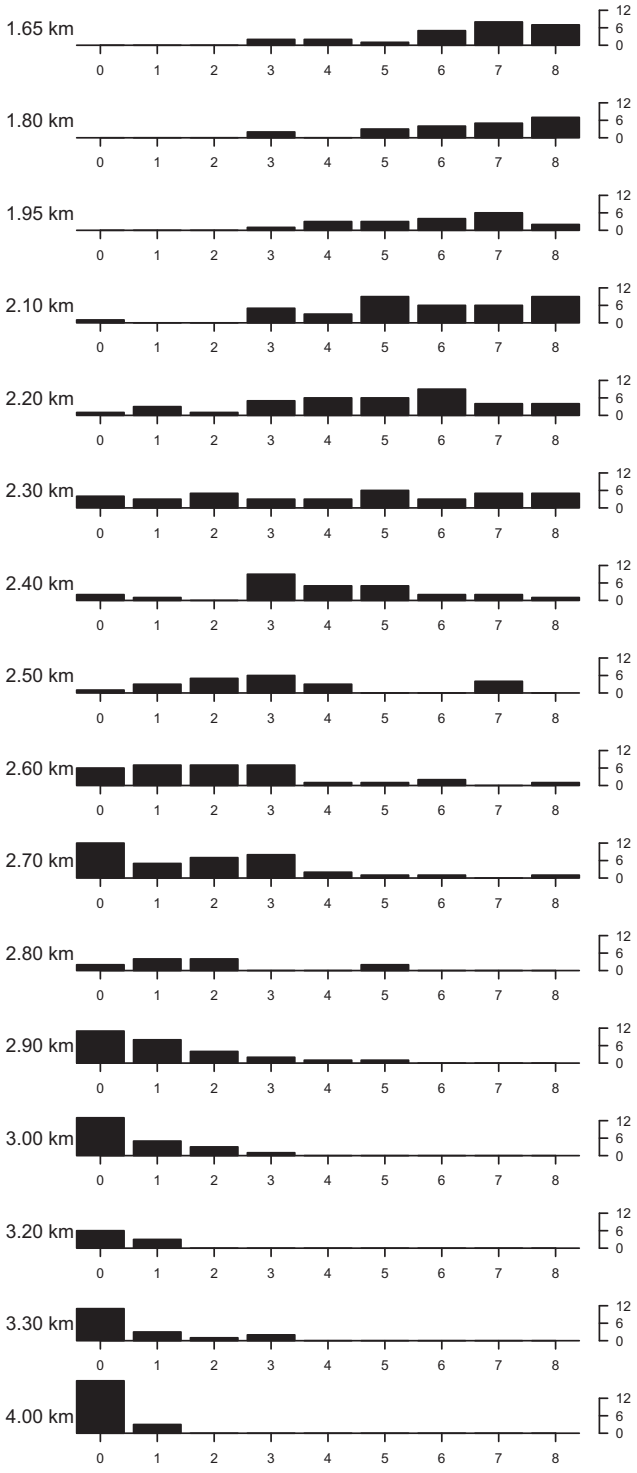

Figure 6. Bar plots of Hybrid Index through the hybrid zone. The Hybrid Index is given by the total number of anadromous alleles an individual carries at *Eda*, *ATP1a1*, P3D05, and P6A10.

populations at the center of the zone and those at the edge, using $s^* = (1.732\sigma/w)^2$ (page 16 of Barton and Gale 1993). We estimated $s^* = 0.097$ for the cline where *Eda*, *ATP1a1*, dorsal fin, 2nd dorsal spine, pectoral fin length, left pelvic spine, and peduncle depth were constrained to share the same slope and center.

Discussion

In this article, we looked for evidence that strong direct selection could uncouple clines for selected loci and traits, thereby reducing the linkage disequilibrium between them and impeding progress toward speciation. In a hybrid zone between anadromous and stream sticklebacks, we found that the best-fit model for our seven selected traits and loci constrained their clines to share the same center and slope, which suggests that they all experience the same amount of total (i.e., direct + indirect) selection. This best-fit model makes sense because five of the seven clines have very similar slopes and centers when each cline is fitted independently (*Eda*, *ATP1a1*, dorsal spine, pelvic spine, and dorsal fin; Table 2 and 3).

