Plasticity and heritability of morphological variation within and between parapatric stickleback demes

R. J. S. McCAIRNS*† & L. BERNATCHEZ†

*Ecological Genetics Research Unit (EGRU), Department of Biosciences, University of Helsinki, Helsinki, Finland
†Québec Océan, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

Abstract

The threespine stickleback (*Gasterosteus aculeatus*) has emerged as an important model organism in evolutionary ecology, largely due to the repeated, parallel evolution of divergent morphotypes found in populations having colonized freshwater habitats. However, morphological divergence following colonization is not a universal phenomenon. We explore this in a large-scale estuarine ecosystem inhabited by two parapatric stickleback demes, each physiologically adapted to divergent osmoregulatory environments (fresh vs. saline waters). Using geometric morphometric analyses of wild-caught individuals, we detected significant differences between demes, in addition to sexual dimorphism, in body shape. However, rearing full-sib families from each deme under controlled, reciprocal salinity conditions revealed no differences between genotypes and highly significant environmental effects. It is also noteworthy that fish from both demes were fully plated, whether found in the wild or reared under reciprocal salinity conditions. Although we found significant heritability for body shape, we also noted significant direct environmental effects for many latent shape variables. Moreover, we found little evidence for diversifying selection acting on body size and shape (Q_{ST}). Nevertheless, uniform compressive variation did exceed neutral expectations, yet despite evidence of both allometry and genetic correlation with body length, we detected no correlated signatures of selection. Taken together, these results suggest that much of the morphological divergence observed in this system is the result of plastic responses to environmental variation rather than adaptive differentiation.

Introduction

The study of adaptive phenotypic variation has contributed significantly to our understanding of contemporary evolution and the rate at which natural selection can act. One area of active interest is the quantification of morphological variation across heterogeneous environments. Although studies of wild populations in relation to their natural ecology can be invaluable in understanding the role of selection in promoting biological diversity (Endler, 1986; Feder & Mitchell-Olds, 2003), interpreting patterns of variation in the wild is a

Correspondence: Scott McCairns, Ecological Genetics Research Unit (EGRU), Department of Biosciences, University of Helsinki, Helsinki 00014, Finland.
Tel.: +358 9 1915 7801; fax: +358 9 1915 7694;
e-mail: scott.mccairns@helsinki.fi

challenge fraught with potential misinterpretations, not least of which is the confounding of evolutionarily neutral differences for divergence induced by natural selection. In this regard, environmental variation itself may be particularly problematic. Covariation between environmental and phenotypic gradients may reflect changing selective pressures, but local adaptation cannot be assumed: such patterns may also belie effectively neutral variation resulting from direct environmental inputs. Expression of all but the most canalized traits will be subject to some degree of direct influence due to myriad environmental effects, and environmental inputs/perturbations during developmental can have pronounced effects on phenotypic variation, often with little affect on fitness (Ghalambor *et al.*, 2007; Georga & Koumoundouros, 2010). This is not to suggest that plasticity cannot be advantageous; indeed, it is well

established that plasticity can be an adaptive response to environmental variation (Gotthard & Nylin, 1995; Ghalambor *et al.*, 2007; Beldade *et al.*, 2011). Moreover, developmental plasticity in particular has been hypothesized to be a significant source of evolutionary novelty (West-Eberhard, 2005; Moczek *et al.*, 2011). Thus, dissecting phenotypic variation into its genetic and plastic components represents an endeavour not only promoting a greater understanding of the proximate mechanisms underlying morphological diversity, but one with potentially broader implications for evolutionary theory. To this end, model systems that are amenable to both experimental and observational approaches may be best-positioned to yield the greatest insights. Such common garden experiments have been used to demonstrate the interaction between genetic and environmental effects on morphology (Marcil *et al.*, 2006; Parsons & Robinson, 2007). These studies are of considerable interest in that they reveal the plastic nature of the phenotype; however, the genetic component often remains implicit. Yet models in which formalized genetic analyses can be applied may be capable of producing even greater insights, permitting a more quantitative comparison of the relative contributions of environmental and genetic sources of variance. Moreover, without evidence for heritable variation underlying any focal trait, subsequent inference regarding its adaptive potential will remain suspect.

Heritability of form has been estimated based on a number of complex shape descriptors ranging from linear combinations of shape variables (Baumgartner, 1995), distances in multivariate shape space (Monteiro *et al.*, 2002), scale-independent landmark coordinate vectors (Klingenberg & Leamy, 2001) and even integrative Fourier shape outlines (Currie *et al.*, 2000). Moreover, analyses of the genetic architecture of shape have identified quantitative trait loci (QTL) for multivariate shape descriptors (Zimmerman *et al.*, 2000; Klingenberg *et al.*, 2001; Workman *et al.*, 2002; Albert *et al.*, 2008). However, although geometric morphometrics can be a powerful and efficient means of describing variation in shape, not all resultant latent variables may be appropriate for genetic analysis. This was highlighted by Berner *et al.* (2011), who suggest that relative warp scores are likely to yield problematic estimates arising through artificial covariance induced by PCA during the calculation of relative warps. Thus, although the central challenge towards a full understanding of whole-body-shape determinism lies in disentangling the relative role of all salient sources of variation, there may be a fundamental conflict between the most biologically/functionally meaningful representation of that variance and its mathematical tractability. The solution to this problem is not obvious, but the need to assess 'shape' within an evolutionary framework still remains. One practical 'work-around' would be to base estimates on partial warp scores – the analytical precursors of relative warps – as partial warp scores are amenable to analysis via traditional statistical tests (Zelditch *et al.*, 2004). We reason that this should also extend to quantitative genetic analyses, which are essentially an exercise in linear mixed-effects modelling (Kruuk, 2004).

The threespine stickleback (*Gasterosteus aculeatus*) has emerged as a premiere model in the study of evolutionary ecology, largely due to the species' multiple, parallel examples of adaptive morphological divergence (Bell & Foster, 1994; Schluter, 1996; McKinnon & Rundle, 2002). Marine populations are ancestral, and this body form has been highly conserved since the Miocene (Bell, 1994; Walker & Bell, 2000); however, colonization of freshwater habitats has resulted in myriad changes in body morphology (Reid & Peichel, 2010). Morphological divergence in freshwater has been shown to have a genetic basis (Schluter *et al.*, 2004; Berner *et al.*, 2011), although the degree of differentiation from the ancestral form appears to be dependent upon how different novel environments are from the marine milieu (Spoljaric & Reimchen, 2007). Numerous studies have used geometric morphometrics to capture the divergence among populations occupying unique habitats and trophic niches (Baumgartner, 1995; Spoljaric & Reimchen, 2007; Sharpe *et al.*, 2008; Aguirre, 2009). Many have focused on variation in specific anatomical structures, from which adaptive significance is argued on functional grounds (Caldecutt & Adams, 1998; Kimmel *et al.*, 2008; Arif *et al.*, 2009), whereas others have associated whole-body-shape variation with adaptive functionality underlying differentiation (Walker, 1997; Sharpe *et al.*, 2008; Hendry *et al.*, 2011). However, striking morphological divergence between habitats is not ubiquitous (Berner *et al.*, 2008; Kaeuffer *et al.*, 2012), and exploring more subtle differences in form may be equally illuminating regarding the conditions favouring and/or impeding differentiation in body shape.

The St Lawrence River estuary represents an ideal system in which to explore the genetic and environmental components of phenotypic variation. One of the principal environmental features of the estuary is its gradation into three hydrological zones, each characterized by unique biological, physiochemical and tidal properties (Vincent & Dodson, 1999). The fluvial estuary (aka upper estuary), although tidal in nature, is a uniquely freshwater zone extending upstream approximately 160 km from the eastern end of Île d'Orléans. The middle estuary, located between the eastern tip of Île d'Orléans and the Saguenay Fjord, is characterized by significant current reversals and strong mixing associated with the diurnal tidal cycles. Consequently, this is a highly turbid and biologically productive section, with salinity ranging from 0.5 to 25 practical salinity units (psu). The maritime estuary (aka lower estuary) is a 230-km stretch ultimately discharging into the Gulf of St Lawrence. Hydraulic dynamics shift from tide-dominated to wave-dominated, and the biological and

physiochemical properties more closely resemble those of the marine environment. Extant stickleback are partitioned into two demes whose geographical ranges correspond to the freshwater/saltwater division of the estuary (McCairns & Bernatchez, 2008). Genetic differentiation is weak ($F_{ST} = 0.006$; $P < 0.001$) yet temporally stable, and ecological factors independent of geographical distance, particularly salinity, explain the greatest proportion of genetic variance. Moreover, these salinity differences represent unique selective pressures that appear to be driving a nascent adaptive, physiological divergence between demes (McCairns & Bernatchez, 2010). Yet, in contrast to the diversity of forms seen in populations inhabiting similar environmental gradients, stickleback in the St Lawrence exhibit no strikingly obvious morphological differences across the range of environmental conditions encountered. Thus, any morphological divergence between demes – whether plastic or adaptive – must be subtle and as such will require more sophisticated analytical means for its detection.

In this study, we quantify shape variation in sticklebacks originating from each deme inhabiting the St Lawrence estuary. Additionally, we compare differences in the number of lateral plates, an important and well-studied meristic trait frequently associated with freshwater adaptation (Colosimo *et al.*, 2005; Barrett *et al.*, 2008). We use geometric morphometric analyses to define latent variables for shape in samples of mature stickleback captured on their spawning grounds and for laboratory crosses from these same demes reared under controlled environmental conditions. Using artificial crosses of known pedigree raised under reciprocal environmental conditions, we are able to quantify both the additive genetic and environmental components of phenotypic variation in shape. Consequently, we are able to better interpret the evolutionary significance and potential of shape differences observed in the wild.

