Molecular Ecology (2008) doi: 10.1111/j.1365-294X.2008.03884.x

Landscape genetic analyses reveal cryptic population structure and putative selection gradients in a large-scale estuarine environment

R. J. SCOTT McCAIRNS and LOUIS BERNATCHEZ
Département de biologie, Québec Océan, Université Laval, Québec, Canada G1V 0A6

Abstract

Disentangling the relative contributions of selective and neutral processes underlying phenotypic and genetic variation under natural, environmental conditions remains a central challenge in evolutionary ecology. However, much of the variation that could be informative in this area of research is likely to be cryptic in nature; thus, the identification of wild populations suitable for study may be problematic. We use a landscape genetics approach to identify such populations of three-spined stickleback inhabiting the Saint Lawrence River estuary. We sampled 1865 adult fish over multiple years. Individuals were genotyped for nine microsatellite loci, and georeferenced multilocus data were used to infer population groupings, as well as locations of genetic discontinuities, under a Bayesian model framework (GENELAND). We modelled environmental data using nonparametric multiple regression to explain genetic differentiation as a function of spatio-ecological effects. Additionally, we used genotype data to estimate dispersal and gene flow to parameterize a simple model predicting adaptive vs. plastic divergence between demes. We demonstrate a bipartite division of the genetic landscape into freshwater and maritime zones, independent of geographical distance. Moreover, we show that the greatest proportion of genetic variation (31.5%) is explained by environmental differences. However, the potential for either adaptive or plastic divergence between demes is highly dependent upon the strength of migration and selection. Consequently, we highlight the utility of landscape genetics as a tool for hypothesis generation and experimental design, to identify focal populations and putative selection gradients, in order to distinguish between phenotypic plasticity and local adaptation.

Keywords: cline, *Gasterosteus aculeatus*, isolation by distance, parapatric, salinity

Received 10 March 2008; revision received 27 June 2008; accepted 2 July 2008

Introduction

From its earliest inceptions, population genetics has been a study of evolutionary processes. Indeed, the formative debates of Fisher and Wright ultimately centred on the relative importance of selection vs. drift, and the nature of genetic variance (i.e. additive vs. dominance and epistasis), underlying adaptive population divergence (Fisher 1930; Wright 1931; Crow 1987). This early work, although inspired by classical genetic experiments, was largely theoretical. It was not until observations of diversity at the nucleotide level, that population genetics shifted focus almost entirely towards a neutral theory of evolution (Kimura 1968). Advances in molecular biology have continued to revolutionized the field, providing researchers with a myriad and diverse toolkit of genetic markers with which to test and refine earlier theoretical results (Cavalli-Sforza 1998). From a practical perspective, the use of molecular markers has proven indispensable in defining heretofore unknown population structure, based largely on differentiation between demes due to neutral processes in the absence or reduction of gene flow (Waples & Gaggiotti 2006). Yet many questions remain, and modern population genetics finds itself in the midst of another technical revolution, this one lead by increasing analytical and computing power. The outcome should lead to a more comprehensive understanding of the spatio-ecological factors underlying population structure.

It is evident, almost taken for granted, that populations have structure in both time and space. Wright (1943) was

Correspondence: Louis Bernatchez, Fax: 1-418-656-7176; E-mail: louis.bernatchez@bio.ulaval.ca

the first to implicitly integrate spatial structure into population genetic theory, with subsequent research becoming more spatially explicit, ultimately leading to refinements of the now classic hypothesis of isolation by geographical distance (Slatkin 1993). Yet until recently, further geographical analyses of population genetic data remained spatially implicit, focusing primarily on post-hoc mapping of patterns of differentiation (Menozzi *et al*. 1978). However, more spatially explicit approaches to analyses of genetic differentiation may be useful in disentangling process from pattern (Latta 2006; Raeymaekers *et al*. 2008). Indeed, the recent integration of landscape ecology with population genetics, linking the interactions between spatial/landscape features and microevolutionary processes, has been championed as a new approach to ascertain the spatio-ecological factors that may promote or constrain divergence (Manel *et al*. 2003; Storfer *et al*. 2007). To date, this approach has been used do explore five major themes: (i) distinguishing historical and contemporary landscape influences on genetic variation; (ii) identifying barriers to gene flow; (iii) exploring source-sink dynamics in relation to habitat quality; (iv) defining spatial and temporal scales of gene flow in relation to landscape features; (v) hypothesis testing for ecological adaptation (Storfer *et al*. 2007). However, the full potential of landscape genetics as a tool for hypothesis generation may be underappreciated, particularly in regard to adaptive population divergence.

Views regarding the utility of landscape genetics for studying adaptive divergence have typically reflected earlier population genetics approaches to the problem, focusing on the identification of outlier loci to infer putative selection gradients (Manel *et al*. 2003; Latta 2006). But natural selection may also leave a genome-wide signature, not necessarily restricted to targeted loci (Charlesworth *et al*. 1997). Conversely, great care must be taken in such analyses and interpretation, given that spatial patterns in genetic variation may also be generated by neutral processes (Latta 2006); furthermore, it has been argued that molecular markers may be most reflective of historical gene flow and drift (McKay & Latta 2002). However, whereas patterns of isolation by geographical distance are most frequently associated with migration/drift equilibrium, that is, neutral processes (Hutchison & Templeton 1999), we might predict that increasing genetic divergence as a function of environmental differentiation may potentially be indicative of adaptive processes (Funk *et al*. 2006; Nosil *et al*. 2008). Therefore, concordant patterns of genetic and environmental divergence, independent of geographical distance, may be informative in both identifying populations warranting further empirical exploration, as well as generating hypotheses regarding any underlying ecological causes of their divergence.