Close concordance in slope would be expected for colocated clines if linkage disequilibria and hence indirect selection are strong, but these conditions should also generate stepped clines (Barton and Szymura 1991; Nürnberger et al. 1995; Gay et al. 2008). Surprisingly, there was no support for a step at any of these seven clines: excluding pectoral fin, the slope of the individually fitted stepped cline was always very similar to the slope of the simpler sigmoid cline (Table 2 and 3).

Our data are thus compatible with two unexpected scenarios. First, direct selection may be dominant, implying that the selection regimes for these traits and loci all transition in the same part of the river and all experience direct selection of $s^* = 0.1$. Given the diverse functions of our selected traits and loci (predator defense, osmoregulation, swimming performance) it seems implausible that they all experience the same selection regime. Alternatively, indirect selection may be strong enough to force these clines into sharing the same slopes and centers, irrespective of their individual selection regimes. However, this indirect selection has somehow not generated stepped clines. We do find substantial LD between *Eda* and *ATP1a1* and significant covariance between the morphological traits in the center of the zone (Table 4), which means they all ought to be experiencing at least some mutual indirect selection. So, although Bonsall Creek exhibits the conditions under which we expect to find stepped clines (steep, concordant clines, and substantial LD), there is no evidence for a step for any of our seven selected traits and loci. This phenomenon warrants further investigation, but this is unfortunately beyond the scope of this paper.

Two anonymous SNPs (P3D05 and P6A10) also had steep clines located close to those for the selected traits and loci (Fig. 4C and D, Table 2). P3D05 is located at 10,349 kb along LGIII and is embedded within the "amyloid beta (A4) precursor protein-binding, family B, member 1 interacting protein" gene. P6A10 is

at 12,601 kb on LGVII in the "mitochondrial calcium uptake 2" gene (stickleback genome assembly BROAD S1, version 86.1 on Ensembl). Neither gene has been explicitly connected to freshwater adaptation in sticklebacks, although Hohenlohe et al. (2010) did highlight five divergent SNPs (at 12,212,041; 12,505,622; 12,525,947; 12,801,733; and 12,987,907) in the vicinity of P6A10 on LGVII (Fig. 8A in Hohenlohe et al. 2010, and P. Hohenlohe, pers. comm.). Terekhanova et al. (2014) also highlight an area of LGVII (between 17,982 kb and 18,002 kb) that is divergent between Russian marine and freshwater stickleback.

Interestingly, there is support for a stepped cline at P6A10 (Table 2 and Fig. 3D). The most plausible explanation for this step is that P6A10 alleles are experiencing both weak direct selection on either side of the hybrid zone ($tslope \times s* = 0.11$, see Szymura and Barton 1991), and a strong combination of direct and indirect selection in the center ($s* = 0.34$). The source of this indirect selection must be linkage disequilibria with other selected loci, either on the same chromosome or elsewhere in the genome.

THE STRUCTURE OF THE BONSALL CREEK HYBRID ZONE

Another puzzling feature of this hybrid zone is that the observed clines are generally very steep compared to the movement capability of the fish. A mark-recapture study conducted on 23rd and 24th May 2006 (T.H. Vines and A.C. Dalziel, unpub. data) found several fish that had moved several hundred meters overnight. If fish do move tens of meters per day in a random direction, selection would need to be extremely strong to maintain these steep clines: almost all fish crossing the hybrid zone would need to be eliminated before they could be sampled. It is difficult to imagine any selective force that could accomplish this. Adult sticklebacks are able to cope with a wide range of salinities without suffering large effects (e.g., Schaarschmidt et al. 1999; Barrett et al. 2009; Gibbons et al. 2016), so the salinity differences between the estuary and freshwater cannot be a sufficient source of mortality. While piscivorous birds and fish are present throughout the hybrid zone, their predation rate for migrant fish would need to be very high to counter the constant flux of individuals to the "wrong" side of the zone.