Materials and methods

We sampled mature sticklebacks from spawning sites using a combination of seine nets, dip nets and minnow traps. Eleven sites were sampled (Fig. 1), each belonging to one of two adaptively divergent demes endemic to the St Lawrence River estuary (McCairns & Bernatchez, 2008, 2010). Each sample contained a random selection of 25 males and 25 females from each site.

Common garden experiment

Broodstock were obtained from two sites (CR and FOR), each belonging to one of the two aforementioned demes. Mature ova were stripped from females *in situ* and transported to wet laboratory facilities, on ice, in sterile Holtfreter's solution. Testes were also dissected *in situ* and transported in sterile Ginzburg's Fish Ringers solution.
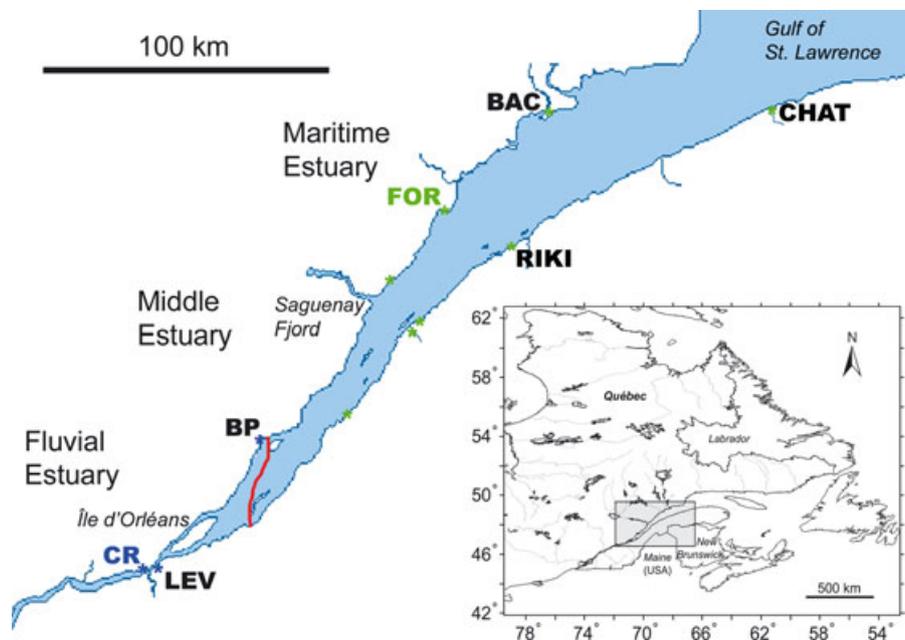
Fig. 1 Location of samples from stickleback spawning sites in the St Lawrence estuary. Sites containing individuals from the maritime deme are indicated with mid-grey asterisks (green online), whereas freshwater sites are in dark grey (blue online). Sites mentioned specifically in the text are labelled on the map, including sources of broodstock for the common garden experiment (CR and FOR). The dark grey line (red online) corresponds to the approximate location of the 80% isocline delimiting the extent of deme ranges (McCairns & Bernatchez, 2008).

First generation families (F_1) were produced following modifications of zebrafish *in vitro* fertilization techniques adapted for stickleback research (University of Oregon Stickleback Research Site, 2008). Crosses were established following a blocked factorial breeding design in which each female was mated with two males, one originating from her deme of origin (purebred crosses) and the other from the reciprocal group (hybrid crosses) – a more detailed description of the breeding design can be found in the methods and supplementary material of the study of McCairns & Bernatchez (2010). A testis from each male was divided into half, and each half macerated in a separate 100-mm-diameter Petri dish. Ova were initially divided into four lots of approximately 50 eggs, and each lot mixed gently with one of the macerated testes and 5‰ embryo medium. Fertilized eggs were incubated at 16 °C, with media changed twice daily.

Experimental conditions consisted of all full-sib families divided into two groups raised under salinity regimes representative of the natal freshwater (< 1‰) and maritime (20‰) environments (McCairns & Bernatchez, 2008). Upon hatch, fry from 5‰ salinity embryo medium were gradually acclimated to alternative salinity conditions. In one group, salinity was increased by 5‰ per day to final experimental conditions (20‰). The second group was immediately transferred to freshwater medium (< 1‰). Upon absorption of yolk sac and beginning of exogenous feeding, each full-sib family was transferred to an individual 2-L container. Fry were fed *ad libitum* twice daily with freshly hatched *Artemia* nauplii. At 45 days post-hatch, all larval fish were photographed for length measurement. Eight factorial blocks, comprising 32 full-sib families with individuals in both salinity treatments, were selected for transfer to experimental aquaria. Experimental tanks consisted of individual aquaria connected to one of two 1600-L recirculating systems, one maintained at 20‰ salinity and the other with dechlorinated tap water (< 1‰ salinity). Water quality was maintained with a biofiltration system and through daily siphoning of waste materials in individual aquaria. Fish were fed *ad libitum* twice daily a mixture of flake food and commercial salmonid fry ration, in addition to once-daily supplements of freeze-dried *Mysis relicta*, frozen chironomid larvae and live *Artemia* nauplii. Families were photographed an additional three times throughout the experiment (120, 180 and 230 days post-hatch), from which size-at-age data were extracted. Experiments continued until 230 days post-hatch, at which time surviving individuals were killed.

Sample preparations

The sex of wild-caught individuals was verified by carefully opening the right lateral side and removing the gonads. Females were stripped of eggs prior to preparation to ensure that morphological comparisons with males were not affected by abdominal inflation due to gravidity. Samples were prepared following Potthoff's (1984) protocols, which included the following: fixation in 5% buffered formalin, bleaching in a 0.5% hydrogen peroxide solution, clearing with trypsin and staining with alizarin red dissolved in potassium hydroxide. Staining facilitated an accurate count of total lateral plate number (left side, including keel plates) and identification of subsequent corporal landmarks. Lateral plate counts were performed for all wild-caught samples ($n = 550$) and all laboratory crosses ($n = 615$). Additionally, we measured standard length (SL) for all these individuals.

A subset of wild-caught sticklebacks and all laboratory-reared individuals aged > 150 days post-hatch were also photographed for geometric morphometric analyses. Four sites from both banks of the river were selected, two corresponding to the freshwater deme (CR and LEV) and two representing the maritime deme (BAC and RIKI). We also selected a site from each deme's range, which represented atypical environmental conditions. These included a site occupied by the freshwater deme, but with pronounced diurnal salinity fluctuations due to tidal processes (BP), and a largely freshwater site (3‰ mean salinity) inhabited by maritime individuals (CHAT): this site was an oxbow of a large tributary near its confluence with the St Lawrence that experienced only infrequent tidal flooding. All samples contained 25 males and 25 females, yielding a total sample size of 300 wild-caught individuals. Purebred F_1 crosses were approximately evenly distributed by cross type (FW–FW and SW–SW) and salinity treatment (freshwater: < 1‰ and saltwater: 20‰), yielding a total sample size of 199 individuals for comparative analyses. Samples for subsequent quantitative genetic analyses also included hybrid families, yielding a total sample size of 364 individuals.

Individuals were pinned in place against a Styrofoam form, providing a level view of the left lateral side. Pins were also placed flush against the body at 14 corporal landmarks, previously identified by Walker (1997). Walker's 15th landmark, located at the posterior edge of the angular, however, was removed from this data set due to inconsistent placement among specimens. All specimens were photographed next to a common scale and at a set focal distance. Landmark coordinates were obtained from digital photographs using the image analysis software IMAGEJ (Abramoff *et al.*, 2004).

Defining latent variables of shape

The most commonly used geometric morphometric analyses involve the translation of homologous corporal locations (landmarks) onto a Cartesian coordinate system, followed by multivariate scaling and rotation of similarly measured individuals onto a common plane or sphere (Bookstein, 1989; Goodall, 1991). One of the major advantages of this approach over traditional truss-based measurements is that the projections represent a

comparison of shapes free of simple isometric size differences. Among biologists, the thin-plate spine (TPS) has gained considerable popularity, largely due to its utility in providing easily interpretable visual representations of shape differences (Parsons *et al.*, 2003; Zelditch *et al.*, 2004). The TPS is a function that can be used to interpolate change between landmarks when a given shape is deformed onto a common or reference shape space. Analytically, this is achieved through eigenanalysis of the matrix of bending energies required to perform such a deformation, yielding vectors of coefficients, known as partial warps, which can be used to calculate individualized scores effectively summarizing variation in form (see Chapter 6 in the study of Zelditch *et al.*, 2004 for a complete derivation). Although partial warp scores themselves have no specific functional/biological meaning, they provide a means of capturing variation in shape. Moreover, partial warp scores can be used in statistical tests without adjusting the degrees of freedom (Zelditch *et al.*, 2004), thus lending themselves to a myriad of analytical techniques capable of detecting extremely subtle differences in form between groups, as well as traditional partitioning of variance components.

Geometric morphometrics were performed using the package 'shapes' (Dryden, 2007), and R functions developed by Claude (2008). Intuitive, visual interpretation of shape variation was facilitated by thin-plate spline (TPS) deformation grids and by Procrustes superimposition of landmark data (Bookstein, 1989; Zelditch *et al.*, 2004). We defined differences between Procrustes superimpositions as the Riemannian (ρ) shape distance (Dryden *et al.*, 2009) and statistically evaluated these differences employing an unbiased test statistic (λ_{min}) and pivotal bootstrap procedure, based on 1000 iterations (Amaral *et al.*, 2007). General shape variation was determined both for affine/uniform deformations requiring zero bending energy (i.e. shearing and compression, the so-called zeroth partial warps) and for nonuniform deformations to a unified shape space (Zelditch *et al.*, 2004). Landmark coordinates from both wild-caught and laboratory-reared fish were analysed simultaneously, as a single data set, to facilitate comparison and interpretation of results. Individual partial warp scores corresponding to each eigenvector capable of describing > 5% of shape variation (i.e. uniform deformation vectors and the first five partial warps) were used as latent variables describing overall shape in subsequent analyses.