The Saint Lawrence River estuary represents a unique natural environment in which to explore the dynamics between microevolutionary processes. This lower 540-km section of the Saint Lawrence River is highly influenced by tidal processes resulting in a gradient of various physicochemical and biological landscape features (Laprise & Dodson 1994; Vincent & Dodson 1999). Additionally, the estuary is characterized by relatively stable freshwater and saltwater zones located upstream and downstream, respectively, of a highly variable freshwater–saltwater transition zone, itself characterized by steep gradients of biological community structure, as well as diurnal variation in physicochemical traits (Vincent *et al*. 1996; Winkler *et al*. 2003). Consequently, environmental heterogeneity may be sufficient to lead to local adaptation, yet no physical barriers exist to prevent the potentially homogenizing effect of gene flow among locally adapted demes (Slatkin 1973; García-Ramos & Kirkpatrick 1997; Hendry *et al*. 2001; Lenormand 2002). Although surprisingly few such studies have been conducted within the estuary, some evidence suggests phenotypic divergence associated with environmental variation in populations of nematodes (Tita *et al*. 1999) and at least one fish species (Lecomte & Dodson 2004, 2005). However, three-spined sticklebacks (*Gasterosteus aculeatus*), a species abundant throughout the estuary, exhibit very little morphological differentiation across the entire range of environmental conditions encountered. For example, a cline of diminishing numbers of lateral plates, commonly classed in terms of armouring (i.e. full, partial and reduced), is typical for numerous estuarine populations, yet throughout the entire Saint Lawrence estuary, sticklebacks are uniquely fully plated (R. J. S. McCairns, unpublished data). This is in stark contrast to the diversity of forms seen in populations inhabiting similar environmental gradients throughout the species' global distribution. Indeed, the three-spined stickleback has emerged as a model organism for the study of contemporary evolution owing to its multiple, parallel examples of adaptive morphological divergence in the face of such environmental heterogeneity (Bell & Foster 1994; Schluter 1996; McKinnon & Rundle 2002). Therefore, the apparent lack of obvious, phenotypic diversity in such an inherently variable model organism inhabiting this heterogeneous environment poses an interesting opportunity for research into the nature of phenotypic variability vis-à-vis spatial and temporal environmental heterogeneity, the relative roles of phenotypic plasticity vs. local adaptation, and the antagonistic effects of gene flow vs. selection. As a first step towards addressing these questions, we used a landscape genetics approach to characterize genetic and environmental variation associated with stickleback populations inhabiting the Saint Lawrence River estuary. Our major objectives were threefold: (i) to ascertain the population structure of sticklebacks inhabiting the estuary, and to determine the spatial location of any barriers to gene flow; (ii) to evaluate environmental differences between the populations' geographical regions, and to explore putative relationships between environmental

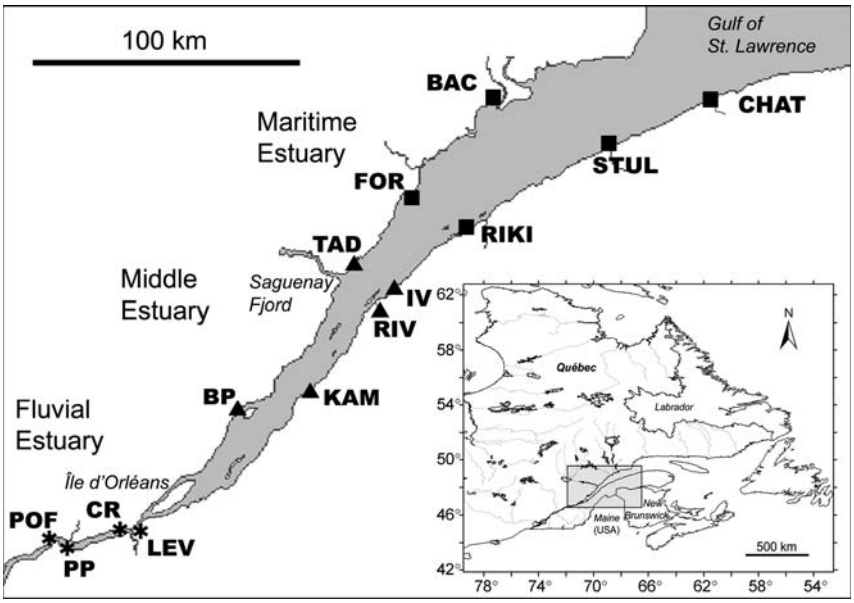

Fig. 1 Sampling locations in relation to the estuarine zones of the Saint Lawrence River. Sites in the fluvial estuary are indicated by asterisks (✱), middle estuary sites with closed triangles (▲), and closed squares (■) denote sites within the maritime estuary. Details regarding the number of individuals sampled by year are contained in Table 1.

and genetic variance; (iii) to estimate the demographic and selective parameters, representative of those experienced by sticklebacks of the Saint Lawrence estuary, that may potentially promote or inhibit adaptive divergence between demes.

Methods

Study area and sampling

This study was conducted over a range of approximately 350 km throughout the Saint Lawrence River estuary in Québec, Canada (Fig. 1). The Saint Lawrence is among the largest rivers in North America, draining an area of approximately 1 320 000 km² (El-Sabh & Silverberg 1990). Its estuary is divided into three sections, each characterized by their biological, physiochemical and tidal properties (Vincent & Dodson 1999). The fluvial estuary (a.k.a. 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 (a.k.a. lower estuary) is a 230-km stretch ultimately discharging into the Gulf of Saint Lawrence. Hydraulic dynamics shift from tide-dominated to wave-dominated, and the biological and physiochemical properties more closely resemble those of the marine environment, with salinity ranging from 25 to 35 psu.

A total of 1865 adult sticklebacks were collected, using seines and minnow traps, from tidal marshes and tributary streams during the breeding season (May–June). In total, 14 sites throughout the Saint Lawrence estuary were sampled from 3 years (Table 1). Ten sites were originally sampled in 2003. In 2004, replicate samples were collected from eight

Table 1 Sampling sites, as per Fig. 1, and the number of individuals collected (*N*) during each year of sampling. Average observed (H_O) and expected (H_E) heterozygosities are presented for each sample, as calculated using ARLEQUIN (Excoffier *et al*. 2005). Inbreeding coefficients (F_{IS}), averaged over all loci, were calculated using FSTAT (Goudet 1995), and evaluated against 3960 randomly permuted distributions

Site	Year	*N*	H_O	H_E	F_{IS}	(*P* Value)
POF	2003	55	0.639	0.645	0.010	(0.296)
PP	2003	50	0.610	0.647	0.058	(0.001)
CR	2003	55	0.669	0.675	0.010	(0.276)
	2004	120	0.664	0.662	–0.004	(0.629)
LEV	2003	54	0.646	0.643	–0.004	(0.581)
	2004	109	0.649	0.654	0.007	(0.295)
BP	2004	99	0.629	0.659	0.046	(0.002)
KAM	2004	116	0.635	0.653	0.030	(0.014)
RIV	2004	124	0.663	0.668	0.008	(0.270)
IV	2003	40	0.636	0.661	0.035	(0.077)
	2004	120	0.631	0.656	0.038	(0.002)
TAD	2003	53	0.621	0.660	0.061	(0.002)
	2004	124	0.630	0.642	0.019	(0.068)
FOR	2003	64	0.635	0.659	0.037	(0.026)
	2004	124	0.647	0.656	0.014	(0.155)
RIKI	2003	60	0.657	0.660	0.007	(0.372)
	2004	124	0.632	0.661	0.045	(0.001)
BAC	2003	53	0.649	0.655	0.008	(0.317)
	2004	122	0.638	0.654	0.024	(0.035)
STUL	2007	47	0.595	0.661	0.101	(< 0.001)
CHAT	2003	56	0.635	0.660	0.039	(0.020)
	2004	96	0.629	0.644	0.026	(0.044)

of the original sites and three additional sites within the middle estuary. An additional site within the maritime estuary (STUL) was sampled in 2007.