A more plausible explanation is that while fish could easily traverse the zone, they instead choose a point along the river according to some function of their phenotype and stay there over the breeding season. For males this would manifest as the location where they choose to build a nest; for females the location where they look for males, or where they rest and forage while developing eggs. This must be the case for the anadromous fish, which spend the winter in the sea and migrate into the estuary and the lower parts of the river to breed (Hagen 1967). The Bonsall Creek hybrid zone is thus similar to hybrid zones between migratory and resident bird species (e.g., Rohwer et al. 2001; Brelsford and Irwin 2009) where one species migrates long distances over the winter and the zone is reformed each breeding season. The key parameter in these zones is not the total distance traveled (which is orders of magnitude greater than the zone width), and instead is the distance between where the organism is hatched and where it reproduces. For Bonsall Creek, the standard deviation of the latter distance is estimated as $\sigma = 0.25$ km gen$^{1/2}$; assuming a normal distribution of parent-offspring dispersal ~68% of fish would breed between 125 m upstream and 125 m downstream of where they hatched.

We hypothesize that the steep concordant clines in Bonsall Creek arise because breeding location along the salinity gradient is determined by a multilocus "breeding salinity preference," and the loci underlying this trait are in tight LD with the traits and loci we examined here. In this scenario, pure anadromous fish prefer to nest in the saltiest part of the river near the sea, while individuals with an increasing proportion of stream alleles at the breeding salinity preference loci are found closer and closer to freshwater. The strong LD between the alleles underlying breeding salinity preference and the alleles for other morphological or genetic clines then must be maintained by strong total selection. Alternatively, there may be no segregating variance for breeding salinity preference, and the fish follow a simple rule such as "breed in the part of the river where the effort to maintain osmotic balance and swimming effort is minimized." A preference of this sort would be analogous to a "one allele" model of speciation (Felsenstein 1981), as only performance differences between anadromous and stream fish are required to separate them along the salinity gradient.

Despite the structuring of the hybrid zone along the salinity gradient, morphologically intermediate hybrids are common at the center (Table 1 and Figs. 5 and 6), such that reproductive isolation is far from complete. Given these ample opportunities for recombination, some form of selection must be responsible for maintaining LD and concordant clines between the various selected traits and loci we study here. What is the source of this selection? Unlike adults, eggs and newly hatched offspring cannot change their location within the river, and must therefore experience the changing salinities associated with the tidal cycle. The effects of salinity on egg hatching success and growth in stickleback appear to be significant. For example, the hatching success of eggs from Belgian freshwater populations is better in low than high salinity, while the converse is true for marine populations (Heuts 1947), and similar patterns were observed by Kassen et al. (1995) for fish from the Georgia Strait (where Bonsall Creek is located). More recent studies have found low hatching success and low growth rate of freshwater fish in higher salinities (Marchinko and Schluter 2007; DeFaveri and Merila 2014).

As hypothesized by Hagen (1967), selection may also occur over the winter. For example, any fish that remains in the stream over the winter will need to cope with the combination of low temperatures and fresh water (which is challenging for anadromous fish; Heuts 1947; Schaarschmidt et al. 1999; Gibbons et al. 2016), and the need to efficiently maintain position in the stream. It will also need to compete effectively with pure stream sticklebacks for (presumably) limited food resources. A hybrid carrying anadromous alleles for loci determining temperature or salinity tolerance, burst swimming ability, and invertebrate predator avoidance will presumably have a lower chance of survival. By contrast, a hybrid or stream fish in the marine environment will need to cope with the higher density of vertebrate predators, long distance migration, and competition for food with the pure anadromous population. Studies with sympatric stickleback species pairs have shown that selection against hybrids via competition for food can be strong (Rundle 2002), particularly when coupled with predation (Vamosi and Schluter 2002; Rundle et al. 2003). It is also possible that some hybrids overwinter in the estuary, which alternates between each environment through the tidal cycle.