Data analysis

All analyses were performed in the R computing language (R Development Core Team, 2007). For all fish sampled from the wild ($n = 550$) and for all purebred laboratory individuals ($n = 615$), lateral plates were analysed via generalized mixed-effects models implemented in the 'MCMCglmm' package (Hadfield, 2010), employing a log link function and Poisson-distributed

error. For wild-caught fish, deme origin, sex and their interaction were treated as fixed effects, and site as a random term. For purebred crosses, genotype (i.e. cross), rearing environment (E) and genotype-by-environment interaction (G × E) were considered fixed effects, and family treated as a random effect. We estimated group means and significance of fixed effects via Markov chain Monte Carlo (MCMC) sampling of their posterior distributions, conditioned on random effects.

Size distributions of wild-caught fish were plotted separately for males and females. We tested for differences in size-at-age in laboratory-reared fish using the aforementioned fixed and random effects in a linear mixed-effects model (LME), implemented in the 'LME4' package (Bates *et al.*, 2007). To test for differences in deme-specific growth, we fit size-at-age data obtained over the course of the experiment to a reparameterized version of the 'von Bertalanffy' growth function (eqn 1) in which asymptotic size does not appear as an explicit parameter, but is rather a reflection of the metabolic parameters first used to derive the model (Stamps *et al.*, 1998; Essington *et al.*, 2001).

$$L(t) = \frac{q}{k}\left(1 - e^{-kt}\right) \tag{1}$$

In metabolic terms, q and k are constants of anabolism and catabolism, respectively (Bertalanffy, 1957). In more general terms, k can also be interpreted as a rate parameter describing the trajectory to a theoretical asymptotic size, whereas q specifies the maximal growth rate, occurring during early development (Stamps *et al.*, 1998). Data comprised SL measurements obtained from scaled, digital photographs taken at four time periods (45, 120, 180 and 230 days post-hatch). These were converted to dates post-fertilization so that the model intercept could pass through the origin (i.e. null size at $t = 0$). SL (L) at time (t) was estimated via nonlinear mixed-effects modelling using the 'NLME' package (Pinheiro & Bates, 2000). Fixed effects included differences between crosses (G), rearing environments (E) and their interaction (G × E); variation among families in both model parameters was treated as random.

Overall differences in shape were formally evaluated by mixed-effects multiresponse models using the 'MCMCglmm' package. Simple, 'MANOVA-like' models testing for aforementioned main effects were contrasted with 'MANCOVA-like' models including centroid size as a covariate to control for potential allometric effects. Random effects included variation among sampling sites within each deme (for wild-caught fish) or variation among full-sib families (for laboratory-reared fish). Model selection was based on the deviance information criterion (DIC), and significance of fixed effects was determined from the posterior distribution of a full 'pseudo-parsimony' model (i.e. with or without allometry effects). Multivariate discrimination among groups in

latent shape variables was also determined by canonical variate analysis (CVA), thereby facilitating a visual representation of overall differences and similarities among groups in multivariate shape space. Additionally, this permitted the evaluation of the relative contributions of partial warps to the respective CVA axes describing divergence among groups (analogous to a traditional relative warps analysis). Degree of differentiation was further evaluated by the proportion of correct assignments of individuals to their respective groups, based on a jackknifed classification procedure, our rationale being that greater group divergence should be reflected in more individuals correctly reclassified to their cluster of origin. This is particularly germane to the case of samples from sites for which environmental salinity was essentially the opposite of the deme-specific norm (BP and CHAT): 'misclassifications' should reflect a means of identifying strong environmental effects on shape in the wild.

To infer signatures of selection on latent shape variables, we calculated the standard index of quantitative genetic differentiation (Q_{ST}; eqn 2) from purebred, laboratory-reared crosses (Spitze, 1993). Quantitative genetic variance between crosses (σ^2_{GB}) and among families within cross (σ^2_{GW}) was estimated by LME variance components analysis, with differences between rearing environments treated as fixed effects. Confidence intervals of Q_{ST} estimates were obtained by nonparametric bootstrapping (5000 iterations; stratification by family in each environment). To facilitate comparisons with other studies, we also computed P_{ST}, the 'phenotypic' equivalent of Q_{ST}, based on partial warp scores of wild-caught individuals. In these analyses, differences between sexes were partitioned out as fixed effects, with random variation between demes and among populations within demes serving as proxies for σ^2_{GB} and σ^2_{GW}, respectively. Both indices were compared to an index of putatively neutral genetic differentiation (F_{ST}) based on nine microsatellite loci described in a previous paper (McCairns & Bernatchez, 2008). We re-evaluated this marker data set, restricting analyses only to those sites/years used in the present study, by performing a hierarchical analysis of molecular variance (Yang, 1998). Ninety-five percent confidence limits were estimated from 10 000 bootstrap iterations using the 'HIERFSTAT' package (Goudet, 2005).

$$Q_{ST} = \frac{\sigma^2_{GB}}{\sigma^2_{GB} + 2(\sigma^2_{GW})} \qquad (2)$$

Finally, we used the 'MCMCGLMM' package to perform Bayesian implementations of the 'animal model' to decompose phenotypic variation into additive genetic (V_A) and direct environmental (salinity, V_E) effects. Variance components were estimated by MCMC sampling of their posterior distributions after first removing effects of fixed model terms. All models were run with an initial burn-in of 100 000 iterations, followed by an additional 100 000 iterations from which each 100th point on the Markov chain were sampled, to reduce autocorrelation between successive estimates. Parameter estimates were based on the posterior mode and bounded by the 95% posterior density interval. V_A was estimated as variation attributable to the individual (i.e. animal) level, weighted by the half-sib pedigree structure among all F_1 crosses. Effects attributable to experimental salinity and cross type, because half-sib families included both maternal and paternal hybrids, were removed as fixed model terms. Significance of V_A was determined by comparison with a simpler model including only fixed effects and using the DIC as a model selection criterion. V_E was estimated as cross-environment trait covariance (Falconer & Mackay, 1996). This was achieved by first imputing 'missing' environment-specific data for each individual on the basis of full-sib family means in the respective salinity treatments. We next ran a series of bivariate model on individualized trait values in both rearing environments. Only cross type was included as a fixed effect in these models, and significance of covariance (i.e. V_E) was assigned on the basis of whether a given trait's 95% PDI excluded zero. Lastly, we estimated the genetic correlation matrix (G) from an iterative series of bivariate models between traits. Reported narrow-sense heritabilities (h^2, diagonal elements) were estimated as the ratio of V_A to total phenotypic variance from aforementioned animal models; significance of off-diagonal elements (correlations) was based on 95% PDI estimates.

Results

The vast majority of individuals, both wild-caught and laboratory-reared, could be classified as fully plated. Although there was some variation in plate number among individuals, nearly all fish possessed between 27 and 33 lateral plates. Only 1/550 mature sticklebacks sampled from spawning sites in the St Lawrence estuary had a reduced number of lateral plates ($n = 14$). Overall, we detected no significant differences in plate number between sexes or between demes (Table 1). Similar results were also obtained under common environmental conditions, with no significant differences between crosses or experimental salinities (Table 1). However, laboratory-reared fish did exhibit somewhat more variation in plate number. Although the majority of these individuals also possessed between 27 and 33 plates, a greater proportion had reduced plate numbers (11–26 plates; 40/615 individuals). This variability was more prevalent in purebred families (FW–FW = 11/16; SW–SW = 16/16) than in hybrid crosses (13/32) and could not be attributable to specific progenitors (12/16 dams and 14/16 sires). Moreover, most of these families contained fewer than three individuals with reduced plate numbers; only three families had 5–6 offspring with this phenotype.

Table 1 Analyses of the conditional mean number of lateral plates (Cond. Mean). Estimates are based on the posterior distributions of GLMM models, conditional on random variation among sampling sites for mature individuals from the freshwater (FW) and maritime (SW) demes (Wild), or among full-sib families of purebred F_1 crosses raised in reciprocal salinity conditions (Lab).

Deme (D)	Sex (S)	Cond. Mean No. of Plates	Fixed effects	*P*-value
Wild				
FW	Female	30.5	D	0.782
	Male	30.9	S	0.672
SW	Female	30.7	D × S	0.892
	Male	30.9		

Cross (G)	Env. (E)	Cond. Mean No. of Plates	Fixed effects	*P*-value
Lab				
FW–FW	Freshwater	29.7	G	0.589
	Saltwater	28.9	E	0.502
SW–SW	Freshwater	30.4	G × E	0.143
	Saltwater	29.7		

Estimates of size frequencies of sticklebacks taken from spawning sites suggested that two age classes were likely sampled (Fig. 2). These data were also suggestive of size differences between sexes and demes: displacement of distributions suggested maritime fish are generally larger than freshwater sticklebacks and that females of the same age class are larger than males. Comparison of size-at-age data for fish reared under experimental conditions revealed no significant differences between demes ($P = 0.360$; Table 2), but a significant effect of salinity ($P = 0.005$). However, growth trajectories did differ significantly between demes (Fig. 2c): both parameters of the von Bertalanffy growth function differed significantly between crosses (i.e. deme genotypes) – parameter estimates were higher for the maritime demes – and the parameter describing initial growth rate (q) also exhibited significant salinity-related variation (Table 2).

Morphological variation

Affine (i.e. uniform) deformations accounted for only 11% of global shape variation. Nonuniform deformations accounted for the remaining 89% of shape variation; however, only the first five partial warps explained at least 5% of total shape variation, respectively. Cumulatively, the combination of these scores captured 81.9% of total shape variation in wild-caught and laboratory-reared sticklebacks (Fig. S1). Significant correlations between centroid size and most latent shape variables, partial warps 4 and 5 excluded, were indicative of potential allometry effects (Fig. S2).