Molecular markers

The right pelvic fin of each captured fish was excised immediately and preserved in 95% ethanol for DNA extraction. Genomic DNA was obtained from tissue samples by salt extraction (Aljanabi & Martinez 1997). Ten dinucleotide microsatellite markers were amplified by simplex polymerase chain reaction (PCR) using fluorescently labelled primers and reaction conditions as specified by the original authors (Largiadèr *et al*. 1999; Peichel *et al*. 2001). Resultant PCR products were uniquely diluted to obtain similar peak intensities when run in multiplexes of five loci on an ABI 3100 capillary sequencer (Applied Biosystems). Electropherograms were scored using GENEMARKER (SoftGenetics), and allelic sizes classified into bins using the FLEXIBIN algorithm (Amos *et al*. 2007).

All electropherograms were scored twice, and their independent bin scores compared to minimize genotyping errors. In the event of a discrepancy between scores, an individual's electropherograms were scored a third time, with the final genotype assessed as the majority score. To estimate total errors within the molecular data set, an independent, random replicate sample of approximately 7% of individuals (2003: $n = 32$; 2004: $n = 96$) was reprocessed entirely, from DNA extraction to genotyping. Per locus error was estimated as the proportion of differing scores between the replicate data sets (Pompanon *et al*. 2005). Estimated error rates for nine of the 10 loci, listed in Table S1, Supplementary material, were deemed acceptable (< 5%). Conversely, one locus (Stn 70) exhibited an excessive error rate, and as such, was removed from the data set before subsequent analyses.

Analysis of population structure

To ascertain the relative degree of temporal stability in population structure, we conducted an analysis of molecular variance (AMOVA) using the software ARLEQUIN 3.11 (Excoffier *et al*. 1992). The first hierarchical level of the analysis consisted of the eight localities for which replicate temporal samples were collected, and the second level consisted of the two sampling periods within each site. We further explored the relative degree of temporal vs. geographical differentiation among sites by constructing a population tree. We calculated 1000 bootstrap matrices of chord distances between each sample (Cavalli-Sforza & Edwards 1967), and constructed a consensus unrooted neighbour-joining tree from the bootstrapped data using PHYLIP 3.66 (Felsenstein 2005). We also conducted a second phenetic analysis employing the same procedure, but using a data set consisting of all sites regardless of sampling year and with replicate samples pooled.

Given the strong signal of temporal stability observed (see Results), replicate temporal samples were pooled for subsequent analyses, and total population structure was inferred using all individuals regardless of sampling year. We began with no a priori assumptions regarding population structure, but rather used georeferenced multilocus genotype data to infer population groupings, as well as locations of genetic discontinuities, under a Bayesian model framework. This model, available as the GENELAND library for R (Team 2007), contains a full suite of commands for varying model priors, as well as for checking model convergence, and postprocessing output for graphical display (Guillot *et al*. 2005b; Estoup *et al*. 2007). When no geographical references are incorporated, the GENELAND model is equivalent to the clustering algorithm without admixture employed in the program, STRUCTURE (Pritchard *et al*. 2000). However, by incorporating geographical data into the analysis, GENELAND is able to detect both subtle population structure, as well as infer the probable locations of genetic transition zones, information that is essential for analysis under a landscape genetics framework (Guillot *et al*. 2005a). GENELAND does this by incorporating geographical information as a weak prior, with the assumption that most populations exhibit some degree of spatial structuring, and that the joint probability of any two individuals belonging to the same population decreases with the geographical distance between them. However, error within the spatial coordinates is incorporated into the model, thus, permitting individuals to be fitted to multiple populations for posterior likelihood testing. Ultimately, the relative positions of different population clusters, and their geographical boundaries, are variables estimated from the posterior likelihood distributions (Guillot *et al*. 2005a). We began with a model assuming no error in spatial coordinates, and allele frequencies following independent Dirichlet distributions. We set a burn-in period of 100 000 iterations, followed by 100 000 iterations from which each 100th observation was sampled from the Markov chain to reduce autocorrelation of parameter estimates. Thus, parameter estimates were based on a distribution of 1000 independent simulations. We verified model convergence by plotting the Markov chain Monte Carlo (MCMC) trace files, and then estimated the number (k) of genetic clusters. Models were then re-run assuming k clusters to infer the probability of individual membership to each group. We compared the results of five separate analyses of the same data set for consistency of cluster estimation. We then proceeded with another series of separate MCMC runs, as previously described, but including a degree of error for the spatial coordinates to reflect connectivity of sites identified as clusters in the previous analysis (delta.coord = 0.5). Each of five independent runs was used to identify the number of genetic

clusters, re-run to infer probability of individual membership within each cluster, and post-processed for graphical display.

In order to estimate the proportion of genetic variance explained by the groupings identified by GENELAND, we performed a nested AMOVA by variance component analysis of the multilocus genotypes (Yang 1998). Individuals were grouped by sampling sites which were nested within two genetic landscapes (see Results). Variance component analysis was performed to partition the proportion of total genetic variance explained by the respective grouping levels, and to estimate hierarchical *F* coefficients. The significance of each *F*-statistic was evaluated by permutation tests, based on 10 000 randomizations. All analyses were performed using the HIERFSTAT package implemented in the R computing language (Goudet 2005, 2006). We contrasted this with the proportion of variance explained based on the natural division of the Saint Lawrence estuary into distinct zones by performing additional variance component analysis and hierarchical AMOVA nesting sites within their respective estuarine zones (see Fig. 1).