Conclusions

This study tested whether direct selection could uncouple clines at loci and traits known to be under ecological selection. We examined two well-studied genes in a hybrid zone between stream and anadromous sticklebacks, and found that the clines at *Eda* and *ATP1a1* had very similar same slopes and centers, as did the five morphological traits we examined. Two anonymous SNPs also had steep clines. There was no evidence for stepped clines at any of the selected traits or loci. The concordance and coincidence of the seven selected clines suggests either that (a) these diverse traits and loci all experience the same selection regime, or (b) that indirect selection has somehow brought these clines in concordance without inducing a step in the center. Even given our inability to distinguish these two scenarios with the current data, our results shed light on the relative roles of direct and indirect selection in promoting (or inhibiting) speciation in these species.

ACKNOWLEDGMENTS

The authors are grateful to Taylor Gibbons for help in the lab, and Jeremy Schmutz for information on the SNP loci. Amelia Mahony, Aleeza Gerstein, Simone Des Roches, Joey Courchesne, Jessica Hill, and Chrissy Spencer for help with sample collection, and to James Thomas and other members of the Halalt first nation for access to their land and information on the history of the creek. We thank Nick Barton, Sean Rogers, and Thomas Lenormand for helpful discussions, while Rowan Barrett and Dave Toews provided comments on an earlier version. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through Discovery and Discovery Accelerator grants to P.M.S., Discovery and Special Research Opportunity grants to D.S., and a Canada Graduate Scholarship to A.C.D. and A.A. T.H.V. was supported by a European Union FP6 Outgoing International Fellowship and a real job. T.V. was supported by the NSERC-CREATE Training Program in Biodiversity Research and NSF grant IOS-1145468.

DATA ARCHIVING

The doi for our data is 10.5061/dryad.v7r0b.

LITERATURE CITED

Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle, and R. Buggs. 2013. Hybridization and speciation. J. Evol. Biol. 26:229–246.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pp. 267–281 *in* B.N. Petrov and F. Csáki, eds. International symposium on information theory. Akad´emiai Kiadi, Budapest, Hungary.

Barrett, R. D. H., S. M. Rogers, and D. Schluter. 2008. Natural selection on a major armor gene in threespine stickleback. Science 322:255–257.

Barrett, R. D. H., T. H. Vines, J. S. Bystriansky, and P. M. Schulte. 2009. Should I stay or should I go? The *Ectodysplasin locus* is associated with behavioural differences in threespine stickleback. Biol. Lett. 5:788–791.

Barton, N. H. 1983. Multilocus clines. Evolution 37:454–471.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Ann. Rev. Ecol. Syst. 16:113–148.

Barton, N. H., and K. S. Gale. 1993. Genetic analysis of hybrid zones. Pp. 13–45 *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford Univ. Press, Oxford.

Barton, N. H., and M. A. R. de Cara. 2009. The evolution of strong reproductive isolation. Evolution 63:1171–1190.

Bazykin, A. D. 1969. Hypothetical mechanism of speciation. Evolution 23:685–687.

Bell, M. A., G. Ortí, J. A. Walker, and J. P. Koenings. 1993. Evolution of pelvic reduction in threespine stickleback fish: a test of competing hypotheses. Evolution 47:906–914.

Bierne, N. J. Welch, E. Loire, F. Bonhomme, and P. David. 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. Mol. Ecol. 20:2044–2072.

Brelsford, A., and D. E. Irwin. 2009. Incipient speciation despite little assortative mating: the yellow-rumped warbler hybrid zone. Evolution 63:3050–3060.

Bulmer, M. G. 1972. Multiple niche polymorphism. Am. Nat. 106:254–257.

Burnham, K. P. and D. L. Anderson. 2002. Model selection and multimodel inference: a practical-theoretic approach. Springer-Verlag, New York.

Canty, A., and B. Ripley. 2014. boot: bootstrap R (S-plus) functions. R Package Version 1.3–13.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, V. J. Guadalupe, M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, and D. M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307:1928–1933.

Dalziel, A. C. and P. M. Schulte. 2012. Correlates of prolonged swimming performance in F2 hybrids of migratory and non-migratory threespine stickleback. J. Exp. Biol. 215:3587–3596.

Dalziel, A. C., T. H. Vines, and P. M. Schulte. 2012. Reductions in prolonged swimming capacity following freshwater colonization in multiple threespine stickleback populations. Evolution 66:1226–1239.