Overall shape differed significantly between demes and sexes, irrespective of random variation among sites and

any potential effects of allometry (Table 3). Multivariate discrimination on the first two CVA axes captured 94% of variation in partial warp scores (ca. 77% total shape variation) and suggested a greater degree of differentiation between sexes (Fig. 3a). Interdeme differences were captured by the secondary CVA axis; moreover, these exhibited a significant degree of overlap (Fig. 3a,b). Compression and partial warp 1 contributed most to this axis, with lesser contributions from eigenvectors for shear and warps 2 and 4 (Fig. 3b). The dominant CVA axis (CV1) defined a region of shape space with no significant overlap between sexes, driven largely by a contribution from partial warp 2 and to lesser extent by partial warps 1, 3 and 4 (Fig. 3c). Procrustes superimpositions also reflected greater sexual dimorphism than differences between demes, revealing smaller interdeme differences between females ($\rho = 0.026$) and males ($\rho = 0.019$), compared to distances between sexes within each deme ($\rho_{FW} = 0.041$; $\rho_{SW} = 0.042$). Jackknife classifications were also informative in this regard, with 100% of specimens correctly separated by sex. Deme classification was less robust, with only 76% of freshwater individuals correctly assigned to their deme of origin and 75% of maritime individuals classified to their proper deme. Interestingly, females showed a greater tendency to be correctly classified by deme of origin (freshwater = 84%; maritime = 78%), compared to males (freshwater = 68%; maritime = 72%).

We detected no significant differences between genotypes of purebred F_1 crosses. ($P = 0.336$; Table 3). Two CVA axes captured 91% of variation in partial warp scores (ca. 74% of total shape variation) and suggested that cross-specific shape differences (Fig. 3d, CV1) were driven almost exclusively by compressive deformations (Fig. 3e). The second axis captured shape variation principally associated with salinity differences and was largely defined by variation in partial warp 1 scores, with lesser contributions from shear deformation and warps 3 and 5. Trends in the degree of overlap of confidence ellipses suggested that shape differences were more pronounced when fish were reared in freshwater compared to saltwater (Fig. 3d). Group classification based on morphological score also reflected this overlap, with only freshwater crosses reared in their native salinity (< 1‰) exhibiting a reasonably high degree of correct identification (78.7%); for the remaining groups, individuals were correctly classified in < 60% of cases (range 52.2–59.3%). Classification by cross type, irrespective of rearing environment (freshwater = 77.4%; maritime = 79.2%), and classification by rearing environment, irrespective of cross (freshwater = 78.7%; saltwater = 79.0%), suggested similar levels of morphological overlap due to genotypic and salinity effects. This was supported by analyses of partial warp scores, which revealed no significant genotype–environment interaction in overall morphological variation ($P = 0.192$; Table 3). Yet, despite considerable overlap among groups in multivariate space (Fig. 3d),
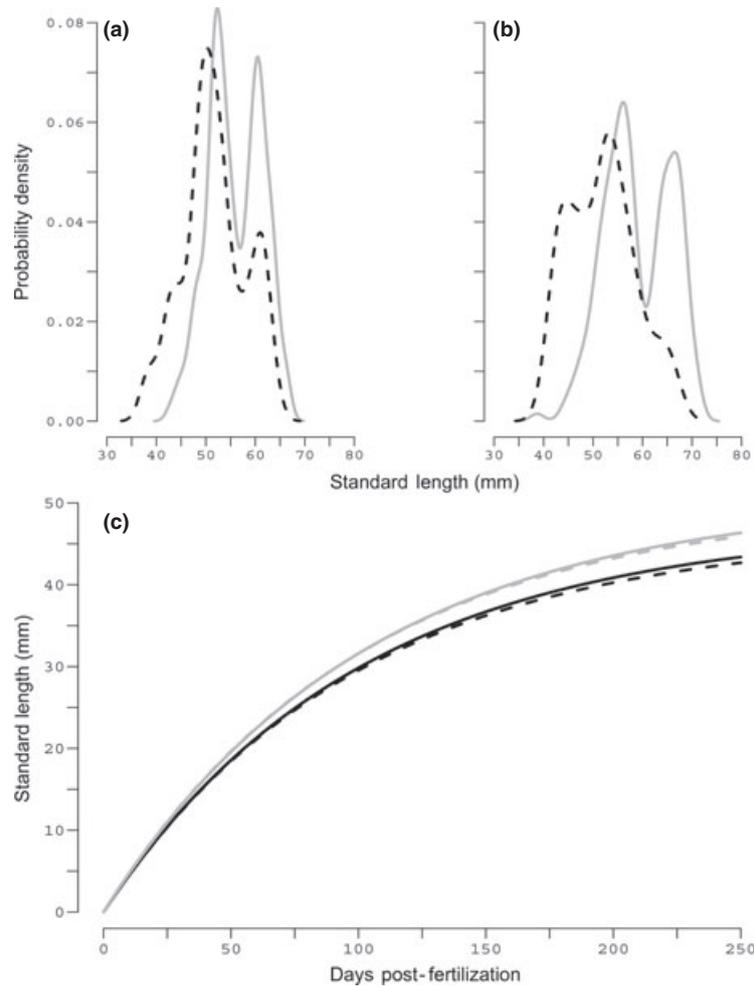
Fig. 2 Frequency distributions of standard length estimated from samples of male (a) and female (b) sticklebacks collected from spawning grounds of the St Lawrence River estuary. Samples from sites within the range of the freshwater deme are plotted as broken, black lines; solid grey lines denote samples corresponding to the maritime deme. (c) Mean growth trajectories for fish with a freshwater (black lines) or maritime (grey lines) genetic background. Individuals sampled randomly from full-sib families were reared in freshwater (< 1‰; broken lines) or saltwater (20‰; solid lines). Growth trajectories were determined by fitting size-at-age data to von Bertalanffy growth curves via nonlinear mixed-effects modelling (see Materials and Methods for model parameterization).

shape vectors differed significantly between rearing environments ($P = 0.008$; Table 3).

Individuals from sites with salinities atypical of their respective demes also exhibited significant shape differences from both deme means, and trends in distance data suggested most samples were closer in Procrustes shape space to that of the alternate deme with similar native salinities (Table 4; see also Fig. S4). In contrast, all laboratory-reared crosses appeared closer to the FW deme average, irrespective of their genetic background (Table 4). For wild fish, Riemannian shape distances between the focal groups (BP and CHAT) and the reference samples were relatively uniform, suggesting that these individuals were not significantly closer in shape space to that of their proper demes. This same

pattern was observed in classification results. Globally, 31.3% of BP fish were correctly grouped with the freshwater deme, and 25.6% of CHAT samples were correctly classified as maritime in shape space. Only 58.3% of BP females were classified as freshwater females, and the others were misidentified as maritime females. Results were far worse for BP males wherein 95.8% were improperly classified as maritime males; only 4.2% were correctly classified. Aberrant classifications were also typical for CHAT fish in which only 26.7% of females and 25.0% of males were correctly classified. Nevertheless, sexual dimorphism was largely evident, with 100% of BP individuals grouped by their correct sex. All CHAT females were also correctly identified as females, and 91.7% this site's males were correctly classified.

Table 2 ANOVA of mixed-effects models for standard length (SL) at 230 days, and for parameters of the von Bertalanffy growth model (q and k; eqn 1) of purebred F_1 crosses raised under reciprocal salinities. Fixed factors describe the effects of cross (i.e. genotype, G), environmental salinity (E) and genotype–environment interaction (G × E). Random variation among full-sib families is reported in square brackets. For SL data, significance of F ratios (P-value) is evaluated against a distribution of F ratios simulated under a null model (10 000 simulations). For growth model parameters, F ratios are evaluated on the basis of Wald's F-test. Note that sum of square variation for this model (SS*) is not provided by the 'NLME' package, so they have been approximated from the reported F ratio and within-group variance.

	d.f.	SS	F	P-value
SL				
Cross (G)	1	24.818	0.922	0.360
Env. (E)	1	216.830	8.055	0.005
G × E	1	5.766	0.214	0.645
[Family]		13.381		
Residual		26.919		

	d.f.	SS*	F	P-value
q				
Cross (G)	2	2.959	845.225	< 0.001
Env. (E)	1	0.041	23.533	< 0.001
G × E	1	4.8×10^{-4}	0.272	0.602
[Family]		1.8×10^{-3}		
k				
Cross (G)	2	2.3×10^{-3}	210.754	< 0.001
Env. (E)	1	6.5×10^{-7}	0.118	0.731
G × E	1	1.3×10^{-8}	0.002	0.961
[Family]		5.5×10^{-6}		
Residual		9.71		

Table 3 Model selection criteria for mixed-effects multiresponse models of total shape variation among mature individuals sampled from the freshwater and maritime demes (Wild) and for purebred F_1 crosses (Lab). Simple models testing for significant effects of deme (D) and sex (S), or cross (i.e. genotype, G) and environmental salinity (E), are contrasted with more complex models including centroid size as a covariate to control for potential allometric effects. Random effects includes variation among sampling sites within each deme or variation among full-sib families. Model selection is based on the deviance information criterion (DIC, parsimony model in bold) and P-values of effects determined from the posterior distribution of the parsimony models.

Model	DIC	Fixed Effects	P-value
Wild			
Null	−11869.29	D	0.002
Deme (D) + Sex (S)	−11911.00	S	< 0.001
D × S	**−11911.23**	D × S	0.106
Centroid Size (C) + D + S	−11908.65		
C + D × S	−11908.87		
C × D × S	−11906.33		
Lab			
Null	−7211.00	G	0.336
Cross (G) + Env. (E)	**−7219.80**	E	0.008
G × E	−7218.20	G × E	0.192
Centroid Size (C) + G + E	−7219.79		
C + G × E	−7215.36		
C × G × E	−7217.51		

Quantitative genetics and differentiation

In general, differentiation in partial warp scores did not exceed neutral expectations, except for uniform compressive deformation (Compr), which was significantly > F_{ST} (Fig. S6). P_{ST} estimates largely mirrored those of F_{ST}, except for that of partial warp 1 scores that exceed neutral expectation in wild-caught fish, but not in the laboratory. With the exception of lateral plate number and two latent shape variables (warp 2 and warp 5), all traits exhibited significant additive genetic variance (Table 5). Phenotypic variation in plate number and body size exhibited significant effects of environmental salinity; however, for latent shape variables, significant V_E was detected only in uniform shape scores and the first partial warp (Table 5). We detected significant genetic correlations between SL and all shape variables, excluding partial warp 3 scores (Table 6). Components of uniform variation were uncorrelated, but both shear (warp 1 and warp 4) and Compr (warp 3) were significantly correlated with at least one nonuniform component of variation. Most partial warp scores exhibited no significant genetic correlations, with the exception of warp 3 and warp 4 (Table 6).