We also analysed the multilocus data set under a model of isolation by distance (IBD) to estimate the proportion of genetic divergence within the estuary as a function of geographical distance between sites (Rousset 1997). As a measure of genetic differentiation, we calculated the ratio of $F_{ST}/(1 - F_{ST})$ for each pairwise combination of sites using ARLEQUIN (Excoffier *et al*. 2005). We then calculated a matrix of geographical distances among sites by measuring the shortest shoreline distances between each pairwise combination of sites. To assess the significance of the IBD model, we estimated the correlation between matrices of genetic and geographical distance using a Mantel test implemented in the VEGAN package for R (Oksanen *et al*. 2007). The significance of the matrix correlation was evaluated by permutation test, based on 10 000 randomizations. Finally, in order to disentangle the relative importance of IBD vs. differentiation between genetic landscapes or estuary zones, we calculated partial correlations to account for multicollinearity. Grouping distance was coded as a dummy variable in which pairwise comparisons between sites are coded as '0' if they belong to the same region identified by GENELAND analysis or estuary zone, and '1' if the sites correspond to different groups, an approach that has previously been used to associate phenetic groupings with genetic variation, also based on partial Mantel tests (Douglas *et al*. 1999). Two partial Mantel correlations were calculated: one testing the proportion of variance in genetic distances among sites explained by geographical distance, controlling for covariation with group membership, another testing the proportion of genetic variance explained by landscape/estuary groupings independent of geographical distance. The significance of each partial correlation was tested against 10 000 permutations, as implemented in VEGAN.

Quantification of environmental variation and its association with genetic differentiation

We collected data for a number of environmental parameters from both the spawning/nursery habitat as well as the Saint Lawrence estuary proper. In 2005, during the period in which juvenile sticklebacks are retained within the tidal marshes of the Saint Lawrence, that is, between the vernal and autumnal floods (Picard *et al*. 1990), we placed max–min thermometers in known spawning/nursery habitats at each sampling site. At the beginning and end of this period, we also measured dissolved oxygen concentration and total salinity *in situ*. Thus, for each site, we had direct measures of minimum and maximum water temperatures experienced by juvenile fish. Given that dissolved oxygen concentrations are likely to decrease throughout the season due to organic decay, although likely with substantial diurnal variation throughout (Reebs *et al*. 1984), and that salinity is likely to increase due to evaporation, measures taken at the beginning and end of the season should represent reasonable proxies for seasonal biological oxygen demand as well as maximum and minimum salinity.

Salinity and water temperature data from the estuary are archived by the Department of Fisheries and Oceans Canada (www.osl.gc.ca/sgdo/en/info-donnees/citation.html). Data consisted of georeferenced surface temperature and salinity measures sampled with shipboard thermosalinographs installed on commercial vessels, in addition to temperature and salinity time series sampled by coastal thermographs. We screened available data for dates between the autumnal and vernal equinoxes, thus, roughly corresponding to the period during which sticklebacks reside within the estuary. Data were restricted to measures taken within 10 m of the surface, and were averaged over all years by month. However, we lacked sufficient data for the months of January and February; consequently, data were pooled and averaged for this period. As such, the data set consisted of georeferenced salinity measures averaged over the entire period (given insignificant intra-annual variation; data not shown), and georeferenced mean water temperatures for the months of September through December, a winter estimate (January–February average), and for the months of March and April. We then employed a spatial interpolation technique, kriging (Cressie 1993), implemented in the FIELDS package for R (Nychka 2007), to estimate mean salinity and monthly temperatures in the waters immediately offshore from each sampling site.

To determine if the regions identified by GENELAND were ecologically divergent, we performed a multivariate analysis of variance (MANOVA), followed by a canonical variate analysis (CVA) to identify the variables contributing most to the differentiation between regions, using the PAST software package (Hammer *et al*. 2001). Data for the spawning/juvenile habitats and the estuary were analysed separately

because the total number of variables exceeded the number of sites per region. We also performed a *k*-means cluster analysis ($k = 2$) based on minimized centroid differences using the same data set of all environmental variables (Venables & Ripley 2002). Our rationale was to determine if sites clustered into the same groups, based strictly on environmental data, and to estimate the proportion of environmental variance explained by this clustering, as determined by the eigenvectors of a principal components analysis (PCA) on the correlation matrices of all environmental variables.

Finally, to determine if genetic divergence among sites was related to differentiation in their ecological characteristics, we modelled pairwise genetic distance as a function of both geographical and ecological distances using the permutation-based multiple regression module in FSTAT (Goudet 1995). We defined a general ecological state for each site based on their respective scores from the first eigenvector (PC1) of the PCA. Ecological distance among sites was then calculated based on a matrix of Euclidean distances of their PC1 scores. We similarly calculated distance matrices for the full suite of environmental variables within both the juvenile/spawning and estuarine habitats. Model building followed a forward stepwise procedure, entering variables one at a time, beginning with geographical distance. The significance of subsequent variable additions was evaluated based on their slopes (β), and the resultant change in the overall model's multiple correlation (R^2). Model building was continued until further variable additions yielded no significant improvement in model fit. However, the unique effect size of all variables was determined by estimating their squared semipartial correlations, controlling for covariation with all significant factors retained within the final multiple regression models.

Modelling the potential for adaptive divergence between landscapes

We employed a simple, qualitative model in order to predict how demographic processes and the general spatial characteristics of the genetic landscapes may affect the potential for adaptive divergence within the estuary. This approach has been previously used to define the necessary conditions leading to adaptive divergence between salmonid populations (Adkison 1995; Hansen *et al*. 2002), and seemed amenable for use in the estuary. The model is based on a numerical approximation of Slatkin's (1973) characteristic length scale of variation in gene frequency (l_c), essentially a metric defining a minimal cline distance under which populations cannot respond to environmental variation. Predictions of genetic homogeneity (*H*) vs. differentiation due to random drift (*R*) or adaptive divergence (*A*) are based on derived variables describing the ratio of migration to drift (β), and *k*, the ratio of the geographical scale at which selection favours a given allele relative to the critical cline scale determined by gene flow and selection (Nagylaki & Lucier 1980). Nagylaki & Lucier (1980) demonstrated that β can be approximated as

$$2N_e\sqrt{s2m},$$

and *k* can be defined as j/l_c, where

$$l_c = \sqrt{m/(2s)}$$

and *j* is the number of local subpopulations, or geographical scale, over which selection (*s*) is the same. Here, we took a stringent view that each landscape represents a single population with differential selection acting in each ($j = 1$). To parameterize the model, we used a variety of approaches to obtain a range of estimates of effective population size (N_e) and migration rate (*m*) likely to capture the actual demographic processes occurring within the estuary. These values were used to infer the conditions likely to lead to adaptive divergence (*A*: $\beta > 1.1; k > 1.1$), random differentiation (*R*: $\beta < 1$), or genetic homogeneity (*H*: $\beta > 1.1; k < 1$).