DeFaveri, J., T. Shikano, Y. Shimada, A. Goto, and J. Merilä. 2011. Global analysis of genes involved in freshwater adaptation in threespine sticklebacks (*Gasterosteus aculeatus*). Evolution 65:1800–1807.

DeFaveri, J., T. Shikano, Y. Shimada, and J. Merila. 2013. High degree of genetic differentiation in marine three-spined sticklebacks (*Gasterosteus aculeatus*). Mol. Ecol. 22:4811–4828.

DeFaveri J. and J. Merila. 2014. Local adaptation to salinity in the three - spined stickleback? J. Evol. Biol. 27:290–302.

Evans, D. H., P. M. Piermarini, and K. P. Choe. 2005. The multifunctional fish gill: dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. Physiol. Rev. 85:97–177.

Feder, J. L., S. M. Flaxman, S. P. Egan, A. A. Comeault, and P. Nosil. 2013. Geographic mode of speciation and genomic divergence. Ann. Rev. Ecol. Evol. Syst. 44:73–97.

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? Evolution 35:124–138.

Flaxman, S. M., J. L. Feder, and P. Nosil. 2013. Genetic hitchhiking and the dynamic buildup of genomic divergence during speciation with gene flow. Evolution 67:2577–2591.

Flaxman, S. M., A. C. Wacholder, J. L. Feder, and P Nosil. 2014. Theoretical models of the influence of genomic architecture on the dynamics of speciation. Mol. Ecol. 23:4074–4088.

Gay, L., P.-A. Crochet, D. A. Bell, and T. Lenormand. 2008. Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. Evolution 62:2789–2806.

Gaunt, T. R., S. Rodriguez, and I. N. M. Day. 2007. Cubic exact solutions for the estimation of pairwise haplotype frequencies: implications for linkage disequilibrium analyses and a web tool 'CubeX'. BMC Bioinformatics 8:428.

Gibbons, T. C., S. M. Rudman, and P. M. Schulte. 2016. Responses to simulated winter conditions differ between threespine stickleback ecotypes. Mol. Ecol. 25:764–775

Glazer A. M., E. E. Killingbeck, T. Mitros, D. S. Rokhsar, and C. T. Miller. 2015. Genome assembly improvement and mapping convergently evolved skeletal traits in sticklebacks with genotyping-by-sequencing. G3 5:1463–1472.

Hagen, D. W. 1967. Isolating mechanisms in threespine sticklebacks (Gasterosteus). J. Fish. Res. Board Can. 24:1637–1692.

Harrison, R. G. 1990. Hybrid zones: windows on evolutionary process. Pp. 69–128 *in* D. Futuyma and J. Antonovics, eds. Surveys in evolutionary biology. Oxford Univ. Press, Oxford, New York, NY.

Heuts, M. J. 1947. Experimental studies on adaptive evolution in *Gasterosteus aculeatus* L. Evolution 1:89–102.

Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet. 6:e1000862.

Jones, F. C., C. Brown, J. Pemberton, and V. Braithwaite. 2006. Reproductive isolation in a threespine stickleback hybrid zone. J. Evol. Biol. 19:1531–1544.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484: 55–61.

Kaplan, J. H. 2002. Biochemistry of Na,K-ATPase. Ann. Rev. Biochem. 2002:511–535.

Kassen, R., D. Schluter, and J. D. McPhail. 1995. Evolutionary history of threespine sticklebacks (*Gasterosteus* spp): insights from a physiological clock. Can. J. Zool. 73:2154–2158.

Kruuk, L. E., S. J. E. Baird, K. S. Gale, and N. H. Barton. 1999. A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. Genetics 153:1959–1971.

Lauder, G. V., J. C. Nauen, and E. G. Drucker. 2002. Experimental hydrodynamics and evolution: function of median fins in ray-finned fishes. Integr. Comp. Biol. 42:1009–1017

Mäkinen, H. S., T. Shikano, and J. Merila. 2008. Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. Mol. Ecol. 17:3565–3582.

Marchinko, K. B. 2009. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. Evolution 63:127–138.

