JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.13264

Adaptation in temporally variable environments: stickleback armor in periodically breaching bar-built estuaries

ANTOINE PACCARD*, BEN A. WASSERMAN†, DIETA HANSON*, LOUIS ASTORG‡, DAN DURSTON§, SARA KURLAND¶, TRAVIS M. APGAR†, RANA W. EL-SABAAWI§, ERIC P. PALKOVACS†, ANDREW P. HENDRY* & ROWAN D. H. BARRETT*

*Redpath Museum and Department of Biology, McGill University, Montreal, QC, Canada
†Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, USA
‡Pavillon des Sciences Biologiques, Université du Québec à Montréal, Montréal, QC, Canada
§Department of Biology, University of Victoria, Victoria, BC, Canada
¶Zoologiska Institutionen: Populations Genetik, Stockholm University, Stockholm, Sweden

Keywords:

armor traits;
ecological stoichiometry;
Eda;
predation;
temporal variation.

Abstract

The evolutionary consequences of temporal variation in selection remain hotly debated. We explored these consequences by studying threespine stickleback in a set of bar-built estuaries along the central California coast. In most years, heavy rains induce water flow strong enough to break through isolating sand bars, connecting streams to the ocean. New sand bars typically re-form within a few weeks or months, thereby re-isolating populations within the estuaries. These breaching events cause severe and often extremely rapid changes in abiotic and biotic conditions, including shifts in predator abundance. We investigated whether this strong temporal environmental variation can maintain within-population variation while eroding adaptive divergence among populations that would be caused by spatial variation in selection. We used neutral genetic markers to explore population structure and then analysed how stickleback armor traits, the associated genes *Eda* and *Pitx1* and elemental composition (%P) varies within and among populations. Despite strong gene flow, we detected evidence for divergence in stickleback defensive traits and *Eda* genotypes associated with predation regime. However, this among-population variation was lower than that observed among other stickleback populations exposed to divergent predator regimes. In addition, within-population variation was very high as compared to populations from environmentally stable locations. Elemental composition was strongly associated with armor traits, *Eda* genotype and the presence of predators, thus suggesting that spatiotemporal variation in armor traits generates corresponding variation in elemental phenotypes. We conclude that gene flow, and especially temporal environmental variation, can maintain high levels of within-population variation while reducing, but not eliminating, among-population variation driven by spatial environmental variation.

Introduction

Spatial variation in selection is known to shape spatial variation in adaptive traits (Endler, 1986; Schluter, 2000; Hendry, 2017); less certain is the role of temporal variation. In fact, different meta-analyses of selection gradients have come to opposite conclusions about the prevalence and importance of temporal variation in selection (Siepielski *et al.*, 2009; Morrissey & Hadfield, 2012). Indeed, although the strength and direction of selection have been shown to greatly vary across time (Reimchen & Nosil, 2002; Hunt *et al.*, 2008; Siepielski

Correspondence: Antoine Paccard, Redpath Museum and Department of Biology, McGill University, Montreal, QC H3A 0C4, Canada.
Tel.: +01 514 398 4086; fax: +01 514 398 3185;
e-mail: antoine.paccard@mcgill.ca

et al., 2009), others found that it was not necessarily the case (Hoekstra *et al.*, 2001; Kingsolver *et al.*, 2001; Morrissey & Hadfield, 2012). Consequently, the effect of temporal variation in selection on phenotypic and genetic divergence remains unclear. Similarly, theoretical models evaluating the evolutionary importance of temporal environmental stochasticity come to variable conclusions that depend on the specific parameters used to calculate fitness at different time points (Coulson & Tuljapurkar, 2008; Chevin *et al.*, 2010; Chevin, 2013; Saether & Engen, 2015). Despite these variable attempts at generalization, many specific instances are known where the direction and magnitude of selection vary through time in correspondence with environmental conditions (Hairston & Dillon, 1990; Grant & Grant, 2002; Reimchen & Nosil, 2002; Mustonen & Lässig, 2007; Sletvold & Grindeland, 2007; Simons, 2009). Indeed, it has been recently argued that temporal variation in environmental conditions can explain an important amount of the temporal variation in selection coefficients analysed across studies (Siepielski *et al.*, 2017). Thus, temporal variation in selection is sometimes strong, but just how important this variation is for evolution remains much debated.

What might be the consequences of temporal variation for evolutionary processes? First, temporal environmental variation dictates that current conditions are not necessarily reflective of past selection and, hence, populations might not appear particularly well adapted to the specific conditions at any given time (Michel *et al.*, 2014). Second, and for the same reason, temporally variable environments might not allow (or favour) strong adaptive divergence across space even if spatial environmental variation is strong at any given time (Bell, 2010). Third, because the particular alleles favoured by selection vary through time, temporal environmental variation can sometimes maintain adaptive genetic variation within populations (Ellner & Hairston, 1994; Sasaki & Ellner, 1997). Fourth, because phenotypic plasticity can sometimes allow a given genotype to quickly adjust its phenotype to fluctuating conditions, it might be favoured over genetic adaptation in temporally fluctuating environments (Chevin & Lande, 2010). Finally, temporal variation can favour bet-hedging strategies, where individuals adopt strategies that reduce long-term variance in fitness even at the expense of short-term mean fitness (Childs *et al.*, 2010). In short, the potential consequences of temporal variation in environments and selection are many – highlighting the need for focused empirical studies in natural ecosystems.

Some of the above theoretical expectations have been confirmed in empirical studies. For instance, stable environments can harbour low genetic variation (Kellermann *et al.*, 2006, 2009), low phenotypic plasticity (Lind & Johansson, 2007; Lind *et al.*, 2010; Baythavong, 2011) and low bet hedging (Simons, 2009). However, the importance of temporal environmental variation in shaping genetic and phenotypic variation within and among populations that experience spatial environmental variation remains uncertain. Some studies have found that spatial differences in adaptive traits are generally maintained through time, suggesting that temporal variation does not overwhelm spatial variation (Mojica *et al.*, 2012; Morrissey & Hadfield, 2012; Gotanda & Hendry, 2014). However, these studies often examine populations known *a priori* to consistently differ in adaptive traits, so one might not expect a strong role for temporal variation (Hendry, 2017). What is needed, then, are studies examining within- and among-population trait variation in systems subject to strong spatial environmental variation but also strong temporal environmental variation.

Stickleback predator defence in bar-built estuaries

We suggest that the evolutionary consequences of temporal environmental variation might be profitably assessed using estuarine threespine stickleback (*Gasterosteus aculeatus*) known to experience extreme seasonal fluctuations. These populations inhabit 'bar-built' estuaries along the central coast of California, USA, which are characterized by fluctuations in ocean connectivity driven by seasonal rainfall patterns. Rainfall connects estuaries to the ocean in times of sufficiently high stream flow (Allen *et al.*, 2006), typically during the winter and/or spring months when heavy rains induce flows strong enough to breach the sand bar and thus connect the estuary to the ocean (Fig. 1b, Fig. S1; Behrens & Bombardelli, 2009; Behrens *et al.*, 2013; Rich & Keller, 2013). Once the high flows stop, a sand bar forms at the mouth of the estuary due to wave action and the deposition of new sand from the stream, forming a brackish-to-freshwater lagoon (Bradley & Griggs, 1976). Owing to these geophysical properties, a given bar-built estuary can greatly and rapidly vary in environmental conditions over the course of a single year, as well as across years. These properties also lead to frequent and dramatic shifts in biotic conditions, including the presence vs. absence of various stickleback fish predators (Becker & Reining, 2008; Frechette *et al.*, 2016).

To consider the evolutionary consequences of this environmental variation associated with bar-built estuaries, we focus on stickleback armor traits, including spines, body shape and lateral plates, all of which differ strongly between marine and freshwater environments, especially in relation to spatial variation in predators (Hoogland *et al.*, 1956; Reimchen, 1980, 1992, 1994, 1995; Reimchen & Nosil, 2002; Marchinko, 2009). Stickleback armor traits are also known for their strong genetic basis (Peichel *et al.*, 2001; Colosimo *et al.*, 2004; Jones *et al.*, 2012). In addition, these traits are expected to have ecological effects on their environment through

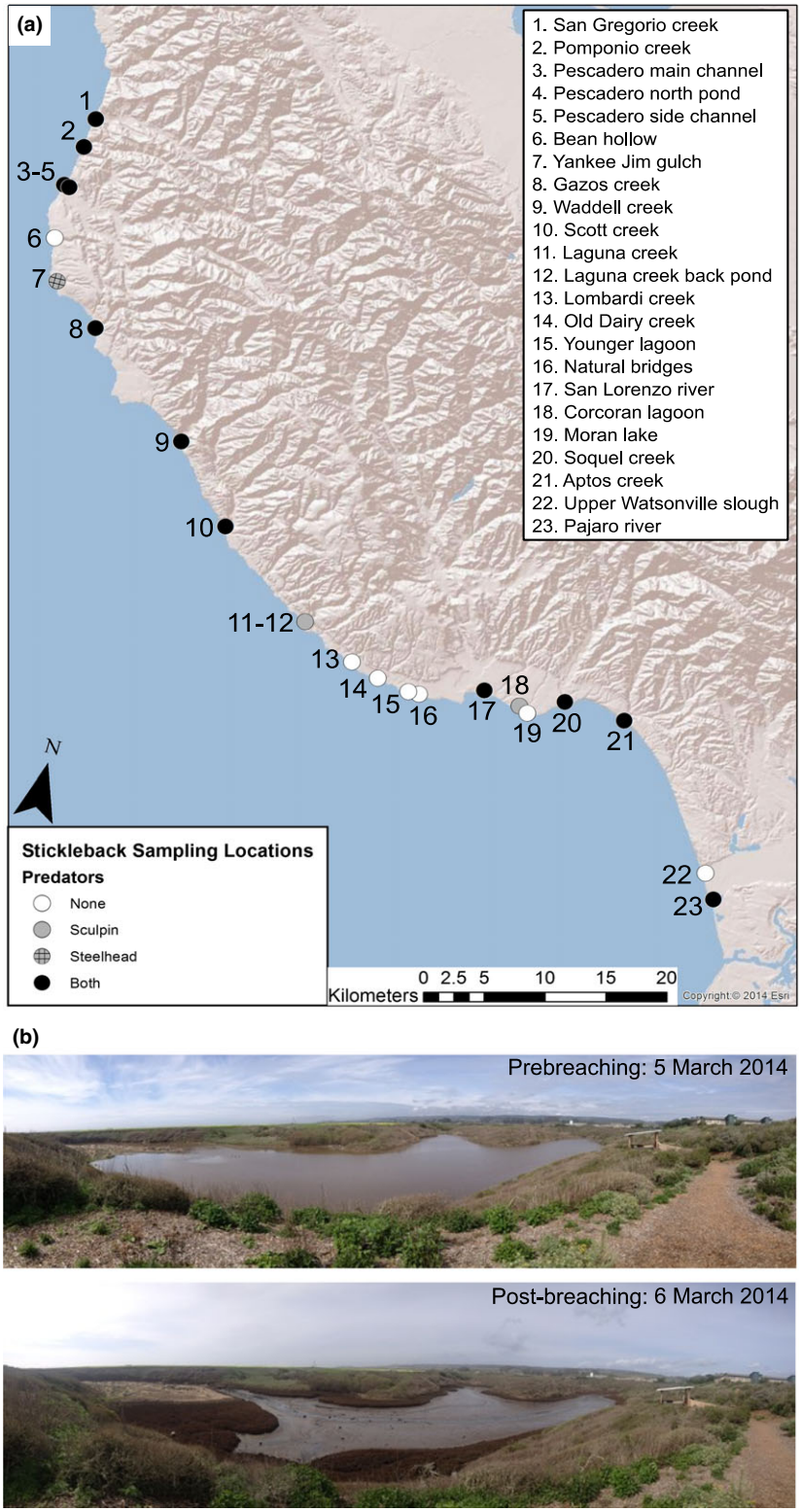

Fig. 1 Map of study sites (a) and photographs of a breaching event in Younger lagoon (b). Coloured markers indicate the presence of known stickleback predators.