Initially we assumed that an upper bound on *m* may be estimated by dispersal between landscape groups. To estimate contemporary rates of dispersal, we used multilocus genotype data to assign each individual to a site of origin using GENECLASS 2 (Piry *et al*. 2004). Individuals were initially grouped according to their original sampling sites, with the exception of certain pairs which were grouped together as single sites due to their geographical proximity and extremely low levels of genetic differentiation ($F_{ST} < 0.0001$). Pairs grouped as single sites were: POF-PP, CR-LEV, IV-RIV and STUL-CHAT (see Fig. 1). For each individual, the likelihood of site membership was calculated based on Rannala & Mountain's (1997) Bayesian criterion. Probability of assignment was based on Monte Carlo re-sampling employing 10 000 simulated individuals (Paetkau *et al*. 2004). Individuals were identified as migrants, and assigned to their location of origin, based on significant likelihood-ratio tests comparing the likelihood of assignment to their location of capture with the location yielding the maximum likelihood of assignment (L_{HOME}/L_{MAX}). Dispersal rates among sites were then estimated as the proportion of individuals assigned to a given location, yet originally captured from a different site. Confidence intervals (95% CI) were estimated from a bootstrap distribution (1000 iterations) of mean dispersal rate for each group. Finally, we repeated the same analysis, but grouping individuals by landscape, rather than site.

We used GENEPOP to estimate the number of effective migrants (N_em) based on the private allele method (Raymond & Rousset 1995). For the subset of temporally replicated data, thus, corresponding to one intergenerational period (Picard *et al*. 1990), we estimated N_e and *m* jointly under an approximation of the two finite demes model, as implemented in the program MLNE (Wang & Whitlock 2003). We also used a coalescent approach to simultaneously

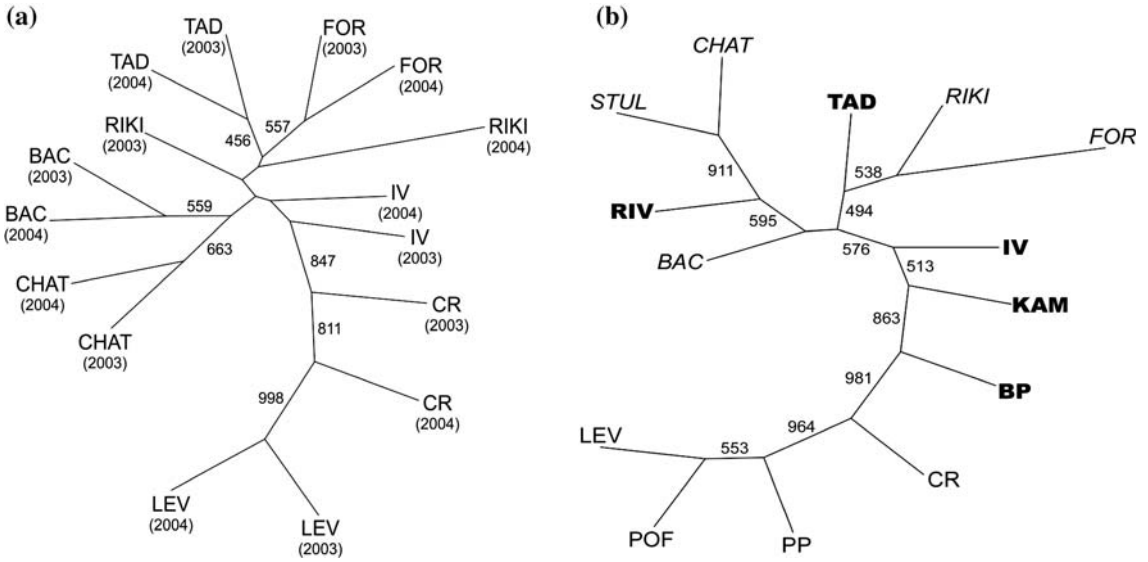

Fig. 2 Unrooted neighbour-joining trees comparing sites with temporal replication (a) and all sampling sites (b). Inter-site divergence was calculated based on Cavali-Sforza and Edwards' chord distance (D_C). Numbers correspond to the consensus value of the adjacent node, based on 1000 bootstrap matrices. In Fig. 2 (b), sites within the fluvial estuary are labelled with plain text (e.g. POF), middle estuary sites with bold text (e.g. **IV**), and sites from the maritime estuary are in italics (e.g. *BAC*).

estimate N_em and m (Beerli & Felsenstein 2001). MIGRATE was parameterized following the author's recommendations, beginning with an initial model assuming mutation approximated by Brownian motion, and initial parameters based on F_{ST} estimates from the data (Beerli 2006). This model was run over 10 short chains, with parameter estimates derived from three long chains and 100 000 sampled genealogies. The resultant estimates were then used to parameterize a second series of runs based on increased short chains ($n = 50$), with final parameter estimates averaged over three independent runs from five long-chains. We estimated migration rate from the MIGRATE output assuming mutation rates ($\mu = 10^{-4}$) corresponding to the average reported for microsatellite markers in fishes (Shimoda *et al*. 1999; Avise 2004; Yue *et al*. 2007). N_e values were estimated iteratively by dividing N_em estimates over the range of migration rate estimates obtained.

Results

Analysis of replicate samples suggested a strong signal of temporal stability in genetic composition within sites. No pairwise differences between replicate temporal samples were evaluated as significant (all $F_{ST} < 0.001$; all $P > 0.184$). Furthermore, results of the AMOVA revealed no significant differences between temporal replicates within each site ($F_{SC} < 0.001$; $P = 0.996$), although differentiation among sampling sites was significant ($F_{CT} = 0.006$; $P < 0.001$). With the exception of one site (RIKI), all temporal replicates tended to cluster together, with the majority of nodes exhibiting bootstrap support greater than 50% majority rule (Fig. 2a). In the case of RIKI, temporal samples tended to cluster with other sites in close geographical proximity (< 50 km), although bootstrap support for this group was generally weak (221/1000 replicates; data not presented). The two sites sampled within the fluvial estuary (LEV, CR) not only exhibited a high degree of temporal stability, but also formed a branch apart from the other samples from the middle and maritime estuaries (847/1000 replicates), thus, providing the first hint of population divergence within the estuary. Subsequent phenetic analysis, including all sampling sites with temporal replicates pooled, further suggested population division along the transitional boundary between the freshwater and saltwater estuarine zones. Sites within the fluvial estuary formed a distinct group apart from sites of the middle and maritime estuaries, with a highly significant degree of bootstrap support for this node (981/1000 replicates; Fig. 2b). Moreover, the general order of branching further suggested a division of sites between the middle and maritime estuaries, although with the exception of sites RIV and TAD, and with a much weaker level of significance (576/1000 replicates).