Marchinko, K. B., and D. Schluter. 2007. Parallel evolution by correlated response: lateral plate reduction in threespine stickleback. Evolution 61:1084–1090.

Nürnberger, B., N. H. Barton, C. J. MacCallum, J. Gilchrist, and M. Appleby. 1995. Natural selection on quantitative traits in a Bombina hybrid zone. Evolution 49:1224–1238.

Peichel, C. L., K. S. Nereng, K. A. Ohgi, B. L. E. Cole, P. F. Colosimo, C. A. Buerkle, D. Schluter, and D. M. Kingsley 2001. The genetic architecture of divergence between threespine stickleback species. Nature 414:901–905.

R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raeymaekers, J. A. M., N. Konijnendijk, M. H. D. Larmuseau, B. Hellemans, L. De Meester, and F. A. M. Volckaert. 2014. A gene with major phenotypic effects as a target for selection vs. homogenizing gene flow. Mol. Ecol. 23:162–181.

Reimchen, T. E. 1994. Predators and evolution in threespine stickleback. Pp. 240–273 *in* M. A. Bell and S. A. Foster, eds. Evolution of the threespine stickleback. Oxford Univ. Press, Oxford.

Rohwer, S., E. Bermingham, and C. Wood. 2001. Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. Evolution 55:405–422.

Rosser, N., K. K. Dasmahapatra, and J. Mallet. 2014. Stable *Heliconius* butterfly hybrid zones are correlated with a local rain-fall peak at the edge of the Amazon basin. Evolution 68:3470–3484.

Rundle, H. D. 2002. A test of ecologically dependent postmating isolation between sympatric sticklebacks. Evolution 56:322–329.

Rundle, H. D., S. M. Vamosi, and D. Schluter. 2003. Experimental test of predation's effect on divergent selection during character displacement in sticklebacks. Proc. Natl. Acad. Sci. USA 100:14943–14948.

Schaarschmidt, T., E. Meyer, and K. Jurss. 1999. A comparison of transport-related gill enzyme activities and tissue-specific free amino acid concentrations of Baltic Sea (brackish water) and freshwater threespine sticklebacks, *Gasterosteus aculeatus*, after salinity and temperature acclimation. Marine Biol. 135:689–697.

Shimada, Y., T. Shikano, and J. Merilä. 2011. A high incidence of selection on physiologically important genes in the three-spined stickleback, *Gasterosteus aculeatus*. Mol. Biol. Evol. 28:181–193.

Slatkin, M. 1975. Gene flow and selection in a two-locus system. Genetics 81:787–802.

Smadja, C. M., and R. K. Butlin. 2011. A framework for comparing processes of speciation in the presence of gene flow. Mol. Ecol. 20:5123–5140.

Szymura, J. M., and N. H. Barton. 1991. The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. Evolution 45:237–261.

Taylor, E. B., and J. D. McPhail. 1986. Prolonged and burst swimming in anadromous and freshwater threespine stickleback, *Gasterosteus aculeatus*. Can. J. Zool. 64:416–420.

Terekhanova, N. V., M. D. Logacheva, A. A. Penin, T. V. Neretina, A. E. Barmintseva, G. A. Bazykin, A. S. Kondrashov, and N. S. Mugue. 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. PLoS Genet. 10:e1004696.

Vamosi, S. M., and D. Schluter. 2002. Impacts of trout predation on fitness of sympatric sticklebacks and their hybrids. Proc. R Soc. London B 269:923–930.

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

Walker, J. A., and M. A. Bell. 2000. Net evolutionary trajectories of body shape evolution within a microgeographic radiation of threespine sticklebacks (*Gasterosteus aculeatus*). J. Zool. 252:293–302

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

Yeaman, S, and M. C. Whitlock. 2011. Genetic architecture of adaptation under migration-selection balance. Evolution 65:1897–1911.

Yeaman, S., and S. P. Otto. 2011. Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. Evolution 65:2123–2129.

Associate Editor: M. Hahn
Handling Editor: M. Servedio

Supporting Information

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

Supplemental Table S1. Primer and sequence details for *Eda, ATP1a1,* and the 13 anonymous SNPs used in this study.