Discussion

Shape variation described by geometric morphometric analyses of wild-caught fish suggests that stickleback demes endemic to the St Lawrence River estuary differ significantly in overall body form. Yet results of the common garden experiment were not concordant with this view, with multivariate analyses of 'total' shape revealing no significant differences between genotypes, but a significant effect of rearing environment. This highlights a potential pitfall of such powerful, multivariate techniques: taken on their own, comparative analyses between samples of wild-caught individuals could easily lead to the type of adaptationist storytelling against which Gould & Lewontin (1979) cautioned. However, by comparing results from wild-caught groups with those reared under controlled environments, we can begin to address underlying proximate causes of shape variation and generate hypotheses of their potential utility and ultimate origins.

Is shape variation plastic or genetic?

Although environmental (plastic) and genetic sources of phenotypic variation are not mutually exclusive, in an evolutionary context greater emphasis is generally placed on heritable variation as this is a necessary condition for a response to selection. This point is certainly reflected in
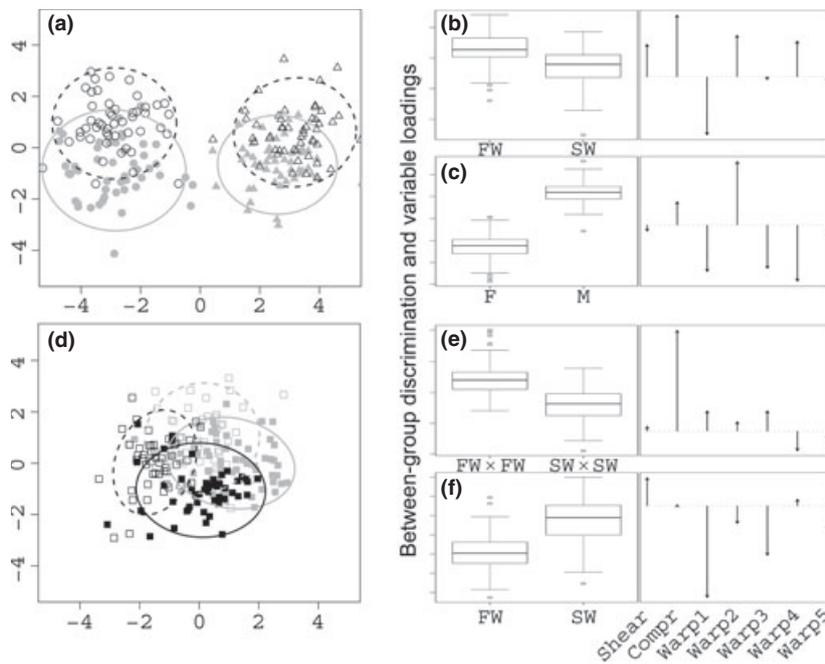
Fig. 3 Multivariate discriminant analysis (CVA) of shape vectors defining affine and partial warp scores of mature sticklebacks sampled from spawning grounds (a). Solid grey symbols denote individuals of the maritime deme, whereas open black symbols denote freshwater individuals. Females are indicated by circles and males by triangles. Respective groupings are bound by approximate 90% confidence ellipses. Mean differentiation between demes (b) and sexes (c) is also shown, with relative contributions (loadings; arrow lengths) of each shape score's contribution to group discrimination. CVA of purebred, laboratory-reared fish is plotted in (d). Grey symbols denote individuals with a maritime genetic background (SW–SW), whereas black symbols denote freshwater crosses (FW–FW). Individuals reared in FW (< 1‰) are plotted as open symbols, with approximate confidence ellipses as broken lines; individuals reared in SW (20‰) are denoted with solid symbols and lines. Discrimination between crosses (e) and rearing environments (f) is as above.

the debate surrounding the most appropriate method of estimating heritability of shape, wherein the crux of the argument hinges on the ability to predict a response to selection (Monteiro *et al.*, 2002; Klingenberg, 2003). One position holds that shape should be defined by the matrix of genetic covariation, because this is the only known method of predicting a multivariate response to selection (Klingenberg & Leamy, 2001; Klingenberg, 2003). The alternate view suggests that univariate measures of compound variables can sufficiently capture this aspect of evolution (Monteiro *et al.*, 2003), likely due to observations from explicit factor analytic modelling of G matrices showing that (co) heritable variation in phenotypic space can be reduced to fewer dimensions than numbers of component traits (McGuigan & Blows, 2007). Although this data set may be insufficient to meaningfully address this debate, we contend that there is at least sufficient evidence for heritable variation of shape within this system, given that both univariate and multivariate decompositions of variance for latent shape variables revealed significant estimates.

Indices of quantitative differentiation were equally informative, allowing for inference into the roles of diversifying selection or drift in shaping phenotypic divergence. In contrast to other freshwater stickleback populations for which strong signals of selective divergence in both standard morphometric and geometric measures of body form are prevalent (Leinonen *et al.*, 2006; Raeymaekers *et al.*, 2007; Berner *et al.*, 2011; Hendry *et al.*, 2011) and despite significant heritable variation upon which selection could act, only one trait index exhibited a signal of differentiation that might be

interpreted as evidence of diversifying selection: uniform compressive variation (Compr). Interestingly, this was also the only shape variable to exhibit significant genotypic differences between crosses reared in the laboratory and had the highest point estimate for heritability of all latent variables. Thus, it could be argued that uniform compressive difference between demes has evolved under directional selection. However, this interpretation should also be tempered by the fact that this accounts for only 6.6% of variation of total morphology. Additionally, Q_{ST} indices for all other latent shape variables fell within the range of neutral divergence. Although most of these shape descriptors exhibited significant additive genetic variance, trait heritability does not preclude the contribution of environment as an important source of phenotypic variation. Indeed, we have demonstrated significant salinity effects on size, uniform shape scores and the partial warp score with the greatest contribution of shape variation (warp 1, 28% of total shape), which suggests a strong potential for direct environmental influences on body form. It is perhaps equally telling that partial warp 1 scores also gave conflicting signals in comparisons of Q_{ST} and P_{ST} estimates (Fig. S6). This 'trait' also made the greatest contribution to the CVA axis capturing salinity-related variation among laboratory crosses (Fig. 3f) while exhibiting relatively similar contributions of additive genetic and environmental variance to its expression (Table 5). Taken together, these observations lead us to conclude that overall differences between demes in general body morphology are predominantly plastic in nature.

Table 4 Differences in Procrustes superimpositions of landmark data. Focal groups are contrasted with mean shape data corresponding to mature individuals sampled from the freshwater (FW) and maritime (SW) demes inhabiting sites with salinities typical of the respective environments (FW: CR and LEV; SW: BAC and RIKI). Fish sampled from sites with atypical environmental conditions (Wild) were contrasted only with individuals of the same sex. Purebred F_1 crosses (Lab) from reciprocal salinity conditions (indicated in parentheses) were contrasted with mean shape data averaged over sexes. ρ is the Riemannian shape distance between Procrustes superimpositions, and $\lambda_{\min}$ represents an unbiased difference statistic, evaluated via bootstrapping (see Materials and Methods for details). Visualizations of superimpositions are available as online supplementary material (Figs S4 and S5).

	FW Deme			SW Deme		
Focal group	ρ	$\lambda_{\min}$	P-value	ρ	$\lambda_{\min}$	P-value
Wild						
BP (females)	0.035	617.4	0.001	0.035	592.6	0.001
BP (males)	0.035	1122.0	0.001	0.030	716.1	0.001
CHAT (females)	0.032	414.3	0.001	0.030	325.9	0.001
CHAT (males)	0.029	307.8	0.001	0.030	394.7	0.001
Lab						
FW–FW (< 1‰)	0.049	1121.9	0.001	0.052	1692.2	0.001
FW–FW (20‰)	0.025	440.9	0.001	0.035	646.8	0.001
SW–SW (< 1‰)	0.041	555.4	0.001	0.041	695.2	0.001
SW–SW (20‰)	0.026	687.6	0.001	0.034	830.2	0.001

Plasticity in stickleback morphology is not unique to the St Lawrence system. Indeed, early experiments into the nature of trophic dimorphism in freshwater sticklebacks demonstrated how morphological variation in the direction of specialized morphotypes could be induced by diet (Day *et al.*, 1994; Day & McPhail, 1996). Interestingly, seeming 'preadaptive' morphological plasticity has been observed in ancestral sticklebacks reared on derived diets (Wund *et al.*, 2008, 2012), although more recent observations suggest that genetic variation masked in the ancestral population/environment may also be released in a novel environment, (McGuigan *et al.*, 2011). Nevertheless, plastic morphological responses have been associated with habitat shifts, likely induced by developmental plasticity (Spoljaric & Reimchen, 2011). Moreover, similar growth-related morphological changes have also been observed in response to olfactory predator cues (Frommen *et al.*, 2011). Although conditions simulated in our microcosms do not reflect the full suite of environmental variation potentially affecting body shape, environmental salinity has been shown to influence morphological variation in pupfish species (Collyer *et al.*, 2007). Similarly, significant shape differences have been observed in transplanted marine sticklebacks after a single generation in freshwater ponds (Kristjánsson, 2005). Both the similarity of shape between genotypes reared in a shared salinity and the observation that wild-caught individuals from sites with atypical salinities exhibited trends in Procrustes superimposition distance metrics placing them 'closer' in shape space to heterodemic individuals (Table 4; Fig. S4) seem to underscore a common sensitivity of St Lawrence demes to environmental inputs. An exception, however, can be found in compressive variation, which displayed the lowest proportional contribution of direct environmental effects, relative to total phenotypic variation. Interestingly, this was also the one latent shape variable to show a signature of divergent selection between demes. Although the confluence of relatively high V_A and low V_E in a trait flagged as significantly divergent may speak to the utility of the Q_{ST} index as a tool to infer signatures of selection, the apparent absence of correlated signals of selection is puzzling.