Estimates from five separate MCMC model runs with no error in spatial coordinates consistently identified two genetic clusters, with each run incorporating the same sites into the respective clusters: sites POF, PP, CR, LEV and BP comprised one group, herein referred to as the freshwater landscape (FW), whereas the remaining populations grouped into the second cluster, the maritime landscape (SW). When an error in spatial coordinates of 0.5 degrees was introduced into the model, three of the independent runs identified the

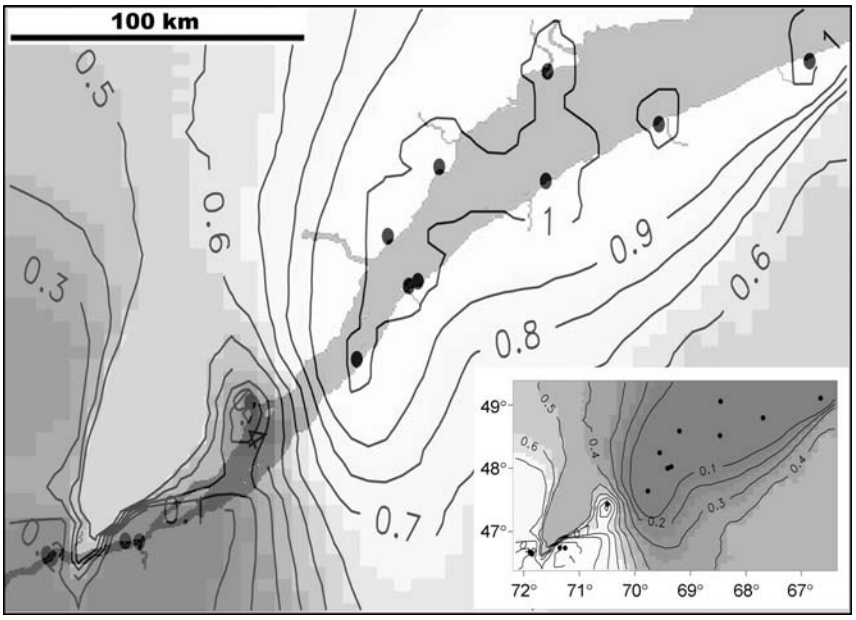

Fig. 3 Probability isoclines denoting the extent of the maritime (SW) genetic landscape. For ease of interpretation, GENELAND output has been cropped, re-scaled, and superimposed over the original site map (Fig. 1). Regions with the greatest probability of inclusion are indicated by white, whereas diminishing probability of inclusion is proportional to the degree of shading. The inverse is true for sites comprising the freshwater landscape (FW). Original GENELAND output denoting isoclines of the FW landscape are presented as an insert in the bottom right corner.

Table 2 Hierarchical AMOVAS partitioning genetic variance into the main effects of genetic landscapes identified by GENELAND analysis (Fig. 3), or estuarine zones (fluvial, middle and maritime; Fig. 1). Effects nested within each main effect are offset and listed in square brackets. Variance components are presented for each main effect. *P* values for each *F*-statistic are derived from permutation tests based on 10 000 randomizations

Genetic landscapes				Estuarine zones			
Effect	Var. comp.	*F*-statistic	*P* value	Effect	Var. comp.	*F*-statistic	*P* value
Landscape	0.031	$F_{LT} = 0.0051$	< 0.001	Estuary	0.025	$F_{ET} = 0.0042$	0.012
[site]	—	$F_{SL} = 0.0036$	< 0.001	[site]	—	$F_{SE} = 0.0029$	< 0.001
[ind.]	—	$F_{IL} = 0.0288$	—	[ind.]	—	$F_{IE} = 0.0284$	—
Site	0.022	$F_{ST} = 0.0088$	< 0.001	Site	0.022	$F_{ST} = 0.0071$	< 0.001
[ind.]	—	$F_{IS} = 0.0253$	—	[ind.]	—	$F_{IS} = 0.0253$	—
Individual	0.149	$F_{IT} = 0.0338$	—	Individual	0.150	$F_{IT} = 0.0321$	—

same groups, whereas a third genetic cluster was initially identified in two runs. However, postprocessing revealed the potential third clusters were artefacts as no sampled sites, nor individuals, were assigned to these populations with a significant level of probability (< 0.01). Moreover, individuals were consistently assigned to the same clusters across all independent MCMC runs with the same level of probability (± 0.04). As was the case for models without error in spatial coordinates, sites POF, PP, CR, LEV and BP formed a distinct genetic cluster (Fig. 3). Individuals from the four sites within the fluvial estuary were assigned to this cluster with high probability (≥ 0.99); however, probability of assignment of individuals from the middle estuary site (BP) was lower, although still within a reasonable level of significance (0.87 averaged over 5 MCMC runs). The remaining sites consistently formed the second cluster, with all individual assignment probabilities greater than 0.99 (Fig. 3).

Results of the nested AMOVA provided weak, but statistically significant, evidence for population subdivision by genetic landscapes (Table 2: $F_{LT} = 0.005$; $P < 0.001$). Population structuring by estuarine zones was similarly supported (Table 2: $F_{ET} = 0.004$; $P = 0.012$). Although genetic variation due to site differences, both total (F_{ST}) and nested within groups ($F_{SL} = 0.004$; $F_{SE} = 0.003$), appeared to increase when sites were clustered into genetic landscapes rather than estuarine zones, so too did the level of differentiation between groups ($F_{LT} = 0.005$ vs. $F_{ET} = 0.004$). Moreover, sites clustered into genetic landscapes captured a slightly greater proportion of genetic variance than groupings by estuary zones (15.2% vs. 12.1%, respectively). Similar patterns were observed in the Mantel correlations between genetic distance and grouping distances (Table 3): grouping by genetic landscapes accounted for up to 35.7% of genetic differences among sites, whereas grouping by estuary zones captured

Table 3 Matrix correlations describing the relative strength of associations between genetic distance and geographical distance, or site groupings within the genetic landscapes identified by GENELAND analysis (Fig. 3) vs. estuary zones. Grouping distances were coded as dummy variables in which pairwise comparisons between sites are coded as '0' if they belong to the same landscape/estuary zone, and '1' if the sites correspond to different groups. Partial Mantel correlations are taken to test the proportion of variance in genetic distances among sites explained by geographical distance, controlling for covariation with group membership vs. the proportion of genetic variance explained by landscape/estuary zone independent of geographical distance. Covariates for each partial Mantel test are listed in square brackets. All *P* values are based on 10 000 permutations