their influence on nutrient dynamics (El-Sabaawi *et al.*, 2016), thus allowing us to consider the potential consequences of temporal variation not only for evolution but also eco-evolutionary dynamics (Hendry, 2017). For instance, variation in fish elemental composition can indicate specific changes in individual behaviour

(e.g. foraging) that influence zooplankton community structure (El-Sabaawi *et al.*, 2016; Durston & El-Sabaawi, 2017). We structured our analysis around four key questions:

1 *Is gene flow sufficiently restricted to enable adaptive divergence among the estuary populations?* We investigate this question by assessing variation in neutral genetic markers that can inform the extent and nature of gene flow among stickleback populations in the different estuaries.
2 *Do stickleback in the different estuaries differ in armor traits, and are these differences associated with spatial variation in predators?* Because the genetic basis of several stickleback armor traits is well known (e.g. *Eda* for lateral plates and *Pitx1* for pelvic structures), we examined variation in both the traits and marker alleles associated with *Eda* and *Pitx1*.
3 *Do estuary stickleback have particularly high levels of (presumed) adaptive variation, as would be expected in their temporally variable environments?* This within-population variation could also be maintained by high among-population gene flow, thus linking to our first question above.
4 *How does an important ecological effect trait, elemental composition* (*phosphorus content, %P*), *vary in relation to phenotypes (armor), genotypes (Eda) and predation regime?* Such variation would indicate the potential for genetically based spatiotemporal variation in traits to impact nutrient dynamics, thus generating potential eco-evolutionary links.

Materials and methods

Field collections

Between April and August 2014, after most estuaries were closed for the summer (i.e. the sand bar separating the estuary from the ocean was in place), we collected threespine stickleback from 23 coastal estuary sites along a 90 km stretch of the central cost of California from San Gregorio State Beach in San Mateo County to the Pajaro River in Santa Cruz County (Table S1, Fig. 1a). Using a combination of minnow traps and beach seines, we collected 30 stickleback of length > 30 mm per site and immediately killed them with an overdose of tricaine methanosulphonate (MS-222). The fish were then placed on ice until they could be stored in a freezer before further processing. At each site, we also visually recorded from seine net catches the presence of known stickleback predators: steelhead trout (*Oncorhynchus mykiss*) and sculpin species (*Cottus asper* and *Leptocottus armatus*). Importantly, predator abundance in bar-built estuaries fluctuates with the frequency of breaching events (Becker & Reining, 2008). We also calculated watershed area for each creek using ArcGIS v. 10.2. Watershed area is a reliable proxy for stream flow, with larger watersheds tending to sustain greater flows and therefore spending longer periods of time with the estuary mouth open (Elwany *et al.*, 1998; Mohamoud & Parmar, 2006). In the laboratory, the collected stickleback were placed in 10% formalin (VWR, Radnor, PA, USA) after the right pectoral fin was removed and stored in 95% ethanol for genetic analyses. Stickleback specimens were then stained using alizarin red dye. To do so, they were first soaked in water for 24 h, then in a solution of alizarin red and 0.5% KOH for 24 h, followed by a second soak in water for 24 h to remove excess dye. Fish were then stored in 40% isopropyl alcohol until further processing.

Population genetics

DNA was extracted from stickleback fin clip tissue using a phenol–chloroform-based protocol. Briefly, tissues were left overnight in tissue digestion buffer and proteinase K at 55 °C, followed by phenol–chloroform and ethanol washes to isolate the DNA. Nine microsatellite markers were amplified on 10–59 individuals per population (Table S1). Two of these markers, stn381 and stn82, are linked to genes *Eda* and *Pitx1*, respectively (Shapiro *et al.*, 2004; Colosimo *et al.*, 2005), and the other seven unlinked loci were chosen for their putative neutrality (stn30, stn173, stn196, stn174, stn185, stn70 and stn199; Peichel *et al.*, 2001). Stn381 is a diagnostic in/del marker for *Eda*, with 'low' and 'complete' alleles that have been shown to be associated with plate count variation (Colosimo *et al.*, 2004). In contrast, although regulatory mutations at *Pitx1* are associated with pelvic spine reduction allelic variation at stn82, a nonintergenic marker, is not directly associated with pelvic spine length (Shapiro *et al.*, 2004; Chan *et al.*, 2010). Nevertheless, stn82 remains a useful marker to test for the effect of selection on *Pitx1* (Mäkinen *et al.*, 2008). Polymerase chain reactions (PCRs) were prepared using the Type-it Microsatellite PCR kit (Qiagen Inc., Valencia, CA, USA) following the manufacturer's protocol. All PCRs were carried out on an Eppendorf™ Mastercycler™ Pro with cycling conditions standardized for all loci: denaturation at 95 °C for 5 min and 28 cycles at 95 °C for 30 s, 60 °C 90 s, 72 °C 30 s and then cooled at 4 °C. The resulting products were sequenced using a ABI 3730XL sequencer at Génome Québec (Montréal, Canada) with a 5-min denaturation step at 95 °C before injection. Peak call analysis was performed using Geneious version 8.8.1 (Biomatters Ltd., Auckland, New Zealand) using the Microsatellite Analysis External Plugin version 1.4.0. To compare the focal estuary populations to a pure marine type, we amplified the same loci on 30 fish from a pure marine population collected from Bodega Bay (Sonoma County, CA, USA).

Using GENEPOP version 4.5.1 (Rousset, 2008), we first tested each neutral locus (those not linked to *Eda* and *Pitx1*) for departures from Hardy–Weinberg equilibrium and for potential linkage between loci after

Bonferroni correction ($\alpha = 0.05$, $K = 601$). A *G*-test (Goudet *et al.*, 1996) performed with the R package *adegenet* (Jombart, 2008) with 99 simulations showed that no F values were greater than expected by chance (simulated $P < 0.01$). With the same R package, we then calculated Nei's pairwise F_{ST} estimates (Nei, 1973).

We explored population structure through several complementary analyses. (i) We performed a correspondence analysis (CA) based on allele frequencies at the seven neutral markers, replacing missing values by the mean of the allele frequency of each locus (similar results were obtained using PCA). (ii) We used STRUCTURE version 2.3.4 (Pritchard *et al.*, 2000) with the admixture model with 10 000 repetitions for burn-in and 200 000 for run length over 10 iterations for $K = 1$–24. We determined the most likely value of K by taking the averaged log-likelihoods across the 24 runs and applying the ΔK method (Evanno *et al.*, 2005). (iii) We performed an analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) on all estuary populations (i.e. excluding Bodega Bay) with the R package *poppr* version 2.2.0 (Kamvar *et al.*, 2014), testing significance by randomly permuting the sample matrices over 500 iterations. (iv) Based on allele frequencies at the seven neutral markers, we calculated – between all population pairs – Edward's genetic distance (Edwards, 1971), which assumes that allele frequencies differ because of drift. These distances were used to compute a hierarchical clustering analysis and build a genetic tree. (5) We tested for isolation by distance between estuary populations (i.e. excluding Bodega Bay) by first computing a matrix of geographic distances based on latitudinal and longitudinal coordinates. We then used a Mantel test (Mantel, 1967) with 999 permutations comparing pairwise Edward's distance to pairwise geographic distance.

Although the assumption that California estuaries represent potential hybrid zones between marine and upstream freshwater population has been historically rejected (Bell, 1976, 1979a, b; Bell, 1981, 1982; Baumgartner & Bell, 1984; Baumgartner, 1986, 1992, 1994; Bell & Richkind, 2015), we used our genetic data to confirm this interpretation for our contemporary samples. Within each population, we selected fish homozygote at the 'complete' *Eda* allele and tested whether those fish were more likely to be assigned to the neutral marine genetic cluster of Bodega Bay. For this inference, we used STRUCTURE version 2.3.4 (Pritchard *et al.*, 2000) with the admixture model with 10 000 repetitions for burn-in and 200 000 for run length over 10 iterations for $K = 1$–19 (five populations did not have any fish homozygote 'complete' at *Eda*). We determined the most likely value of K by taking the averaged log-likelihoods across the 19 runs and applying the ΔK method (Evanno *et al.*, 2005). As described above, we considered whether (as would be expected for hybrid zones) our populations were out of Hardy–Weinberg equilibrium at neutral loci. If fish with the homozygote 'complete' at *Eda* do not cluster with the Bodega Bay neutral marine population cluster, and if our populations are in Hardy–Weinberg equilibrium at neutral markers, then our estuaries are – as historically inferred – not hybrid zones.

Divergence associated with predator regimes

To test whether *Eda* and *Pitx1* have experienced divergent selection among estuaries, we used an F_{ST}-outlier detection method implemented in LOSITAN version 1.44 (Antao *et al.*, 2008). Lositan is an allele frequency-based method that identifies outliers from the joint distribution of F_{ST} and expected heterozygosity, using coalescent simulations to determine the F_{ST} null distributions and assuming an island model. In this analysis, the distribution of F_{ST} is characterized by estimating the quantiles of the distribution and defining a window in which 95% of the data points are expected to lie (Beaumont & Nichols, 1996). Based on the simulated distribution, it is possible to calculate *P*-values for loci of interest. Loci with a high F_{ST} value are putatively under directional selection (P-value > 0.975), whereas loci with a low F_{ST} value are putatively under balancing selection (P-value < 0.025). We used the infinite alleles model with 50 000 simulations, a 95% confidence interval and a false discovery rate of 0.1. Finally, we tested for associations between particular *Eda* alleles and predator regime by regressing the 'complete' allele frequency (*Eda* C allele), which is strongly associated with high plate counts (Colosimo *et al.*, 2005), in a given population against the environmental predictors of watershed area, presence of steelhead and presence of sculpin.

Univariate morphometrics

We first took ventral and left lateral photographs of all stained fish with a Canon EOS Rebel *X3i* digital camera fitted with a 50-mm lens under standardized light conditions with a millimetre ruler in the image for scale. Small pins were inserted into the fish to help indicate anatomical points for placing digital landmarks (e.g. Kaeuffer *et al.*, 2012). We then blotted the fish dry and measured mass to the nearest hundredth of a gram on an electronic balance. We next used digital callipers to measure, to the nearest hundredth of a millimetre: standard length from the tip of the upper jaw to the end of the vertebral column on the caudal peduncle, the lengths of the first and second dorsal spine and the length of the left pelvic spine. We also counted lateral plates on each side of the fish under a dissecting microscope, not including any keel plates at the end of the caudal peduncle (Bell, 1981). Finally, we dissected all fish and inspected the gonads to visually identify sex.