Sexual dimorphism and allometry revisited

Shape differences between the sexes, even within demes, were two-fold greater than those estimated between demes. Such strong sexual dimorphism in body shape has been observed in other stickleback

Table 5 Partitioning of phenotypic variance into additive genetic effects (V_A), direct effects of environmental salinity (V_E) and residual error ($V_{residual}$). Parameter estimates are based on the posterior mode of 1000 MCMC samples; 95% posterior density interval estimates are in parentheses. Nonsignificant parameter estimates are italicized and are assumed to include zero (lower PDI).

Trait	DIC null	DIC V_A	V_A (95% PDI)	V_E (95% PDI)	$V_{residual}$ (95% PDI)
No. plates	3408.0	3945.5	*0.033 (0–0.037)*	0.002 (1.4×10^{-4}–0.005)	0.021 (0.014–0.027)
SL (45 days)	8012.9	6145.5	5.623 (3.456–6.812)	1.404 (1.013–1.631)	0.217 (0.001–1.032)
SL (230 days)	2345.6	2266.9	14.552 (4.545–35.736)	6.94 (1.845–10.159)	18.097 (1.925–31.153)
Shear	−2126.5	−2153.2	3.4×10^{-5} (1.2×10^{-5}–8.8×10^{-5})	1.9×10^{-5} (1.3×10^{-7}–4.2×10^{-5})	1.2×10^{-4} (5.5×10^{-5}–1.7×10^{-4})
Compr	−2024.0	−2071.8	6.4×10^{-5} (2.4×10^{-5}–1.5×10^{-4})	1.0×10^{-5} (6.5×10^{-9}–5.0×10^{-5})	1.5×10^{-4} (6.0×10^{-5}–2.1×10^{-4})
Warp 1	−1587.0	−1604.9	1.4×10^{-4} (4.3×10^{-5}–3.7×10^{-4})	1.1×10^{-4} (4.8×10^{-6}–1.8×10^{-4})	5.0×10^{-4} (2.6×10^{-4}–7.4×10^{-4})
Warp 2	−1790.2	−1786.6	*4.3×10^{-5} (0–1.0×10^{-4})*	*4.2×10^{-5} (0–1.1×10^{-4})*	3.6×10^{-4} (2.2×10^{-4}–4.6×10^{-4})
Warp 3	−1898.4	−1956.6	1.1×10^{-4} (3.2×10^{-5}–2.3×10^{-4})	*1.5×10^{-5} (0–6.0×10^{-5})*	2.0×10^{-4} (8.1×10^{-5}–2.8×10^{-4})
Warp 4	−1951.6	−1972.5	4.8×10^{-4} (1.2×10^{-5}–1.4×10^{-4})	*1.3×10^{-5} (0–6.0×10^{-5})*	2.0×10^{-4} (1.0×10^{-4}–2.7×10^{-4})
Warp 5	−2051.0	−2048.2	*2.1×10^{-5} (0–4.9×10^{-5})*	*5.9×10^{-6} (0–4.4×10^{-5})*	1.8×10^{-4} (1.2×10^{-4}–2.2×10^{-4})

DIC, deviance information criterion; MCMC, Markov chain Monte Carlo; SL, standard length.

Table 6 G matrix of size and shape variables. Narrow-sense heritability estimates (h^2) and their 95% PDIs are boldfaced on the diagonal. Point estimates of the genetic correlation (G) between traits are in the lower triangle. Nonsignificant estimates are presumed to be zero; however, point estimates are in the upper triangle, with nonsignificant estimates italicized. 95% PDI estimates are in the parentheses in the upper triangle, with nonsignificant estimates italicized.

	SL	Shear	Compr	Warp 1	Warp 2	Warp 3	Warp 4	Warp 5
SL	**0.404** **(0.149 to 0.760)**	(0.419 to 0.954)	(0.314 to 0.853)	(−0.986 to −0.867)	(−0.983 to −0.832)	(−0.776 to 0.071)	(0.210 to 0.911)	(−0.936 to −0.353)
Shear	0.791	**0.195** **(0.066 to 0.444)**	(−0.575 to 0.574)	(−0.930 to −0.236)	(−0.807 to 0.209)	(−0.844 to 0.209)	(0.144 to 0.935)	(−0.428 to 0.782)
Compr	0.638	−0.159	**0.245** **(0.116 to 0.571)**	(−0.877 to 0.189)	(−0.833 to 0.280)	(−0.952 to −0.593)	(−0.642 to 0.547)	(−0.914 to 0.128)
Warp 1	−0.958	−0.764	−0.509	**0.179** **(0.059 to 0.439)**	(−0.362 to 0.815)	(−0.649 to 0.550)	(−0.888 to 0.101)	(−0.595 to 0.662)
Warp 2	−0.969	−0.664	−0.460	0.265	**0.108** **(0.041 to 0.229)**	(−0.779 to 0.342)	(−0.769 to 0.413)	(−0.751 to 0.379)
Warp 3	−0.301	−0.531	−0.845	−0.014	−0.494	**0.272** **(0.135 to 0.619)**	(−0.895 to −0.181)	(−0.732 to 0.521)
Warp 4	0.721	0.723	−0.113	−0.728	−0.147	−0.749	**0.172** **(0.065 to 0.461)**	(−0.374 to 0.795)
Warp 5	−0.806	0.133	−0.564	0.126	−0.485	−0.055	−0.494	**0.105** **(0.035 to 0.220)**

SL, standard length.

populations (Kitano *et al.*, 2007; Aguirre *et al.*, 2008). Further, the degree of sexual dimorphism has been shown to vary among populations and habitats (Spoljaric & Reimchen, 2008; Leinonen *et al.*, 2011a), although many of these populations also exhibit potentially confounding relationships between partial warp scores and centroid size. Moreover, as female-specific differences in body form, independent of gravity, may only be apparent after the initial diversion of bioenergetic resources from growth to sexual reproduction, size-related differences may be an important determinant of sexual dimorphism (Kitano *et al.*, 2007). Conversely, QTL mapping has suggested a strong genetic component to these differences: although many QTL for morphological landmarks are common between sexes, QTL with the greatest effect sizes are located within a sex-determining chromosomal region (Albert *et al.*, 2008). As no size QTL were detected in the same region, it was presumed that sexual dimorphism could not be attributable to allometric differences. Likewise, we have concluded that allometric effects are unlikely to be driving patterns of sexual dimorphism in this system (Table 3), despite observed size differences between males and females.

Both ontogenetic allometric effects and sexual differences have been reported to explain shape variation in marine shiner perch over a broad latitudinal cline (Woods, 2007). Thus, it is conceivable that relationships between size and shape could influence functional morphology independent of sexual dimorphism, although Reid & Peichel (2010) hypothesize an interesting counterpoint if antagonistic interactions between sexual and natural selection were to differ between males and females. Observations of greater divergence between sexes than between demes may also pose the question of whether sexual dimorphism is potentially mitigating diversifying selection between habitats. For instance, Bolnick & Lau (2008) observed that sexual dimorphism in sticklebacks was strongest in habitats where disruptive selection should have been predominant and that dimorphism and disruptive selection were negatively correlated. Certainly disentangling sexual dimorphism from other putatively adaptive functional polymorphisms may be complicated by a tendency for many sexually based dimorphisms to reflect ecologically based differences between trophic specialists (Reimchen & Nosil, 2006). Further difficulties of interpretation are likely inevitable in the face of size-based differences between focal groups, in as much as size can be an ultimate target of selection, but also a potential mechanistic link between plastic and genetic determinants of phenotypic expression (Schaefer & Bookstein, 2009; although see McGuigan *et al.*, 2011). And although McGuigan *et al.* (2010) have shown that evidence of allometry alone does not predict a correlated response to selection between growth/size and body form, it is generally assumed that allometry is

ultimately pleiotropic in nature (Klingenberg, 2010). Given genotypically based differences in growth trajectories (Table 2), in addition to significant genetic correlations between SL and compression, it was surprising that we observed no correlated signatures of selection on body size. We cannot fully explain this result, although such body shape/size coevolution may be tempered by constraints found in the G matrix of ancestral/colonizing populations (Leinonen *et al.*, 2011b).

Lateral plate homogeneity: does an exception help prove an emerging rule?

The existence of three distinct morphotypes defined by the number and corporal location of protective lateral plates is a polymorphism common throughout the stickleback's global distribution (Bell & Foster, 1994; Bell, 2001). The ancestral marine form is fully plated, and certain estuarine populations may exhibit varying numbers of plates (Klepaker, 1996; Bell, 2001). However, the colonization of freshwater has typically favoured morphs lacking the full complement of lateral plates (Bell *et al.*, 2004; Raeymaekers *et al.*, 2005; Aguirre *et al.*, 2008). Yet, this distribution of morphotypes was not observed in the St Lawrence; in fact, all sticklebacks within freshwater are fully plated. Given the Mendelian architecture of the lateral plate polymorphism (Cresko *et al.*, 2004; Colosimo *et al.*, 2005), it was perhaps not surprising that common garden experiments revealed no significant additive variance for plate number. Interestingly, we did detect a significant environmental component to phenotypic variance; however, the magnitude of this effect was substantially smaller than total phenotypic variance (Table 5). Thus, it is unlikely that early hypotheses proposed to explain the distribution of plate morph focusing on environmental effects, such as differential survival and plate genesis as a function of ion concentration (Heuts, 1947), or climatic effects (Hagen & Moodie, 1982), can explain the apparent monomorphism of St Lawrence sticklebacks. We contend, based upon our observations, in addition to independent experimental tests of morph-specific salinity-dependant mortality (Marchinko & Schluter, 2007), that Heuts (1947) hypothesis may be put to rest in favour of alterative explanations, namely differential predation by vertebrate or macroinvertebrate enemies (Reimchen, 1994, 2000; Bergstrom & Reimchen, 2003; Patankar *et al.*, 2006). Important piscine stickleback predators are common throughout the estuary, although the species compositions do vary with environmental salinity. The homogeneity of plate morphs within the St Lawrence estuary most likely reflects the complexity of this ecosystem, rather than a fundamentally divergent genetic architecture from other stickleback systems.