Association	Mantel *R*	*P* Value
$F_{ST}/(1-F_{ST})$ ~ Geographical distance	0.344	0.010
$F_{ST}/(1-F_{ST})$ ~ Landscape grouping	0.597	0.001
$F_{ST}/(1-F_{ST})$ ~ Estuary zone	0.401	< 0.001
	Partial *R*	*P* value
$F_{ST}/(1-F_{ST})$ ~ Geo. dist. [landscape]	– 0.124	0.774
$F_{ST}/(1-F_{ST})$ ~ Geo. dist. [estuary]	0.131	0.189
$F_{ST}/(1-F_{ST})$ ~ Landscape [geo. dist.]	0.530	< 0.001
$F_{ST}/(1-F_{ST})$ ~ Estuary [geo. dist.]	0.254	0.020

only 16.1% of this variation. Differences between groupings were even more pronounced when controlling for covariance with geographical distance (partial $R = 0.530$ vs. 0.254). In comparison, the isolation-by-distance model alone did not explain a significantly greater proportion of variance than either grouping, accounting for only 11.9% of genetic differentiation among sites (Table 3: Mantel $R = 0.344$). Moreover, the effect of geographical distance was not significant when covariance with either landscape or estuary groupings were included in the test (Table 3).

Environmental variation and its relationship with genetic differentiation

MANOVA results suggested that the two identified genetic landscapes exhibited only weak, marginally nonsignificant, differences between environmental characteristics of their juvenile/spawning habitats (Wilk's $\Lambda = 0.259$; $P = 0.070$). In contrast, the estuary conditions within these regions were significantly different (Wilk's $\Lambda = 0.001$; $P < 0.001$). Subsequent canonical analyses indicated that salinity was the principal variable associated with any differentiation between juvenile/spawning habitats; however, CV1 loadings for variables within the estuary suggested that temperature differences were also significant. Data reduction by PCA indicated that 61% of variation in environmental data could be explained by one eigenvector (PC1), loading differentially between salinity variables and estuary temperatures (Fig. 4b). Moreover, based on environmental data, sampling sites independently clustered into the same two groups identified by GENELAND (Fig. 4a). The majority of sites' centroid differences with their respective clusters were sufficiently small to suggest membership between 0.83 and 0.93 probability. In contrast, the membership coefficient for site BP clustering with other sites belonging to the FW landscape was significantly less (0.55).

Simple Mantel correlations indicated significant correspondence between matrices of genetic distances and

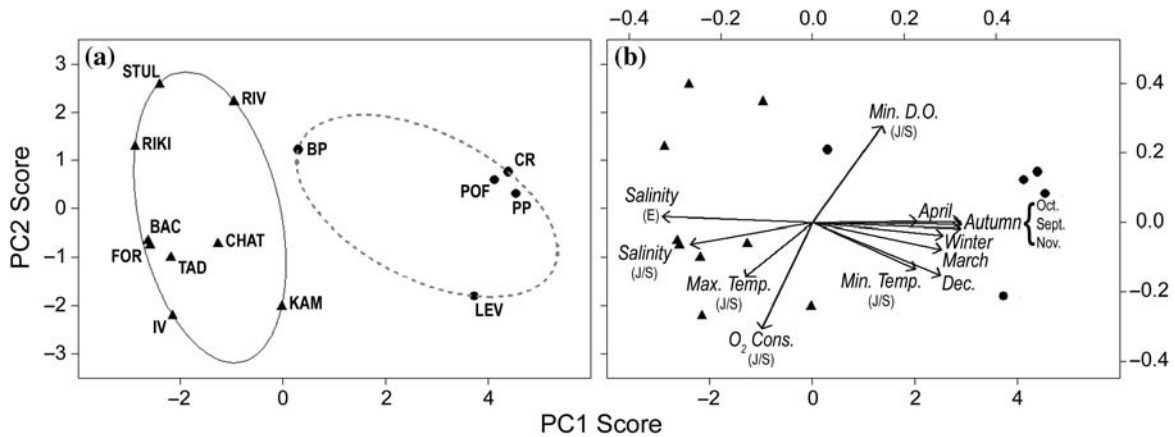

Fig. 4 Site clustering based on minimized centroid distances of environmental variables comprising the reduced data set (a). The plotting coordinate system corresponds to a principal component analysis in which the abscissa describes a vector based primarily on salinity differences and estuary temperature, and the ordinate captures variation in dissolved oxygen. Sites belonging to the FW landscape are labelled with closed circles (●), whereas triangles (▲) denote sites within the SW landscape. Ellipses encompass the sites clustered by minimized centroid distances. A bi-plot summarizing the variables comprising the first two principal components is also presented (b). Positions of sites based on their respective PC1 and PC2 scores is the same as in A; however, additional axes (upper and right) describe the vector loadings of the respective variables.

Model	Model evaluation		Effect size
Term	β (*P* value)	R^2 (*P* value)	Part r^2 (*P* value)
Reduced data set (PC1)			
(1) distance	–0.000019 (0.045)	0.119 (0.001)	0.043 (0.048)
(2) ecological score	0.001618 (< 0.001)	0.434 (< 0.001)	0.315 (< 0.001)
Juvenile and spawning habitat (J/S)			
(1) distance	0.000014 (0.030)	0.119 (0.001)	0.051 (0.031)
(2) salinity	0.000151 (0.002)	0.229 (0.002)	0.110 (0.002)
(3a) min. DO	–0.000055 (0.806)	0.229 (0.806)	< 0.001 (0.806)
(3b) min. temp.	–0.000012 (0.964)	0.229 (0.963)	< 0.001 (0.963)
(3c) max. temp.	–0.000148 (0.364)	0.238 (0.364)	0.009 (0.364)
Estuarine environment (E)			
(1) distance	–0.000023 (0.027)	0.119 (0.001)	0.053 (0.030)
(2) salinity	0.000404 (< 0.001)	0.382 (< 0.001)	0.263 (< 0.001)
(3a) September temp.	0.000134 (0.715)	0.383 (0.714)	0.002 (0.714)
(3b) October temp.	0.000674 (0.079)	0.417 (0.079)	0.035 (0.079)
(3c) November temp.	0.000244 (0.876)	0.382 (0.876)	< 0.001 (0.876)
(3d) December temp.	0.001522 (0.097)	0.412 (0.097)	0.030 (0.097)
(3e) winter temp.	0.003096 (0.113)	0.409 (0.113)	0.028 (0.113)
(3f) March temp.	–0.000847 (0.620)	0.384 (0.620)	0.003 (0.620)
(3g) April temp.	–0.000167 (0.781)	0.382 (0.781)	< 0.001 (0.781)
Salinity and temperature			
(1) distance	–0.000024 (0.031)	0.119 (0.001)	0.053 (0.027)
(2) salinity (J/S)	0.000105 (0.034)	0.229 (0.001)	0.050 (0.030)
(3) salinity (E)	0.000364 (< 0.001)	0.432 (< 0.001)	0.203 (< 0.001)
(4a) October temp.	0.000475 (0.239)	0.448 (0.239)	0.016 (0.239)
(4b) December temp.	0.001380 (0.135)	0.457 (0.135)	0.025 (0.135)
(4c) winter temp.	0.003191 (0.100)	0.461 (0.100)	0.030 (0.100)