Morphological analyses were performed on up to 30 individuals per population of standard length > 30 mm

(Table S1). All spine length measurements were standardized to a common body size following the allometric approach: $M_S = M_o(L_S/L_o)^b$, where M_S is the standardized spine length measurement, M_o is the unstandardized spine length measurement, L_S is the overall mean body length of all fish, and L_o is the body length of the individual (Lleonart *et al.*, 2000). The exponent b was calculated as the common within-group slope from a linear mixed-effects model regressing $\log_{10}(M_o)$ on $\log_{10}(L_o)$ with population as the random factor (Reist, 1986; Hendry & Taylor, 2004).

We used linear mixed-effects models to find the best set of predictors for the length of each size-corrected spine using the R package nlme (Pinheiro *et al.*, 2016). We included a random intercept term for population and fixed terms for watershed area, presence of steelhead and presence of sculpin. As the larger predatory fishes tended to be found in estuaries with larger watersheds (and therefore more upstream habitat), we tested for multicollinearity of predictors by examining variance inflation factors (VIFs). All VIFs were within acceptable limits: VIF < 3 (Zuur *et al.*, 2009). Log-transformed plate counts were analysed in a separate model with the same structure as above.

Geometric morphometrics

We placed 18 homologous landmarks on the lateral photographs using tpsDig software (Rohlf, 2006; Fig. S2; Table S1). Immature fish and fish with large internal parasites were discarded from the analysis. The 18 landmarks were then superimposed using the generalized Procrustes analysis of *geomorph* (Adams & Otarolla-Castillo, 2013), yielding 36 Procrustes residuals representing shape differences among individuals after removing effects of (isometric) scale, rotation and translation. A Procrustes ANOVA (Goodall, 1991; Adams & Otarolla-Castillo, 2013) using body shape as the response variable and sex as the predictor variable revealed a significant effect of sex ($F = 62.14$, $P < 0.01$). To correct for this effect, residuals from this Procrustes ANOVA were added to the mean consensus shape of all individuals. This sexual dimorphism-free shape dataset was used for further analysis.

We performed a multivariate analysis of variance (MANOVA) using Wilks' lambda (λ) as the test statistic. The PCs derived from the 36 Procrustes residuals were allometrically adjusted for centroid size and body depth using the common within-group slope approach described above (Reist, 1986; Lleonart *et al.*, 2000; Rolshausen *et al.*, 2015). The PCs were then used as the dependent variables with presence of steelhead, presence of sculpin and population as fixed explanatory variables. We performed a canonical variates analysis (CVA) using fish facing different predator regimes as separate factors (Webster & Sheets, 2010). This method allows for the identification of different patterns of shape among populations by providing an ordination of the population in morphological space (Leinonen *et al.*, 2006). Thus, the canonical vector (or divergence vector) extracted from this analysis maximizes the morphometric variance for a specific factor (here predator presence/absence). We used the mean individual scores from this divergence vector for each population to visualize body shape differences along this factor.

Elemental composition

Whole fish elemental composition was analysed for 10 fish from each of 15 populations, except for Gazos Creek ($N = 9$) and Younger Lagoon ($N = 20$; Table S1). These fish were different individuals from those analysed above because the two analysis procedures were incompatible on the same fish. Individuals analysed for elemental composition came from estuaries where the two predator types (steelhead and sculpin) were either both present or both absent. We quantified the following phenotypes for each of these fish: standard length, head length (cm), body depth (cm), pelvis length (combined length of anterior and posterior processes, in cm) and lateral plate count (left side). For these traits, we then applied the allometric standardizations as described above (Reist, 1986; Hendry & Taylor, 2004).

Digestive and reproductive tissues were discarded prior to elemental analysis (El-Sabaawi *et al.*, 2012). Stickleback specimens were freeze-dried for 72 h using a LABCONCO 77545-00-J (Kansas City, MO, USA). Dry mass was then recorded and relative condition calculated based on the length–mass relationship (Froese, 2006). Phosphorus content (%P) was determined as the mean of three 9–11 mg subsamples of the ground body tissue. These samples were ashed at 500 °C for 2 h and digested with 1N HCl at 105 °C for 2 h before assay with a Mandel UVmini-1240 spectrophotometer using an acid molybdate method (Murphy and Riley 1962). The mean coefficient of variance was < 3% between fish replicates, and extraction efficiency was > 95% for bonemeal (NIST 1486) and spinach (NIST 1570a) standards.

Two different statistical inferences were explored. First, to test whether lateral plates or *Eda* genotypes predict elemental composition, we created two GLMMs. The first GLMM used only size-corrected phenotypic traits as main effects (standard length, pelvis length, head length, body depth, condition and lateral plate count), whereas the second replaced lateral plate count with *Eda* genotype. All models included population as a random effect, and collinearity was again (as above) within acceptable limits. We then used an AICc-based model search conducted in the MuMIn package to select the best model from each global model (Grueber *et al.*, 2011; Bartoń, 2016). Second, we used GLM to test whether the presence of predatory fish (fixed factor) is associated with stickleback %P, with condition as the only other predictor.

Comparing within- and among-population variation to other stickleback systems

We first verified whether our estuary populations would display greater levels of within than among-population variation. For each trait, we calculated the proportion of the total variation attributable to within vs. between-population variation in our system using a nested ANOVA with trait as the dependent variable and individuals nested in populations as the predictor variable. Within- and among-population variance explained (η^2) was calculated by dividing the sum of squares of each fixed term (individual nested in population and population, respectively) by the total sum of squares. We tested for differences in percentage of variance explained across traits using a two-sided *t* test.

To test whether strong temporal environmental fluctuations would lead to high levels of trait variation, we compared levels of within-population variation in our estuaries to within-population variation from stickleback populations that experience comparatively lower temporal environmental variation. We are not asserting here that populations from these other systems are completely temporally stable, but that they are typically less variable than those in bar-built estuaries subject to dramatic and rapid breaching events, which can lead to extreme changes in abiotic and biotic conditions over a period of hours (Fig. 1b, Fig. S1). Using Tukey's honest significance tests, we compared our within-population standard deviation values with equivalent within-population values from temporally stable lakes, streams and marine environments reported in the literature (Table S7; Whitlock & Schluter, 2009).

Finally, we tested whether environmental variation would lead to relatively lower between-population differences by comparing levels of among-population variation in plate counts in our system to among-population variation from relatively temporally stable stickleback populations experiencing divergent predator regimes (Table S8). To calculate among-population variation, we computed ANOVAS in each system separately with mean population plate counts as dependent variable and populations as predictor. Among-population variation was calculated by dividing the population term sum of squares by the total number of populations in each system, respectively.

Results

Population genetics

No indication of linkage disequilibrium was found between our microsatellite markers (Fisher's exact test, average $\chi^2 = 29.37$, average d.f. = 43.33 and average *P* between pairwise testing = 0.83), as was expected based on their positions on separate linkage groups (Peichel *et al.*, 2001). The markers also showed no within-population departures from Hardy–Weinberg equilibrium after Bonferroni correction.

Correspondence analysis did not reveal obvious discontinuous structuring of the estuary populations (Fig. S3) – apart from our reference marine population, which was a clear outlier (results for the other estuary populations did not change when excluding the marine population). However, STRUCTURE revealed two somewhat distinct groups with the ΔK method identifying two clusters as most likely (Fig. 2 and Table S2 for F_{ST}-based measures of pairwise genetic differentiation). At one end of the spectrum was the marine population composed almost entirely of genotypes from that cluster. At the other end of the spectrum were Lombardi Creek, Old Dairy Creek and Younger Lagoon composed mostly of genotypes from the other cluster. These later three populations were geographically close to each other and had smaller watershed areas as compared to the other estuary populations (mean of 3.4 ± 3 km^2 and 414.9 ± 1015 km^2, respectively). Smaller watersheds tend to have lower stream flows and therefore spend shorter periods of time with the estuary mouth open, suggesting that these populations will be less often connected to the ocean, thus explaining their partial genetic isolation from other populations. Their geographic proximity also means that they are likely to breach at similar times and then exchange migrants with each other, thus explaining their genetic similarity to each other. The other populations contained a variable mixture of alleles from the two clusters. AMOVA revealed *Phi* (*Ø*) statistics below 0.2 (Table S3), confirming low population differentiation that was nevertheless significant (Table S4, Fig. S4). The hierarchical clustering tree showed again that the marine population from Bodega Bay was distinct from the estuary populations, with the estuaries appearing to branch mostly based on geographic proximity (Fig. S5). The Mantel test performed on the estuary populations alone (excluding the marine population) revealed low but significant isolation by distance (Fig. S6 simulated $P = 0.02$).

As noted above, our neutral markers showed no signs of deviations from Hardy–Weinberg equilibrium. Further, when considering only fish homozygote 'complete' at *Eda*, STRUCTURE revealed two distinct groups with the ΔK method identifying two clusters as most likely: one representing the genetic cluster of Bodega Bay and the other representing our estuaries (Fig. S7). Both outcomes support historical analyses in inferring that stickleback in bar-built estuaries are coherent populations, as opposed to hybrid zones.

Divergence associated with predator regimes

LOSITAN revealed that *Eda* was putatively under directional selection ($H_e = 0.88$, $F_{ST} = 0.12$, $P_{Simul.\ Fst < sample\ Fst} = 0.97$), whereas *Pitx1* was putatively under balancing selection ($H_e = 0.45$, $F_{ST} = 0.25$, $P_{Simul.\ Fst < sample\ Fst} < 0.02$). Stickleback in estuaries with sculpin showed a higher frequency of the C allele at *Eda* than did

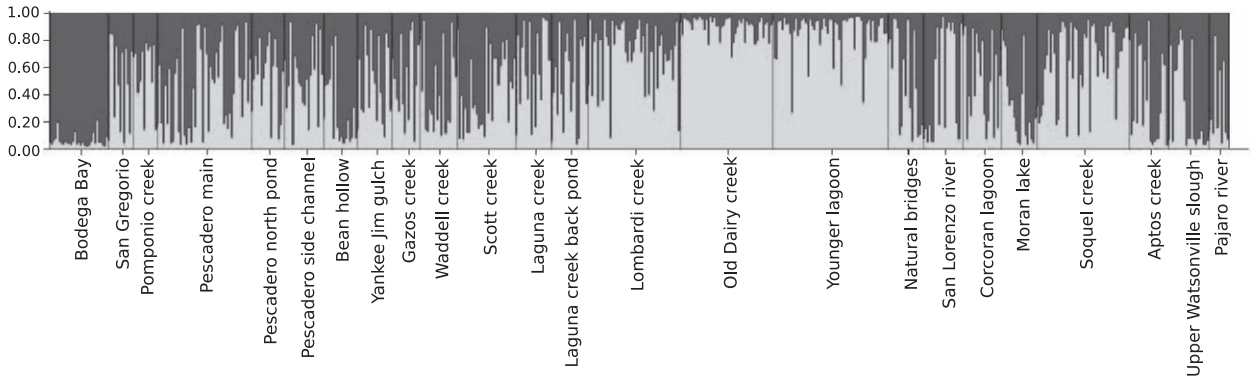

Fig. 2 Individual assignment to population structure inferred by STRUCTURE. Each bar represents an individual. The *y*-axis represents the probability of classification to a cluster.

stickleback in estuaries without sculpin (mean across populations: 0.46 vs. 0.18; Table 1, Fig. 3).