Conclusions

Geometric morphometric analyses of wild-caught individuals suggest that stickleback demes endemic to the St Lawrence River estuary differ in body form; however, purebred crosses reared under controlled environmental conditions indicate that variation within this system is most likely plastic. Although multivariate descriptors of shape are heritable, many are equally influenced by direct environmental contributions. Additionally, we show evidence of both allometry and genetic correlations, but no concomitant correlated signatures of selection on morphological variation, despite evidence of adaptive divergence between demes – previous work in this system has demonstrated weak population structure linked to environmental (i.e. salinity) heterogeneity, and adaptive differences in osmoregulatory physiology. Taken in their entirety, observations on stickleback demes of the St Lawrence estuary would seem to validate the perspective that inherently 'plastic' traits (e.g. behaviour and physiology) may be among the first to diverge among populations. Yet, one must also consider that the most divergent trait within this relatively young system, differential osmoregulatory capacity, retains a pattern of environmental sensitivity found within the ancestral phenotype. Thus, a predominant role of plasticity reiterated for morphology serves also as an important reminder for the unique capacity of the threespine stickleback to cope with heterogeneous environments over the species' broad geographical range. In more general terms, these observations may speak to the fundamental importance of trait plasticity in modulating the evolutionary process, given that a species contemporary distribution may rely as much and perhaps more on plasticity than local adaptation. This may be particularly the case for *G. aculeatus* in many of its freshwater forms and localities.

Acknowledgements

References

Abramoff, M.D., Magelhaes, P.J. & Ram, S.J. 2004. Image processing with ImageJ. *Biophotonics Int.* **11**: 36–42.

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.

Aguirre, W.E., Ellis, K.E., Kusenda, M. & Bell, M.A. 2008. Phenotypic variation and sexual dimorphism in anadromous threespine stickleback: implications for postglacial adaptive radiation. *Biol. J. Linn. Soc.* **95**: 465–478.

Albert, A.Y.K., Sawaya, S., Vines, T.H., Knecht, A.K., Miller, C.T., Summers, B.R. et al. 2008. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution* **62**: 76–85.

Amaral, G.J.A., Dryden, I.L. & Wood, A.T.A. 2007. Pivotal bootstrap methods for k-sample problems in directional statistics and shape analysis. *J. Am. Stat. Assoc.* **102**: 695–707.

Arif, S., Aguirre, W.E. & Bell, M.A. 2009. Evolutionary diversification of opercle shape in Cook Inlet threespine stickleback. *Biol. J. Linn. Soc.* **97**: 832–844.

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

Bates, D.M., Maechler, M. & Bolker, B.M. 2007. lme4: linear mixed-effects models using S4 classes.

Baumgartner, J.V. 1995. 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.

Beldade, P., Mateus, A.R.A. & Keller, R.A. 2011. Evolution and molecular mechanisms of adaptive developmental plasticity. *Mol. Ecol.* **20**: 1347–1363.

Bell, M.A. 1994. Paleobiology and evoslution of threespine stickleback. In: *The Evolutionary Biology of the Threespine Stickleback* (M.A. Bell & S.A. Foster, eds), pp. 438–469. Oxford University Press, Oxford, UK.

Bell, M.A. 2001. Lateral plate evolution in the threespine stickleback: getting nowhere fast. *Genetica* **112**: 445–461.

Bell, M.A. & Foster, S.A. 1994. Introduction to the evolutionary biology of the threespine stickleback. In: *The Evolutionary Biology of the Threespine Stickleback* (M.A. Bell & S.A. Foster, eds), pp. 1–27. Oxford University Press, Oxford, UK.

Bell, M.A., Aguirre, W.E. & Buck, N.J. 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* **58**: 814–824.

Bergstrom, C.A. & Reimchen, T.E. 2003. Asymmetry in structural defenses: insights into selective predation in the wild. *Evolution* **57**: 2128–2138.

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

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

Bertalanffy, L.V. 1957. Quantitative laws in metabolism and growth. *Q. Rev. Biol.* **32**: 217–231.

Bolnick, D.I. & Lau, O.L. 2008. Predictable patterns of disruptive selection in stickleback in postglacial lakes. *Am. Nat.* **172**: 1–11.

Bookstein, F.L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**: 567–585.

Caldecutt, W.J. & Adams, D.C. 1998. Morphometrics of trophic osteology in the threespine stickleback, *Gasterosteus aculeatus*. *Copeia* **1998**(4): 827–838.

Claude, J. 2008. *Morphometrics with R*. Springer, New York, NY.

Collyer, M.L., Stockwell, C.A., Dean, C.A. & Reiser, M.H. 2007. Phenotypic plasticity and contemporary evolution in introduced populations: evidence from translocated populations of white sands pupfish (*Cyrpinodon tularosa*). *Ecol. Res.* **22**: 902–910.

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

Cresko, W.A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P. et al. 2004. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc. Natl Acad. Sci. USA* **101**: 6050–6055.

Currie, A.J., Ganeshanandam, S., Noiton, D.A., Garrick, D., Shelbourne, C.J.A. & Oraguzie, N. 2000. Quantitative evaluation of apple (*Malus × domestica* Borkh.) fruit shape by principal component analysis of Fourier descriptors. *Euphytica* **111**: 219–227.

Day, T. & McPhail, J.D. 1996. The effect of behavioural and morphological plasticity on foraging efficiency in the threespine stickleback (*Gasterosteus* sp.). *Oecologia* **108**: 380–388.

Day, T., Pritchard, J. & Schluter, D. 1994. A comparison of two sticklebacks. *Evolution* **48**: 1723–1734.

Dryden, I.L., Koloydenko, A. & Zhou, D.W. 2009. Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Ann. Appl. Stat.* **3**: 1102–1123.

Dryden, I.L. 2012. *The Shapes Package: Statistical Shape Analysis in R*. R Foundation for Statistical Computing, Vienna, Austria. Contributed package. http://www.stat.sc.edu/~dryden/shapes/.

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

Essington, T.E., Kitchell, J.F. & Walters, C.J. 2001. The von Bertalanffy growth function, bioenergetics, and the consumption rates of fish. *Can. J. Fish. Aquat. Sci.* **58**: 2129–2138.

Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*. Prentice Hall, Upper Saddle River, NJ, USA.

Feder, M.E. & Mitchell-Olds, T. 2003. Evolutionary and ecological functional genomics. *Nat. Rev. Genet.* **4**: 651–657.

Frommen, J.G., Herder, F., Engqvist, L., Mehlis, M., Bakker, T.C.M., Schwarzer, J. et al. 2011. Costly plastic morphological responses to predator specific odour cues in three-spined sticklebacks (*Gasterosteus aculeatus*). *Evol. Ecol.* **25**: 641–656.

Georga, I. & Koumoundouros, G. 2010. Thermally induced plasticity of body shape in adult zebrafish *Danio rerio* (Hamilton, 1822). *J. Morphol.* **271**: 1319–1327.

Ghalambor, C.K., McKay, J.K., Carroll, S.P. & Reznick, D.N. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct. Ecol.* **21**: 394–407.

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

Gotthard, K. & Nylin, S. 1995. Adaptive plasticity and plasticity as an adaptation: a selective review of plasticity in animal morphology and life-history. *Oikos* **74**: 3–17.

Goudet, J. 2005. hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* **5**: 184–186.

Gould, S.J. & Lewontin, R.C. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. Series B Biol. Sci.* **205**: 581–598.

Hadfield, J. 2010. MCMC methods for multi-response generalised linear mixed models: the MCMCglmm r package. *J. Stat. Softw.* **33**: 1–22.

Hagen, D.W. & Moodie, G.E.E. 1982. Polymorphism for plate morphs in *Gasterosteus aculeatus* on the east coast of Canada and an hypothesis for their global distribution. *Can. J. Zool.* **60**: 1032–1042.

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

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

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.

Kimmel, C.B., Aguirre, W.E., Ullmann, B., Currey, M. & Cresko, W.A. 2008. Allometric change accompanies opercular shape evolution in Alaskan threespine sticklebacks. *Behaviour* **145**: 669–691.

Kitano, J., Mori, S. & Peichel, C.L. 2007. Sexual dimorphism in the external morphology of the threespine stickleback (*Gasterosteus aculeatus*). *Copeia* **2007**(2): 336–349.

Klepaker, T. 1996. Lateral plate polymorphism in marine and estuarine populations of the threespine stickleback (*Gasterosteus aculeatus*) along the coast of Norway. *Copeia* **1996**(4): 832–838.

Klingenberg, C.P. 2003. Quantitative genetics of geometric shape: heritability and the pitfalls of the univariate approach. *Evolution* **57**: 191–195.

Klingenberg, C.P. 2010. Evolution and development of shape: integrating quantitative approaches. *Nat. Rev. Genet.* **11**: 623–635.

Klingenberg, C.P. & Leamy, L.J. 2001. Quantitative genetics of geometric shape in the mouse mandible. *Evolution* **55**: 2342–2352.