Univariate morphometrics

Mixed models with population as a random effect significantly improved the fit of linear models for spine length and plate count as measured by a likelihood ratio test (Table S5). None of our fixed predictors (presence of sculpin, presence of steelhead and watershed size) for the length of the first dorsal spine were significant (Table 1). However, stickleback had somewhat longer second dorsal and pelvic spines, as well as more lateral plates, in estuaries with sculpin than in estuaries without sculpin (Table 1, Fig. 3).

Geometric morphometrics

The first two axes explained 49% of the total shape variation (33% for PC1 and 16% for PC2), with both axes mainly related to body depth. In particular, stickleback scoring negatively were shallower bodied whereas fish scoring positively were deeper bodied, in the posterior part of the body (PC1) or the anterior part of the body (PC2) (Fig. S8). MANOVA on all 36 PCs revealed a significant influence of population ($\lambda = 0.01$, d.f. = 22, $F = 2.61$ and $P < 0.01$), sculpin ($\lambda = 0.71$, d.f. = 1, $F = 4.99$ and $P < 0.01$) and steelhead ($\lambda = 0.83$, d.f. = 1, $F = 2.37$ and $P < 0.01$). Testing the effect of presence vs. absence of each predator type alone yielded a similar outcome ($\lambda = 0.74$, d.f. = 1, $F = 4.31$ and $P < 0.01$). Overall, stickleback tend to be shallower bodied in the presence than absence of predatory fishes (Fig. 4), although most estuaries showed a great diversity of body shapes, with individuals scoring positively and negatively.

Elemental composition

Phosphorus content ranged from 2.8% to 6.9% among the collected stickleback. In the best phenotypic model (using plate number rather than *Eda* genotype), five main effects explained over one-third of the total variation ($R^2_{\text{Marg.}} = 0.35$) and, when combined with population as a random effect, explained double that ($R^2_{\text{Cond.}} = 0.72$). Of these factors, condition had the largest effect on %P ($P < 0.001$), with high condition fish showing reduced phosphorus content (Table S6). Standard length, head length and lateral plate count were also significant predictors of %P ($P < 0.001$) and had similar effect sizes ($\eta^2 = 0.35$–0.50, Table S6). In each case, %P was positively correlated with trait values (Fig. S9). The best genotypic model (using *Eda* genotype rather than plate number) showed similar relationships and explanatory power (Table S6, Fig. 5a). In this case, six main effects explained 0.42% of the variation and, when combined with population as a random effect explained 0.77%. Again, condition had the

Table 1 Results of mixed-models analysis testing the effect of the presence of sculpin, steelhead and watershed size.

Response	Predictor	Coef.	SE	*T*-value	*P*-value
First spine length	Sculpin	0.41	0.26	1.59	0.121
	Steelhead	−0.11	0.26	−0.42	0.656
	Watershed size	0.01	0.01	1.12	0.281
Second spine length	Sculpin	0.45	0.22	2.07	0.052
	Steelhead	−0.13	0.22	−0.59	0.543
	Watershed size	0.01	0.01	0.77	0.439
Pelvic spine length	Sculpin	0.58	0.32	1.82	0.081
	Steelhead	−0.21	0.33	−0.62	0.535
	Watershed size	−0.01	0.01	−0.95	0.419
Log plate count	Sculpin	0.73	0.23	3.14	**0.005**
	Steelhead	−0.22	0.24	−0.89	0.382
	Watershed size	−0.01	0.01	−1.45	0.165
C allele frequency	Sculpin	0.43	0.16	2.75	**0.013**
	Steelhead	−0.15	0.16	−0.93	0.363
	Watershed size	−0.01	0.01	−0.94	0.359

Coefficient (Coef.), standard error (SE) *T* and *P*-values are reported. d.f. were 19 for all variables. Intercepts and random effects are not shown. $P \leq 0.05$ are in bold.

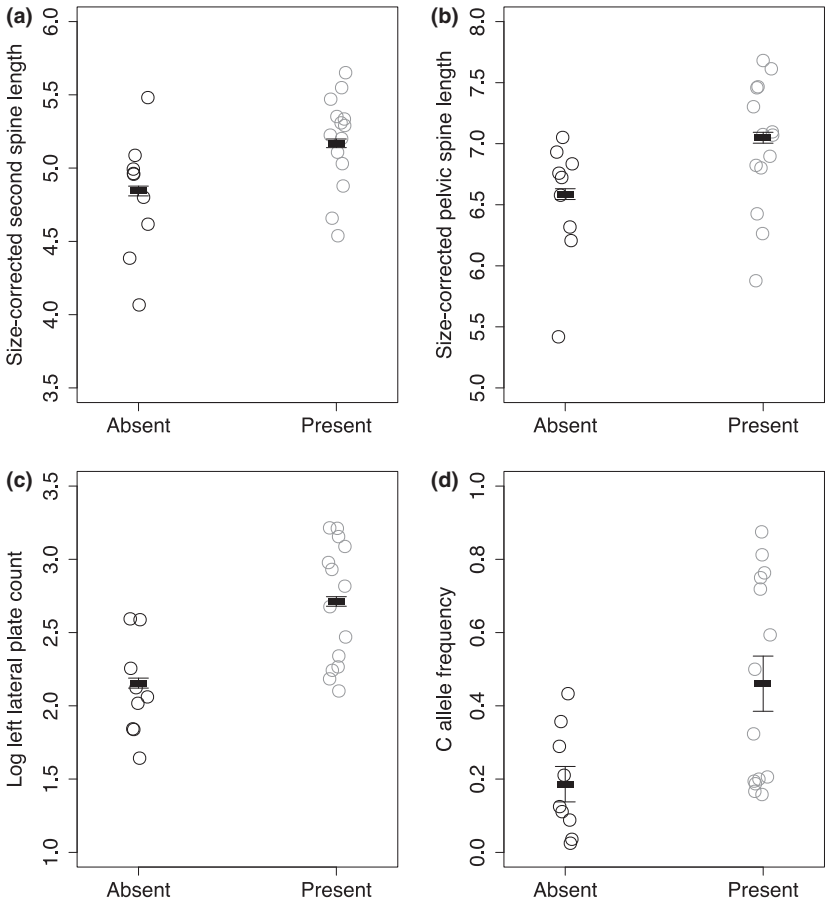

Fig. 3 Armor morphology in the presence or absence of sculpin. (a) Size-corrected first dorsal spine length. (b) Size-corrected pelvic spine length. (c) Log left lateral plate count. (d) *Eda* complete allele frequency. Each circle depicts the mean value of a particular estuary population in the absence (black) or presence (grey) of sculpin. Bars represent the overall mean value (± SE) in the presence of absence of sculpin.

largest effect ($\eta^2 = -1.11$) with *Eda* genotype having the second largest effect ($\eta^2 = -0.68$ for LL vs CC genotypes). Predation and condition influenced %P differences among populations ($\eta^2 = 0.24$, $P < 0.001$), predation having a greater effect than condition ($\eta^2 = 0.24$ vs 0.15). Across the 15 populations, those in estuaries with predators were 20% higher in %P (5.1% vs 4.2%; Fig. 5b).

Within- and among-population variation

In our study system, the proportion of variance explained (η^2) was significantly greater within than among populations for all traits, except %P ($t = -2.72$, d.f. = 12, $P < 0.01$, Fig. S10). Within-population variation in plate count, *Eda* complete allelic count, and shape was significantly greater in our Santa Cruz estuaries than in presumed more stable environments documented in the literature, except in lakes for the *Eda* complete allelic count (Table 2, Fig. 6a–c). Among-population variation in plate counts was lower in Santa Cruz populations than in other systems, including systems with populations exposed to divergent predator regimes (Haida Gwaii and Vancouver Island, Table 3, Fig. 6d; Reimchen *et al.*, 2013; Miller *et al.*, 2015). Note that, as compared to our bar-built system, the other systems used in this comparative analysis face much lower gene flow. For instance, the lakes in Québec and on Vancouver Island are completely geographically isolated from each other, ensuring no gene flow between populations (Lacasse & Aubin-Horth, 2012; Miller *et al.*, 2015). For Alaska, most of the populations reported in Table 3 are also geographically isolated, except for those present in the Matanuska-Susitna valley, which nevertheless have a mean F_{ST} much greater (0.111: Bell & Orti, 1994; Aguirre, 2009; W.E. Aguirre, 2010 unpublished data) than in our system (0.003). The same is true from populations from North Uist in Scotland (mean F_{ST} of 0.199).

Discussion

We considered potential consequences of the extreme temporal environmental variation present in bar-built estuaries for within- and among-population variation in stickleback armor traits and their potential ecological effects. We first describe our main results and then discuss the nuances and implications in more detail. First, stickleback gene flow was high among many of the estuaries, but not so high as to entirely prevent divergence in armor traits in response to different predation

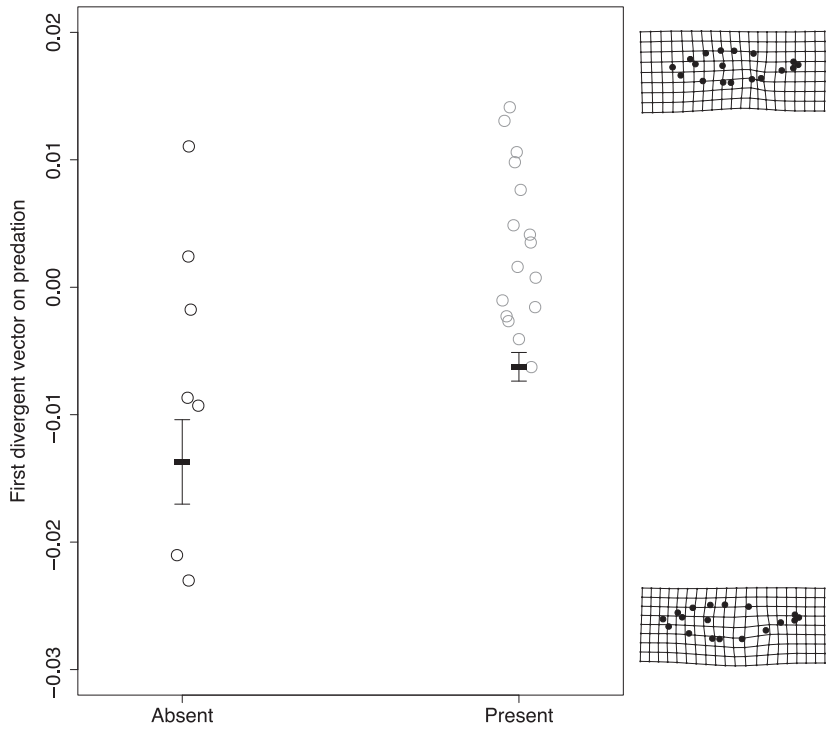

Fig. 4 Divergence scores extracted from the first divergent vector of each population and obtained through a canonical vector analysis (CVA). Each circle depicts the mean value of a particular estuary population in the absence (black) or presence (grey) of predators. Bars represent the overall mean value (± SE) in the presence of absence of predators. Populations with mean negative divergence scores have deeper bodies whereas population scoring positively are more streamlined. Thin-plate spline transformation grids of CVA divergent vectors display the shape difference between positive and negative scores.

regimes. Second, this divergence in armor traits was – as expected from the high gene flow – generally weaker than that observed in other (not bar-built) systems, including among stickleback populations exposed to divergent predator regimes in more temporally stable environments. Third, within-population variation was very high for stickleback in the estuaries, including in comparison to stickleback from other study systems where temporal environmental variation is presumably lower. Fourth, an essential element for ecological stoichiometry (%P) – a trait potentially linked to the ecological effects of stickleback – was strongly associated with armor traits and *Eda* allele frequency. Overall, our results suggest that strong temporal environmental variation – in conjunction with high gene flow – can have important consequences for within- and among-population variation in adaptive traits, and the potential ecological effects of those traits.

Population structure reveals high gene flow between estuaries

Despite frequent breaching events that disrupt the isolation of estuary populations (Allen *et al.*, 2006), we detected some evidence for population structure across the system. The greatest contribution to this structure was that stickleback in several estuaries were clearly distinct from the Bodega Bay marine population, with stickleback in the other estuaries showing apparent admixture between the two genotypic clusters (Fig. 2, Fig. S5). These results concur with the expectation that breaching events promote dispersal between bar-built estuary stickleback and marine stickleback, but not so much as to prevent the latter from diverging genetically in at least some cases. Consistent with this interpretation, we detected weak but significant isolation by distance (Table S2, Fig. S6) and population differentiation (Table S4, Fig. S4), indicating the potential for adaptive divergence among populations. However, it was also clear that many of the estuaries experienced high gene flow with each other and with marine stickleback. Together, these results indicate that gene flow between the bar-built estuaries along this coast is sufficiently low to allow population divergence in at least some cases, but also sufficiently high to constrain the magnitude of that divergence.

Trait differentiation is associated with divergent predator regimes

Spatial variation in the presence of piscivorous fishes was correlated with spatial variation in stickleback armor traits. In particular, when sculpin were present, stickleback had slightly longer spines, more lateral plates, shallower bodies and a higher frequency of the complete *Eda* allele (Fig. 2). Sculpin are well-known predators of stickleback and prey on eggs, fry and adults (Moodie, 1972; Pressley, 1981; Reimchen, 1994; Ingram *et al.*, 2012). These findings parallel many previous studies of stickleback, where populations experiencing greater levels of predation from fish display longer spines, more lateral plates (and therefore a higher

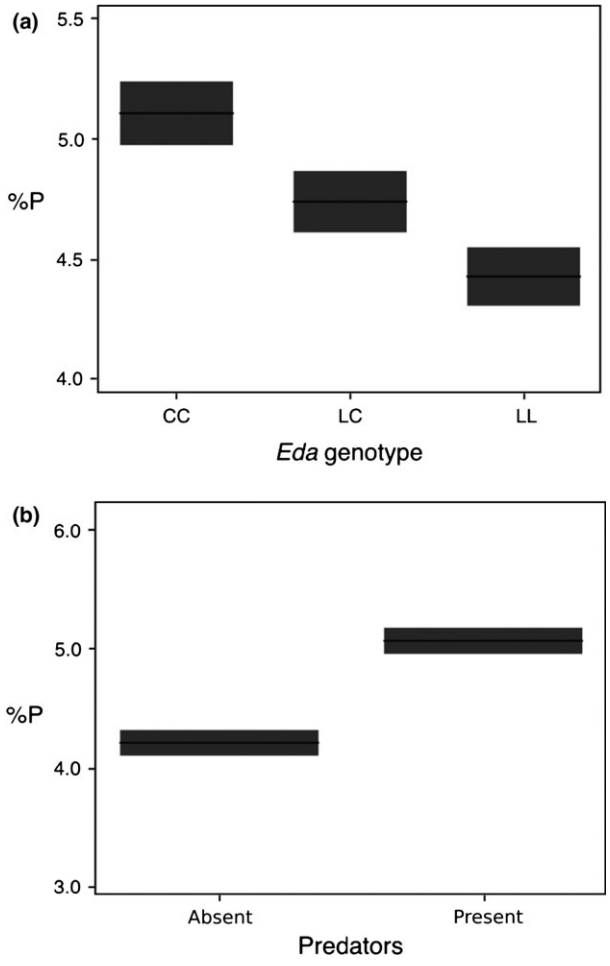

Fig. 5 Modelled relationship between %P and *Eda* from 'phenotype + *Eda*' GLMM (a) and between %P and predation (b). Shaded regions depict ± 1 SE from mean.

frequency of the complete *Eda* allele) and shallower bodies (Reimchen, 1992, 1994; Leinonen *et al.*, 2011; Lescak & von Hippel, 2011), with these patterns being especially strong in the presence of sculpin (Ingram *et al.*, 2012; Miller *et al.*, 2015). In our study, however, the presence of sculpin only modestly affected spine length. Perhaps, one contributor to this comparative subtlety is that longer spines will be less effective against predators without significant gape limitation, such as the Pacific staghorn sculpin (*Leptocottus amatus*), which are able to swallow stickleback with large spines (Moyle, 1976; Hyatt, 1979). Therefore, the only modest effect of sculpin presence on spine length differentiation between estuaries could be due to this trait not providing an effective defence against the functional capabilities of the local predators. Taken together, these results show, despite extreme temporal variation in environmental conditions and high gene flow among estuaries, spatial variation consistent with local adaptation was evident in stickleback armor traits.

Beyond phenotypes, genetic markers associated with *Eda* and *Pitx1* showed evidence for directional and balancing selection, respectively. Consistent with the above results for lateral plates, the frequency of the complete *Eda* allele was higher in the presence of sculpin (Fig. 3d). This pattern is consistent with predation-induced selection, similar to that documented in previous studies of other stickleback systems (Marchinko, 2009; Zeller *et al.*, 2012; Raeymaekers *et al.*, 2014). Although phenotypic plasticity could explain some of this variation in armor phenotypes, its role is likely minimal given that *Eda* explains about 75% of the variation in plate counts (Colosimo *et al.*, 2004; Kitano *et al.*, 2008). Thus, the inferred directional selection at *Eda* likely reflects the importance of lateral plate defence against the predatory sculpin. Interpretations for *Pitx1* are quite different. In other stickleback

Table 2 Results of Tukey *post hoc* test testing for differences in standard deviations between our estuary populations and environmentally stable lake, marine and stream populations for plate counts, *Eda* complete allele count and procrustes variance.

	Plate count				Complete *Eda* allele count				Procrustes variance			
	d.f.	Sum.Sq	Mean.Sq	*F*-value	d.f.	Sum.Sq	Mean.Sq	*F*-value	d.f.	Sum.Sq	Mean.Sq	*F*-value
ANOVA	3	509.30	169.80	**77.06**	3	1.18	0.39	**8.13**	2	0	0	**8.28**
	Diff.	Lower	Upper	*P*-value	Diff.	Lower	Upper	*P*-value	Diff.	Lower	Upper	*P*-value
Estuary–Lake	−5.154	−6.038	−4.269	**< 0.001**	0.104	−0.098	0.306	0.905	−0.006	−0.012	−0.002	**0.010**
Estuary–Marine	−4.246	−5.828	−2.664	**< 0.001**	−0.484	−0.878	−0.091	**0.011**	*na*	*na*	*na*	*na*
Estuary–Stream	−4.787	−5.983	−3.592	**< 0.001**	−0.263	−0.554	0.027	0.087	0.000	−0.012	−0.006	**0.007**
Marine–Lake	0.908	−0.504	2.321	0.343	−0.389	−0.642	−0.138	**< 0.001**	*na*	*na*	*na*	*na*
Marine–Stream	−0.542	−2.166	1.083	0.823	0.073	−0.248	0.395	0.929	*na*	*na*	*na*	*na*
Stream–Lake	0.367	−0.591	1.325	0.753	−0.316	−0.629	−0.003	**0.04**	−0.000	−0.006	0.006	0.991

Mean differences (Diff.), 95% confidence intervals (lower and upper) and *P*-values are reported. Degrees of freedom (d.f.), sum of squares (Sum.Sq), mean sum of squares (Mean.Sq) and *F*-values are reported for a standard ANOVA. $P < 0.05$ and significant *F*-values are in bold.

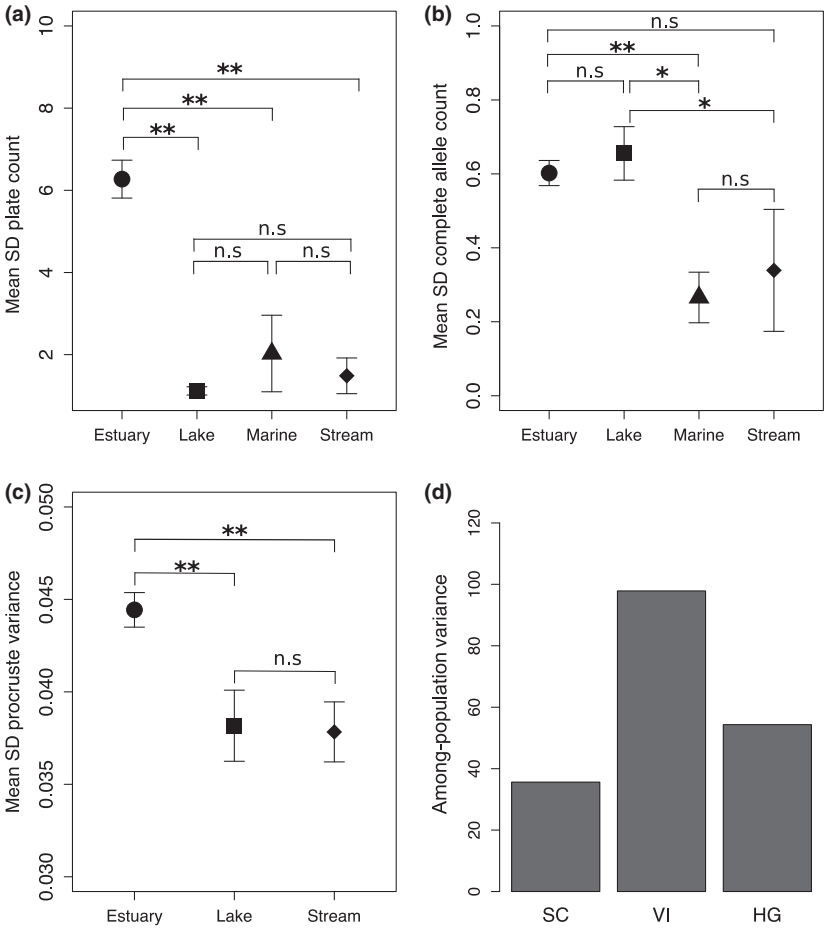

Fig. 6 Within-population mean standard deviations (SD) between Santa Cruz estuaries and less temporally variable lake, marine and stream environments (± SE, panels a–c) and among-population variance in plate counts in the Santa Cruz estuaries (SC), Vancouver Island (VI) and Haida Gwaii (HG) (panel d). *P*-values < 0.01 are presented by two stars, and *P*-values < 0.05 by one star. Nonsignificant differences are represented by n.s.