Klingenberg, C.P., Leamy, L.J., Routman, E.J. & Cheverud, J.M. 2001. Genetic architecture of mandible shape in mice: effects of quantitative trait loci analyzed by geometric morphometrics. *Genetics* **157**: 785–802.

Kristjánsson, B.K. 2005. Rapid morphological changes in threespine stickleback, *Gasterosteus aculeatus*, in freshwater. *Environ. Biol. Fishes* **74**: 357–363.

Kruuk, L.E.B. 2004. Estimating genetic parameters in natural populations using the 'animal model'. *Philos. Trans. R. Soc. Lond. Series B Biol. Sci.* **359**: 873–890.

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., Cano, J.M. & Merilä, J. 2011a. Genetic basis of sexual dimorphism in the threespine stickleback *Gasterosteus aculeatus*. *Heredity* **106**: 218–227.

Leinonen, T., Cano, J.M. & Merilä, J. 2011b. Genetics of body shape and armour variation in threespine sticklebacks. *J. Evol. Biol.* **24**: 206–218.

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

Marcil, J., Swain, D.P. & Hutchings, J.A. 2006. Genetic and environmental components of phenotypic variation in body shape among populations of Atlantic cod (*Gadus morhua* L.). *Biol. J. Linn. Soc.* **88**: 351–365.

McCairns, R.J.S. & Bernatchez, L. 2008. Landscape genetic analyses reveal cryptic population structure and putative selection gradients in a large-scale estuarine environment. *Mol. Ecol.* **17**: 3901–3916.

McCairns, R.J.S. & Bernatchez, L. 2010. Adaptive divergence between freshwater and marine sticklebacks: insights into the role of phenotypic plasticity from an integrated analysis of candidate gene expression. *Evolution* **64**: 1029–1047.

McGuigan, K. & Blows, M.W. 2007. The phenotypic and genetic covariance structure of drosphilid wings. *Evolution* **61**: 902–911.

McGuigan, K., Nishimura, N., Currey, M., Hurwit, D. & Cresko, W.A. 2010. Quantitative genetic variation in static allometry in the threespine stickleback. *Integr. Comp. Biol.* **50**: 1067–1080.

McGuigan, K., Nishimura, N., Currey, M., Hurwit, D. & Cresko, W.A. 2011. Cryptic genetic variation and body size evolution in threespine stickleback. *Evolution* **65**: 1203–1211.

McKinnon, J.S. & Rundle, H.D. 2002. Speciation in nature: the threespine stickleback model systems. *Trends Ecol. Evol.* **17**: 480–488.

Moczek, A.P., Sultan, S., Foster, S., Ledón-Rettig, C., Dworkin, I., Nijhout, H.F. *et al.* 2011. The role of developmental plasticity in evolutionary innovation. *Proc. R. Soc. B Biol. Sci.* **278**: 2705–2713.

Monteiro, L.R., Diniz, J.A.F., dos Reis, S.F. & Araújo, E.D. 2002. Geometric estimates of heritability in biological shape. *Evolution* **56**: 563–572.

Monteiro, L.R., Diniz, J.A.F., dos Reis, S.F. & Araújo, E.D. 2003. Shape distances in general linear models: are they really at odds with the goals of morphometrics? A reply to Klingenberg. *Evolution* **57**: 196–199.

Parsons, K.J. & Robinson, B.W. 2007. Foraging performance of diet-induced morphotypes in pumpkinseed sunfish (*Lepomis gibbosus*) favours resource polymorphism. *J. Evol. Biol.* **20**: 673–684.

Parsons, K.J., Robinson, B.W. & Hrbek, T. 2003. Getting into shape: an empirical comparison of traditional truss-based morphometric methods with a newer geometric method applied to New World cichlids. *Environ. Biol. Fishes* **67**: 417–431.

Patankar, R., von Hippel, F.A. & Bell, M.A. 2006. Extinction of a weakly armoured threespine stickleback (*Gasterosteus aculeatus*) population in Prator Lake, Alaska. *Ecol. Freshw. Fish* **15**: 482–487.

Pinheiro, J.C. & Bates, D.M. 2000. *Mixed-Effects Models in S and S-Plus*. Springer, New York, NY, USA.

Potthoff, T. 1984. Clearing and staining techniques. In: *Ontogeny and Sytematics of Fishes* (H.G. Moser, W.J. Richards, D.M. Cohen, M.P. Fahay, A.W. Kendall Jr & S.L. Richardson, eds), pp. 35–37. American Society of Ichthyologists and Herpetologists, Lawrence, KS.

R Development Core Team 2007. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raeymaekers, J.A.M., Maes, G.E., Audenaert, E. & Volckaert, F.A.M. 2005. Detecting Holocene divergence in the anadromous-freshwater three-spined stickleback (*Gasterosteus aculeatus*) system. *Mol. Ecol.* **14**: 1001–1014.

Raeymaekers, J.A.M., Van Houdt, J.K.J., Larmuseau, M.H.D., Geldof, S. & Volckaert, F.A.M. 2007. Divergent selection as

revealed by Pst and QTL-based Fst in three-spined stickleback (*Gasterosteus aculeatus*) populations along a coastal-inland gradient. *Mol. Ecol.* **16**: 891–905.

Reid, D.T. & Peichel, C.L. 2010. Perspectives on the genetic architecture of divergence in body shape in sticklebacks. *Integr. Comp. Biol.* **50**: 1057–1066.

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

Reimchen, T.E. 2000. Predator handling failures of lateral plate morphs in *Gasterosteus aculeatus*: functional implications for the ancestral plate condition. *Behaviour* **137**: 1081–1096.

Reimchen, T.E. & Nosil, P. 2006. Replicated ecological landscapes and the evolution of morphological diversity among Gasterosteus populations from an archipelago on the west coast of Canada. *Can. J. Zool.* **84**: 643–654.

Schaefer, K. & Bookstein, F.L. 2009. Does geometric morphometrics serve the needs of plasticity research? *J. Biosci.* **34**: 589–599.

Schluter, D. 1996. Ecological speciation in postglacial fishes. *Philos. Trans. R. Soc. Lond. Series B Biol. Sci.* **351**: 807–814.

Schluter, D., Clifford, E.A., Nemethy, M. & McKinnon, J.S. 2004. Parallel evolution and inheritance of quantitative traits. *Am. Nat.* **163**: 809–822.

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

Spitze, K. 1993. Population-structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* **135**: 367–374.

Spoljaric, M.A. & Reimchen, T.E. 2007. 10 000 years later: evolution of body shape in Haida Gwaii three-spined stickleback. *J. Fish Biol.* **70**: 1484–1503.

Spoljaric, M.A. & Reimchen, T.E. 2008. Habitat-dependent reduction of sexual dimorphism in geometric body shape of Haida Gwaii threespine stickleback. *Biol. J. Linn. Soc.* **95**: 505–516.

Spoljaric, M.A. & Reimchen, T.E. 2011. Habitat-specific trends in ontogeny of body shape in stickleback from coastal archipelago: potential for rapid shifts in colonizing populations. *J. Morphol.* **272**: 590–597.

Stamps, J.A., Mangel, M. & Phillips, J.A. 1998. A new look at relationships between size at maturity and asymptotic size. *Am. Nat.* **152**: 470–479.

University of Oregon Stickleback Research Site 2008. Crossing and rearing protocols.

Vincent, W.F. & Dodson, J.J. 1999. The St. Lawrence River, Canada-USA: the need for an ecosystem-level understanding of large rivers. *Jpn J. Limnol.* **60**: 29–50.

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

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

West-Eberhard, M.J. 2005. Developmental plasticity and the origin of species differences. *Proc. Natl Acad. Sci. USA* **102**: 6543–6549.

Woods, P.J. 2007. Habitat-dependent geographical variation in ontogenetic allometry of the shiner perch *Cymatogaster aggregata* Gibbons (Teleostei : Embiotocidae). *J. Evol. Biol.* **20**: 1783–1798.

Workman, M.S., Leamy, L.J., Routman, E.J. & Cheverud, J.M. 2002. Analysis of quantitative trait locus effects on the size and shape of mandibular molars in mice. *Genetics* **160**: 1573–1586.

Wund, M.A., Baker, J.A., Clancy, B., Golub, J.L. & Fosterk, S.A. 2008. A test of the ''Flexible stem'' model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *Am. Nat.* **172**: 449–462.

Wund, M.A., Valena, S., Wood, S. & Baker, J.A. 2012. Ancestral plasticity and allometry in threespine stickleback reveal phenotypes associated with derived, freshwater ecotypes. *Biol. J. Linn. Soc.* **105**: 573–583.

Yang, R.C. 1998. Estimating hierarchical F-statistics. *Evolution* **52**: 950–956.

Zelditch, M.L., Swiderski, D.L., Sheets, H.D. & Fink, W.L. 2004. *Geometric Morphometrics for Biologists: A Primer*. Elsevier Academic Press, London, UK.

Zimmerman, E., Palsson, A. & Gibson, G. 2000. Quantitative trait loci affecting components of wing shape in *Drosophila melanogaster*. *Genetics* **155**: 671–683.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Screen plot of eigenvalues, expressed as a proportion of the sum of all eigenvalues, for partial warp scores.

Figure S2 Correlations between partial warp scores and centroid size.

Figure S3 Thin-plate spline (TPS) deformation grids for freshwater (a) and maritime (b) females, and for freshwater (c) and maritime (d) males; symbols denoting locations of landmarks are as per Fig. 3a.

Figure S4 Procrustes superimpositions of mean shape for females (a) and males (b) originating from the Baie St-Paul (BP) site, and for females (c) and males (d) from Cap Chat (CHAT).

Figure S5 Procrustes superimpositions of mean shape for purebred F_1 crosses contrasted with those of wild-caught sticklebacks from two demes of the St Lawrence estuary.

Figure S6 Box-percentile plots (description inset; lower left panel) of the distribution of bootstrapped Q_{ST} and P_{ST} values for partial warp scores and total number of lateral plates (No. Plates; data square root transformed).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.