Table 3 Among-population variance (Variance), total number of populations (*N*) and population sum of squares (Sum.Sq) in the Santa Cruz, Vancouver Island, Haida Gwaii and Iceland systems.

	N	Sum.Sq	Variance
Santa Cruz	23	818.88	35.61
Vancouver Island	49	6002.83	97.86
Haida Gwaii	30	1630.06	54.34
Iceland	10	501.23	50.13

systems, regulatory mutations at *Pitx1* are generally associated with molecular signatures of positive directional selection in pelvic-reduced populations that colonized freshwater from the ocean (Chan *et al.*, 2010). In contrast, we detected evidence of balancing selection at this locus. Balancing selection is thought to be an important mechanism responsible for the maintenance of genetic polymorphism (Hedrick, 1986), especially in heterogeneous environments (Hedrick, 1986; Spichtig & Kawecki, 2004). Thus, whereas patterns for lateral plates likely reflect consistent directional selection on a defensive trait owing to spatial variation in predatory fishes, balancing selection at *Pitx1* could be reflective of the temporal fluctuations in environmental conditions present in these estuaries.

At the same time, it is important to recognize that population divergence in the bar-built system is considerably weaker than that in other stickleback systems (Table 3, Fig. 6d). This contrast among systems is consistent with the expected effects of both temporal variation and gene flow. First, when temporal variation is high, spatial differences are expected to be compromised, as suggested by some previous theoretical and empirical analyses (Kawecki & Ebert, 2004; Siepielski *et al.*, 2009; Bell, 2010; Chevin *et al.*, 2015). Second, when gene flow is high, spatial population divergence is often low, as shown in theory (Slatkin, 1973; Felsenstein, 1976; Kawecki, 2008) and empirical systems including stickleback (e.g. Hendry & Taylor, 2004; Stuart *et al.*, 2017).

One additional consideration is that the relatively high within-population variation observed in these estuary populations could occur because they represent a hybrid zone between marine and stream freshwater populations (e.g. Jones *et al.*, 2006; Vines *et al.*, 2016).

This hypothesis was historically investigated and rejected (Bell, 1976, 1979a, b, 1981, 1982; Baumgartner & Bell, 1984; Baumgartner, 1986, 1992, 1994; Bell & Richkind, 2015). Indeed, a freshwater form was never found upstream of California estuaries, and plate counts were – in fact – often greater upstream than downstream (Bell, 1976, 1979a, b, 1981, 1982). In addition, all of our neutral markers showed no departure from Hardy–Weinberg equilibrium and individuals homozygote 'complete' at *Eda* did not group with the neutral marine cluster of Bodega Bay (Fig. S7). These results confirm historical evidence that our estuaries do not represent hybrid zones but rather coherent populations in their own right. In summary, spatial patterns of phenotypic and genetic variation for stickleback in bar-built estuaries match some important aspects of previous studies, while also suggesting additional nuances and effects.

Trait variation within populations

We found that stickleback in bar-built estuaries of the central California coast exhibits very high levels of within-population variation. This result held for all traits, ranging from spine length to body shape to plate count to *Eda* genotype to %P (Fig. S10). This within-population variation appears much greater than that documented in previous stickleback studies that focused on populations in presumably more stable environments (Table 3, Table S7, Fig. 6a–c). An exception that could prove the rule is the very low among-population differentiation and very high within-population variation in stickleback from ephemeral streams and adjacent vineyard reservoirs in Napa, California (Hendry *et al.*, 2013), another system where temporal environmental variation (and likely gene flow) is extremely high. These differences among systems are consistent with arguments that constantly shifting environmental conditions prevent temporally consistent selection, thereby impeding the ability of directional selection to eliminate variation from the populations (Bell, 2010; Michel *et al.*, 2014). Valuable additional steps would be to examine the fitness consequences of this high genetic variation – such variation could impose a substantial genetic load on populations (Lande & Shannon, 1996; Arnold *et al.*, 2001). On the other hand, high genetic variation should maintain the potential for strong selection and rapid evolutionary responses, which could aid responses to future environmental changes (Mackay, 1981; Kirkpatrick & Barton, 1997; Kawecki & Ebert, 2004).

As alluded to several times already, there are two likely mechanisms driving the observed high within-population and low among-population variation: high temporal environmental variation and high gene flow. Although gene flow could certainly contribute to reduced divergence – as has been inferred by our group for other stickleback systems (e.g. Hendry & Taylor, 2004; Stuart *et al.*, 2017) – we do not think that this mechanism alone explains patterns of variation in the bar-built system. The reason is that high gene flow is most effective at maintaining high within-population variation if among-population variation is also high. In the bar-built system, however, among-population variation is low (Table 3, Fig. 6), which means that gene flow will not be moving novel variants among estuaries and inflating the variation within each of those populations. Hence, we suggest that high temporal variation is responsible for the observed high within-population variance and low among-population variance, as also suggested by some previous theoretical and empirical analyses (Kawecki & Ebert, 2004; Siepielski *et al.*, 2009; Bell, 2010; Chevin *et al.*, 2015).

Elemental composition

Previous studies have shown that investment in bony structures can increase phosphorus demand, which can potentially alter how fish forage and recycle nutrients (El-Sabaawi *et al.*, 2016; Durston & El-Sabaawi, 2017; Leal *et al.*, 2017). We find that, despite dramatic environmental fluctuations, the expected association between %P and armor remains strong. Moreover, we find that genotypes at a single locus (*Eda*) explain a large amount of the variation in %P (Fig. 5a), which is not surprising given that variation in *Eda* explains much of the variation in lateral plates (Colosimo *et al.*, 2004). Importantly, lateral plates and *Eda* vary dramatically within and among the estuaries, generating the high levels of variation in %P. This variation should have a major influence on whole fish elemental ratios and thus the observed variation in %P is likely to influence the balance between excretion rates and diet choice (El-Sabaawi *et al.*, 2016; Durston & El-Sabaawi, 2017). Given that allelic variation at *Eda* appears to be driving variation in elemental composition, and because *Eda* is generally subject to strong natural selection in the wild (Colosimo *et al.*, 2004, 2005; Barrett *et al.*, 2008; Jones *et al.*, 2012), it is likely that elemental composition in %P can evolve just as rapidly as can lateral plates (see Durston & El-Sabaawi, 2017). As a result, this contemporary evolution of elemental composition should then feedback to influence selection on stickleback armor and elemental composition, thus influencing ecological interactions (Leal *et al.*, 2017; individuals with different elemental demands compensate through consumer–resource interactions). These eco-evolutionary hypotheses seem a profitable avenue for future studies.

The among-population variation in %P was closely associated with predator regime, being higher in stickleback populations coexisting with sculpins. This variation could arise for two main reasons: (i) stickleback evolving with predators are more heavily armored and

therefore have greater %P or (ii) stickleback exposed to predators forage less, resulting in lower lipid stores and higher %P due to the lower body mass (Sterner & Elser, 2002). Both effects seem possible here given that (i) predation regime influences stickleback armor traits (Fig. 3c), which then influences %P (Fig. 5b); and (ii) fish condition influences %P (Par. $\eta^2 = 0.24$, $P < 0.001$) and predator presence can lead to lower foraging rates in stickleback (Milinski & Heller, 1978). Here, then, we have the potential for both genetic variation (armor adaptive divergence) and perhaps plasticity (decreased foraging) to jointly influence ecological effects, which has been suggested (Hendry, 2017), but not yet demonstrated. In addition, predator-driven selection on armor traits could lead to changes in environmental stoichiometry, which may then alter selection regimes further, thereby facilitating eco-evolutionary feedbacks (Pelletier *et al.*, 2009; Hendry, 2017).

Conclusions and implications

Our study provides empirical support for the expectation that temporal variation in environmental conditions can maintain high levels of variation in adaptive traits, even in traits that show differentiation associated with spatial variation in predator regime. In this system, breaching events cause each estuary to be periodically open to the ocean, which likely increases within-population variation through two mechanisms that prevent the fixation of adaptive alleles: (i) temporal variation in selection within populations (Kawecki & Ebert, 2004; Bell, 2010) and (ii) high gene flow between populations (Slatkin, 1973; Felsenstein, 1976; Kawecki, 2008). Although the latter effect is likely important, the former is too because gene flow alone is an insufficient explanation for the high within-population variation given only modest among-population variation. Our results thus support the hypothesis that temporal variation helps to maintain variation in adaptive traits within populations.

At the same time, temporal variation and high within-population variation did not eliminate population divergence in response to spatial variation in selection. Specifically, we documented associations between predator regime (presence or absence of sculpin) and stickleback armor traits (lateral plates and the gene that controls them, *Eda*). Although this differentiation is not as great as that found among other stickleback populations experiencing divergent selection regimes, it is notable for occurring in the face of frequent temporal fluctuations and relatively high gene flow. It therefore seems likely that selection occurring during the periods when estuaries are closed from the ocean and isolated from each other is sufficiently strong to drive some differentiation – even if it is later erased or reduced when the estuaries are open to the ocean. Certainly, some other studies have found that adaptive divergence can persist despite temporal variation in selection and high gene flow (Mojica *et al.*, 2012; Gotanda & Hendry, 2014); yet we argue that the divergence documented here is especially noteworthy given the extreme and rapid shifts in environmental conditions that these populations experience (Fig. 1b, Fig. S1).

It has long been debated whether selection in nature is typically 'strong' or 'weak' (Endler, 1986; Hoekstra *et al.*, 2001; Kingsolver *et al.*, 2001; Hereford *et al.*, 2004). What can be safely asserted is that selection should be stronger when environmental conditions change more rapidly (Chevin *et al.*, 2010; Michel *et al.*, 2014). Hence, we predict that these bar-built estuary stickleback population experience extremely strong selection at periodic intervals – and that this selection likely differs depending on temporal proximity to the breaching event. We suggest that selection is constantly driving contemporary evolution in these populations but that this nascent adaptation is frequently impeded or reversed by the rapidly changing conditions (i.e. fluctuating selection generating evolution in 'fits and starts'). These highly dynamic conditions should provide an excellent system for studies of contemporary evolution and its ecological consequences.

Acknowledgments

DD was funded by an NSERC graduate fellowship and the University of Victoria. RES was funded by an NSERC Discovery grant. EPP received support from the NOAA Cooperative Institute for Marine Ecosystems and Climate, University of California Institute for the Study of Ecological and Evolutionary Climate Impacts, and the US National Science Foundation (NSF-DEB #1457333; NSF-DEB #1556378). APH was supported by an NSERC Discovery Grant. RDHB received support from a Canada Research Chair, a Fonds de Recherche Nature et Technologies New Researcher grant and an NSERC Discovery grant. We thank the Younger Lagoon Natural Reserve, the Santa Cruz County Fish and Game Advisory Commission, and the Friends of Corcoran Lagoon Beach, and the Friends of the San Lorenzo River for funding. Special thanks to Gage Dayton, Elizabeth Howard, Kristin Kittleson, Jean Brocklebank and Georgina Monahan. We thank Katie Peichel, Yoel Stuart, Jun Kitano, Matthew Morris, Sean Rogers, Kerry Marchinko, Tom Reimchen, Andrew MaColl, Isabel Magalhaes, Daniele D'Agostino and Mike Bell for providing data on armor traits. Windsor Aguirre provided F_{ST} data from Alaskan populations. The Kingsley lab provided samples from Bodega Bay.

Contributions

AP, BAW, EPP, APH and RDHB designed the study. BAW, TMA and EPP collected samples. BAW took phenotypic measurements. AP, DH and SK performed the

molecular work. LA took geomorphometric measurements. DD and RWE took stoichiometric measurements. AP, BAW, DH, LA and DD analysed the data. AP wrote the manuscript with inputs from DH, RWE, EPP, EPH and RDHB.

References

Adams, D.C. & Otarolla-Castillo, E. 2013. geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**: 393–399.

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

Allen, L.G., Yoklavich, M.M., Cailliet, G.M. & Horn, M.H. 2006. Bays and Estuaries. In: *The Ecology of Marine Fishes California and Adjacent Waters* (L.G. Allen, D.J. Pondella & M.H. Horn, eds), pp. 119–148. University of California Press, Berkeley, CA.

Antao, T., Lopes, A., Lopes, R.J., Beja-Pereira, A. & Luikart, G. 2008. LOSITAN: a workbench to detect molecular adaptation based on a F_{st}-outlier method. *BMC Bioinformatics* **9**: 323.

Arnold, S.J., Pfrender, M.E. & Jones, A.G. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* **112–113**: 9–32.

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

Bartoń, K. 2016. MuMln: multi-model inference. R package.

Baumgartner, J.V. 1986. The genetics of differentiation in a stream population of the threespine stickleback, *Gasterosteus aculeatus*. *Heredity* **57**: 199–208.

Baumgartner, J.V. 1992. Spatial variation of morphology in a freshwater population of the threespine stickleback, *Gasterosteus aculeatus*. *Can. J. Zool.* **70**: 1140–1148.

Baumgartner, J.V. 1994. Phenotypic, genetic, and environmental integration of morphology in a stream population of the threespine stickleback, *Gasterosteus aculeatus*. *Can. J. Fish Aquat. Sci.* **52**: 1307–1317.

Baumgartner, J.V. & Bell, M.A. 1984. Lateral plate morph variation in California populations of the threespine stickleback, *Gasterosteus aculeatus*. *Evolution* **38**: 665–674.

Baythavong, B.S. 2011. Linking the spatial scale of environmental variation and the evolution of phenotypic plasticity: selection favors adaptive plasticity in fine-grained environments. *Am. Nat.* **178**: 75–87.

Beaumont, M.A. & Nichols, R.A. 1996. Evaluating loci for use in the genetic analysis of population structure. *Proc. Biol. Sci.* **263**: 1619–1626.

Becker, G.S. & Reining, I.J. 2008. *Steelhead/Rainbow Trout (Oncorhynchus mykiss)*. Resources South of the Golden Gate, California. Center for Ecosystem Restoration and Management, Oakland, CA.

Behrens, D.K. & Bombardelli, F.A. 2009. Characterization of time and spatial scales of a migrating rivermouth. *Geophys. Res. Lett.* **36**: l–4.

Behrens, D.K., Bombardelli, F.A., Largier, J.L. & Twohy, E. 2013. Episodic closure of the tidal inlet at the mouth of the Russian River – a small bar-built estuary in California. *Geomorphology* **189**: 66–80.

Bell, M.A. 1976. Evolution of phenotypic diversity in *Gasterosteus aculeatus* superspecies on the Pacific coast of North America. *Syst. Zool.* **25**: 211–227.

Bell, M.A. 1979a. Low-plate morph of the threespine stickleback breeding in salt water. *Copeia* **1979**: 529.

Bell, M.A. 1979b. Persistence of ancestral-sister species. *Syst. Zool.* **28**: 85.

Bell, M.A. 1981. Lateral plate polymorphism and ontogeny of the complete plate morph of threespine sticklebacks (*Gasterosteus aculeatus*). *Evolution* **35**: 67–74.

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

Bell, G. 2010. Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **365**: 87–97.

Bell, M.A. & Orti, G. 1994. Pelvic reduction in threespine stickleback from Cook Inlet lakes: geographical distribution and intrapopulation variation. *Copeia* **1994**: 314–325.

Bell, M.A. & Richkind, K.E. 2015. Clinal variation of lateral plates in threespine stickleback fish. *Am. Nat.* **117**: 113–132.

Bradley, W.C. & Griggs, G.B. 1976. Form, genesis, and deformation of central California wave-cut platforms. *Geol. Soc. Am.* **87**: 433–449.

Chan, Y.F., Marks, M.E., Jones, F.L.C., Villarreal, G., Shapiro, M.D., Brady, S.D. *et al.* 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**: 302–305.

Chevin, L.M. 2013. Genetic constraints on adaptation to a changing environment. *Evolution* **67**: 708–721.

Chevin, L.M. & Lande, R. 2010. When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated population? *Evolution* **64**: 1143–1150.

Chevin, L.M., Lande, R. & Mace, G.M. 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biol.* **8**: 1–8.

Chevin, L.M., Visser, M.E. & Tufto, J. 2015. Estimating the variation, autocorrelation, and environmental sensitivity of phenotypic selection. *Evolution* **69**: 2319–2332.

Childs, D.Z., Metcalf, C.J.E. & Rees, M. 2010. Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. *Proc. Biol. Sci.* **277**: 3055–3064.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. *et al.* 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol.* **2**: e109.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal Jr, G., Dickson, M., Grimwood, J. *et al.* 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**: 1928–1933.

Coulson, T. & Tuljapurkar, S. 2008. The dynamics of a quantitative trait in an age-structured population living in a variable environment. *Am. Nat.* **172**: 599–612.

Durston, D.J. & El-Sabaawi, R.W. 2017. Bony traits and genetics drive intraspecific variation in vertebrate elemental composition. *Funct. Ecol.* **31**: 2128–2137.

Edwards, A.W. 1971. Distances between populations on the basis of gene frequencies. *Biometrics* **27**: 873–881.

Ellner, S. & Hairston Jr, N.G. 1994. Role of overlapping generations in maintaining genetic-variation in a fluctuating environment. *Am. Nat.* **143**: 403–417.

El-Sabaawi, R.W., Zandona, E. & Kohler, T.J. 2012. Widespread intraspecific organismal stoichiometry among populations of the Trinidadian guppy. *Funct. Ecol.* **26**: 666–676.

El-Sabaawi, R.W., Warbanski, M.L., Rudman, S.M. & Hovel, R. 2016. Investment in boney defensive traits alters organismal stoichiometry and excretion in fish. *Oecologia* **181**: 1209–1220.

Elwany, M., Flick, R.E. & Aijaz, S. 1998. Opening and closure of a marginal southern California lagoon inlet. *Estuaries* **21**: 246–254.

Endler, J.A. 1986. *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ.

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**: 2611–2620.

Excoffier, L., Smouse, P.E. & Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Felsenstein, J. 1976. The theoretical population genetics of variable selection and migration. *Annu. Rev. Genet.* **10**: 253–280.

Frechette, D.M., Satterthwaite, W.H., Osterback, A.-M.K. & Hayes, S.A.S.A. 2016. Steelhead abundance in seasonally closed estuaries estimated using mark recapture methods. *NOAA Technical Memorandum NMFS* 1–30.

Froese, R. 2006. Cube law, condition factor and weight–length relationships: history, meta-analysis and recommendations. *J. Appl. Ichthyol.* **22**: 241–253.

Goodall, C. 1991. Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc. B* **53**: 285–339.

Gotanda, K.M. & Hendry, A.P. 2014. Using adaptive traits to consider potential consequences of temporal variation in selection: male guppy colour through time and space. *Biol. J. Linn. Soc.* **112**: 108–122.

Goudet, J., Raymond, M., de Meeüs, T. & Rousset, F. 1996. Testing differentiation in diploid populations. *Genetics* **144**: 1933–1940.

Grant, P.R. & Grant, B.R. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**: 707–711.

Grueber, C.E., Nakagawa, S., Laws, R.J. & Jamieson, I.G. 2011. Multimodel inference in ecology and evolution: challenges and solutions. *J. Evol. Biol.* **24**: 699–711.

Hairston, N.G. & Dillon, T.A. 1990. Fluctuating selection and response in a population of freshwater Copepods. *Evolution* **44**: 1796–1805.

Hedrick, P.W. 1986. Genetic polymorphism in heterogeneous environments: a decade later. *Annu. Rev. Ecol. Syst.* **17**: 535–566.

Hendry, H. 2017. *Eco-Evolutionary Dynamics*. Princeton University Press, Princeton, NJ.

Hendry, A.P. & Taylor, E.B. 2004. How much of the variation in adaptive divergence can be explained by gene flow? An evaluation using lake-stream stickleback pairs. *Evolution* **58**: 2319–2331.

Hendry, A.P., Hendry, A.S. & Hendry, C.A. 2013. Hendry Vineyard stickleback: testing for contemporary lake–stream divergence. *Evol. Ecol. Res.* **15**: 343–359.

Hereford, J., Hansen, T.F. & Houle, D. 2004. Comparing strengths of directional selection: how strong is strong? *Evolution* **58**: 2133–2143.

Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hoang, A., Hill, C.E. *et al.* 2001. Strength and tempo of directional selection in the wild. *Proc. Natl. Acad. Sci. USA* **98**: 9157–9160.

Hoogland, R., Morris, D. & Tinbergen, N. 1956. The spines of sticklebacks (*Gasterosteus* and *Pygosteus*) as means of defence against predators (*Perca* and *Esox*). *Behaviour* **10**: 205–236.

Hunt, G., Bell, M.A. & Travis, M.P. 2008. Evolution toward a new adaptive optimum: phenotypic evolution in a fossil stickleback lineage. *Evolution* **62**: 700–710.

Hyatt, K.D. 1979. Feeding strategy. In: *Fish Physiology Vol. 8: Bioenergetics and Growth* (W.S. Hoar, D.J. Randall & J.R. Brett, eds), pp. 71–113. Academic Press, New York, NY.

Ingram, T., Svanbäck, R., Kraft, N.J.B., Kratina, P., Southcott, L. & Schluter, D. 2012. Intraguild predation drives evolutionary niche shift in threespine stickleback. *Evolution* **66**: 1819–1832.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

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

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

Kaeuffer, R., Peichel, C.L., Bolnick, D.I. & Hendry, A.P. 2012. Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution* **66**: 402–418.

Kamvar, Z.N., Kamvar, Z.N., Tabima, J.F., Tabima, J.F., Grünwald, N.J. & Grünwald, N.J. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**: e281.

Kawecki, T.J. 2008. Adaptation to marginal habitats. *Annu. Rev. Ecol. Evol. Syst.* **39**: 321–342.

Kawecki, T.J. & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecol. Lett.* **7**: 1225–1241.

Kellermann, V.M., van Heerwaarden, B., Hoffmann, A.A. & Sgrò, C.M. 2006. Very low additive genetic variance and evolutionary potential in multiple populations of two rainforest *Drosophila* species. *Evolution* **60**: 1104–1106.

Kellermann, V., van Heerwaarden, B., Sgro, C.M. & Hoffmann, A.A. 2009. Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. *Science* **325**: 1244–1246.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E. *et al.* 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* **157**: 245–261.

Kirkpatrick, M. & Barton, N.H. 1997. Evolution of a species' range. *Am. Nat.* **150**: 1–23.

Kitano, J., Bolnick, D.I., Beauchamp, D.A., Mazur, M.M., Mori, S., Nakano, T. *et al.* 2008. Reverse evolution of armor plates in the threespine stickleback. *Curr. Biol.* **18**: 769–774.

Lacasse, J. & Aubin-Horth, N. 2012. A test of the coupling of predator defense morphology and behavior variation in two threespine stickleback populations. *Curr. Zool.* **58**: 53–65.

Lande, R. & Shannon, S. 1996. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* **50**: 434–437.

Leal, M.C., Best, R.J., Durston, D., El-Sabaawi, R.W. & Matthews, B. 2017. Stoichiometric traits of stickleback: effects of genetic background, rearing environment, and ontogeny. *Ecol. Evol.* **7**: 2617–2625.

Leinonen, T., Cano, J.M., Mäkinen, H. & Merilä, J. 2006. Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *J. Evol. Biol.* **19**: 1803–1812.

Leinonen, T., Herczeg, G., Cano, J.M. & Merilä, J. 2011. Predation-imposed selection on threespine stickleback (*Gasterosteus aculeatus*) morphology: a test of the refuge use hypothesis. *Evolution* **65**: 2916–2926.

Lescak, E.A. & von Hippel, F.A. 2011. Selective predation of threespine stickleback by rainbow trout. *Ecol. Freshw. Fish* **20**: 308–314.

Lind, M.I. & Johansson, F. 2007. The degree of adaptive phenotypic plasticity is correlated with the spatial environmental heterogeneity experienced by island populations of *Rana temporaria*. *J. Evol. Biol.* **20**: 1288–1297.

Lind, M.I., Ingvarsson, P.K., Johansson, H., Hall, D. & Johansson, F. 2010. Gene flow and selection on phenotypic plasticity in an island system of *Rana temporaria*. *Evolution* **65**: 684–697.

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

Mackay, T.F.C. 1981. Genetic variation in varying environments. *Genet. Res.* **37**: 79–93.

Mäkinen, H.S., Shikano, T., Cano, J.M. & Merilä, J. 2008. Hitchhiking mapping reveals a candidate genomic region for natural selection in three-spined stickleback chromosome VIII. *Genetics* **178**: 453–465.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**: 209–220.

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

Michel, M.J., Chevin, L.M. & Knouft, J.H. 2014. Evolution of phenotype-environment associations by genetic responses to selection and phenotypic plasticity in a temporally autocorrelated environment. *Evolution* **68**: 1374–1384.

Milinski, M. & Heller, R. 1978. Influence of a predator on the optimal foraging behaviour of sticklebacks (*Gasterosteus aculeatus* L.). *Nature* **275**: 642–644.

Miller, S.E., Metcalf, D. & Schluter, D. 2015. Intraguild predation leads to genetically based character shifts in the threespine stickleback. *Evolution* **69**: 3194–3203.

Mohamoud, Y.M. & Parmar, R.S. 2006. Estimating streamflow and associated hydraulic geometry, the mid-atlantic region, USA. *J. Am. Water Resour. Assoc.* **42**: 755–768.

Mojica, J.P., Lee, Y.W., Willis, J.H. & Kelly, J.K. 2012. Spatially and temporally varying selection on intrapopulation quantitative trait loci for a life history trade-off in *Mimulus guttatus*. *Mol. Ecol.* **21**: 3718–3728.

Moodie, G.E.E. 1972. Predation, natural selection and adaptation in an unusual threespine stickleback. *Heredity* **28**: 155–167.

Morrissey, M.B. & Hadfield, J.D. 2012. Directional selection in temporally replicated studies is remarkably consistent. *Evolution* **66**: 435–442.

Moyle, P.B. 1976. *Inland Fishes of California*. University of California Press, Berkeley, CA.

Murphy, J. & Riley, I.P. 1962. A modified single solution method for the determination of phosphate in natural waters. *Anal. Chim. Acta* **27**: 31–36.

Mustonen, V. & Lässig, M. 2007. Adaptations to fluctuating selection in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **104**: 2277–2282.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **70**: 3321–3323.

Peichel, C.L., Nereng, K.S., Ohgi, K.A., Cole, B.L., Colosimo, P.F., Buerkle, C.A. *et al.* 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* **414**: 901–905.

Pelletier, F., Garant, D. & Hendry, A.P. 2009. Eco-evolutionary dynamics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**: 1483–1489.

Pinheiro, J., Bates, D., Debroy, S. & Sarkar, D. 2016. nlme: Linear and nonlinear mixed effects models.

Pressley, P.H. 1981. Parental effort and the evolution of nest-guarding tactics in the threespine stickleback, *Gasterosteus aculeatus* L. *Evolution* **35**: 282–295.

Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

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

Reimchen, T.E. 1980. Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus*: an adaptation to predators? *Can. J. Zool.* **58**: 1232–1244.

Reimchen, T.E. 1992. Injuries on stickleback from attacks by a toothed predator (*Oncorhynchus*) and implications for the evolution of lateral plates. *Evolution* **46**: 1224.

Reimchen, T.E. 1994. Predators and morphological evolution in threespine stickleback. In: *The Evolutionary Biology of the Threespine Stickleback* (M. Bell & S.A. Foster, eds), pp. 240–276. Oxford University Press, Oxford.

Reimchen, T.E. 1995. Predator-induced cyclical changes in lateral plate frequencies of *Gasterosteus*. *Behaviour* **132**: 1079–1094.

Reimchen, T.E. & Nosil, P. 2002. Temporal variation in divergent selection on spine number in threespine stickleback. *Evolution* **56**: 2472–2483.

Reimchen, T.E., Bergstrom, C. & Nosil, P. 2013. Natural selection and the adaptive radiation of Haida Gwaii stickleback. *Evol. Ecol. Res.* **15**: 241–269.

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

Rich, A. & Keller, E.A. 2013. A hydrologic and geomorphic model of estuary breaching and closure. *Geomorphology* **191**: 64–74.

Rohlf, F.J. 2006. tpsDig version 2.10.

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

Rousset, F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**: 103–106.

Saether, B.-E. & Engen, S. 2015. The concept of fitness in fluctuating environments. *Trends Ecol. Evol.* **30**: 273–281.

Sasaki, A. & Ellner, S. 1997. Quantitative genetic variance maintained by fluctuating selection with overlapping generations: variance components and covariances. *Evolution* **51**: 682–696.

Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.

Shapiro, M.D., Marks, M.E., Peichel, C.L. & Blackman, B.K. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**: 717–723.
Siepielski, A.M., DiBattista, J.D. & Carlson, S.M. 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecol. Lett.* **12**: 1261–1276.
Siepielski, A.M., Morrissey, M.B., Buoro, M., Carlson, S.M., Caruso, C.M., Clegg, S.M. *et al.* 2017. Precipitation drives global variation in natural selection. *Science* **355**: 959–962.
Simons, A.M. 2009. Fluctuating natural selection accounts for the evolution of diversification bet hedging. *Proc. Biol. Sci.* **276**: 1987–1992.
Slatkin, M. 1973. Gene flow and selection in a cline. *Genetics* **75**: 733–756.
Sletvold, N. & Grindeland, J.M. 2007. Fluctuating selection on reproductive timing in *Digitalis purpurea*. *Oikos* **116**: 473–481.
Spichtig, M. & Kawecki, T.J. 2004. The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. *Am. Nat.* **164**: 70–84.
Sterner, W.S. & Elser, J.J. 2002. *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere*. Princeton University Press, Princeton, NJ.
Stuart, Y.E., Veen, T., Weber, J.N., Hanson, D., Ravinet, M., Lohman, B.K. *et al.* 2017. Contrasting effects of environment and genetics generate a continuum of parallel evolution. *Nat. Ecol. Evol.* **1**: 158.
Vines, T.H., Dalziel, A.C., Albert, A.Y.K., Veen, T., Schulte, P.M. & Schluter, D. 2016. Cline coupling and uncoupling in a stickleback hybrid zone. *Evolution* **70**: 1023–1038.
Webster, M. & Sheets, D.H. 2010. A practical introduction to landmark-based geometric morphometrics. In: *Quantitative Methods in Paleobiology* (J. Alroy & G. Hunt, eds), pp. 163–168. The Paleontological Society Papers, Cambridge.
Whitlock, M.C. & Schluter, D. 2009. *The Analysis of Biological Data*. Roberts and Company Publishers, Greenwood Village, CO.
Zeller, M., Lucek, K. & Haesler, M.P. 2012. Signals of predation-induced directional and disruptive selection in the threespine stickleback. *Evol. Ecol. Res.* **14**: 193–205.
Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A. & Smith, G.M. 2009. *Mixed Effects Models and Extensions in Ecology with R*. Springer-Verlag, New York, NY.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:
Table S1 Population sample size used in each analysis.
Table S2 Pairwise genetic differentiation between estuaries represented by F_{ST} values.
Table S3 Summary statistics of AMOVA test.
Table S4 Results of permutation tests performed on the AMOVA test statistics.
Table S5 For each element of armor morphology: standard deviation (SD) of the random intercepts for estuary for the mixed model, ΔAIC for the comparison of a model without a random effect to the full model, log-likelihood-ratio (LLR), and *P*-value of a likelihood ratio test comparing the full model with a model lacking the random effect of population for each dependent variable.
Table S6 Best models based on AICc for %P. Marginal R^2 for model fit of main effect is .35 (phenotype model) and .43 (phenotype + *Eda* model).
Table S7 Within-population standard deviations from the estuary populations and from known environmentally stable populations.
Table S8 Mean lateral plate counts (Plate) and standard deviations (SD) from six different systems.
Figure S1 Photos and videos of breaching events.
Figure S2 Location of the 18 landmarks for geometric morphometrics.
Figure S3 Representation of correspondence analysis (CA) performed on allele frequencies on axes 1 and 2.
Figure S4 Histograms representing the simulated values of the randomized values of the AMOVA test. The black line represents the observed values. A. Variation within samples, B. Variation between samples, C. Variation between populations.
Figure S5 Hierarchical clustering tree based on a cluster analysis constructed with Edward's genetic distance. The axis is a measure of closeness of clusters (Distance).
Figure S6 Isolation by distance in Santa Cruz estuary populations. Scatter plot of pairwise genetic distances (Edward's) against pairwise geographic distances (km) in our 23 sampling points.
Figure S7 Individual assignment to population structure inferred by STRUCTURE using only fish homozygote 'complete' at *Eda*. Each bar represents an individual. The Y axis represents the probability of classification to a cluster.
Figure S8 Body shape variation on PC1 (top) and PC2 (bottom) of Procrustes residuals.
Figure S9 Relationships between phenotypic traits and %P from the phenotypic GLMM. %P rose significantly ($P < 0.001$) with standard length (Panel A; SL), head length (C) and lateral plate count (D), while declining with condition (B). Shaded regions depict 95% confidence ranges.
Figure S10 Proportion of variance explained (η^2) for the studied traits. Between-population variance is in grey and within-population variance in black.

Data deposited at Dryad: https://doi.org/10.5061/dryad.7h4s265

Received 17 July 2017; revised 23 January 2018; accepted 3 March 2018