Table 4 Modelling pairwise genetic distance as a function of both geographical and ecological distances among sites. Models presented include an overall ecological score for each site, based on the first principal component of the total data set, in addition to models exploring variables within both the juvenile/spawning and estuarine habitats, and overall effects of salinity and coldwater temperature. Model building follows a forward stepwise multiple regression procedure in which variables are added at each step, with step numbers indicated in parentheses. The significance of variable addition is evaluated based on its slope (β), and the resultant change in the overall model's multiple correlation (R^2). Model building is continued until further variable addition yields no significant improvement in model fit. Nonsignificant model steps are enumerated in parentheses, with the step number followed by a letter (e.g. 3a). The unique effect size of each variable is defined by the squared semipartial correlation (part r^2). All *P* values are based on 5000 permutations, as calculated using FSTAT (Goudet 1995)

differences in PC1 scores among sites (Mantel $R = 0.625$; $P < 0.001$). However, neither PC2 (Mantel $R = -0.124$; $P = 0.829$), nor PC3 (Mantel $R = 0.035$; $P = 0.404$), were significantly correlated with genetic divergence. Consequently, we modelled general ecological score as a function of PC1 only. Overall model fit incorporating the effects of geographical distance and general environmental differences, based on data reduction by PCA, suggested a strong relationship between spatio-ecological and genetic divergence, potentially explaining up to 43.4% of total genetic variance (Table 4). Moreover, the unique effect of geographical distance was only marginally significant (part $r^2 = 0.043$; $P = 0.048$), whereas the unique effect of ecological differences accounted for 31.5% of variation in genetic divergence (part $r^2 = 0.315$; $P < 0.001$). Subsequent analyses within the juvenile/spawning and estuarine habitats, respectively, suggested that salinity differences were significantly associated with genetic differentiation, whereas temperature variation was not (summarized in Table 4).

Modelling adaptive divergence

Contemporary rates of dispersal, based on proportions of first-generation migrants identified using GENECLASS from individuals classified by their respective genetic landscapes (i.e. FW or SW), were estimated to be 0.042 into FW and 0.058 into SW. Estimates from which individuals were grouped according to site of origin were significantly lower: dispersal into the FW landscape was estimated to be 0.027 (95% CI = 0.020–0.033), whereas dispersal from FW into SW was 0.032 (95% CI = 0.016–0.048). GENEPOP analysis of the multilocus data set yielded an overall N_em value of 20.6, whereas the averaged estimates obtained from MIGRATE suggested 14.2 and 74 effective migrants from the FW and SW landscapes, respectively. Additionally, migration estimates (m) calculated from the MIGRATE output were 0.001 for migrants from FW into SW, and 0.004 into the FW landscape. Based on the subset of temporally replicated data analysed with MLNE, the overall maximum likelihood estimate of m was 0.046 (95% CI = 0.008–0.107), and effective population size (N_e) was estimated to be 371 (95% CI = 190–1861). Given varying m estimates, from 0.001 to 0.05, N_e values calculated from the GENEPOP N_em estimate ranged from 412 to approximately 20 000. When based strictly on their respective values of N_em and m obtained via MIGRATE, indirect estimates of N_e were calculated as 9892 for the FW landscape and 19 566 for SW. Consequently, we parameterized the model of adaptive vs. plastic divergence to include scenarios of N_e ranging

Table 5 Predicted values of effective population size (N_e), migration rate (m), and strength of selection (s) likely to lead to adaptive divergence (A) or genetic homogenization (H) between FW and SW landscapes. Predictions are based on the model of Nagylaki & Lucier (1980)

N_e	m	s = 0.001	s = 0.01	s = 0.1
500	0.001	A	A	A
1 000	0.001	A	A	A
20 000	0.001	A	A	A
500	0.002–0.01	H	A	A
1 000	0.002–0.01	H	A	A
20 000	0.002–0.01	H	A	A
500	0.02–0.1	H	H	A
1 000	0.02–0.1	H	H	A
20 000	0.02–0.1	H	H	A

from 500 to 20 000 (Table 5). Model parameterization also included migration rates spanning the range of estimates obtained.

Predictions of adaptive divergence appeared most heavily influenced by migration rate and intensity of selection (s). In contrast, model predictions were not influenced by effective population size. At the lowest predicted rate of migration (m = 0.001), adaptive divergence was predicted regardless of selection intensity. Conversely, at the highest level of selection (s = 0.1), equivalent to a 10% fitness advantage to favoured genotypes, adaptive divergence was predicted at all levels of gene flow (Table 5). However, at intermediate combinations of migration rate and/or selection intensities, genetic homogeneity remained a possibility. Interestingly, random differentiation was not predicted, based on the given demographic parameterization.

Discussion

We have identified two distinct genetic groups of three-spined sticklebacks within the Saint Lawrence River estuary. Although the low level of genetic divergence between these groups may warrant a cautious interpretation regarding the significance of this division (F_{SL} = 0.005), we suggest that there is sufficient evidence, both based on our results and observation from other systems, to conclude that some degree of true population structuring is inherent within the estuary. First, we analysed microsatellite data using both difference-based and individual multilocus genotype approaches, each providing mutually corroborating lines of evidence of population structuring. Moreover, these separate lines of evidence converge upon the same conclusion of two distinct genetic clusters within the estuary. The temporal stability revealed in Fig. 2a may be indicative of either intergenerational spawning site fidelity and/or minimal movement within the estuary, both conditions likely to promote divergence at neutral molecular markers. When considering the neighbour-joining tree comparing chord distances among all localities (Fig. 2b), the node with the greatest level of bootstrap support (981/1000 iterations) also clearly separated the four most upstream sites from the others. These were also the sites with the highest level of assignment (P > 0.99) to the FW genetic cluster identified by GENELAND (Fig. 3). Moreover, the other site assigned to this group (BP) also formed a contiguous node with the aforementioned sites. Finally, although we reported some evidence for isolation by distance (IBD), this result is best explained by overall divergence between genetic groups (Table 3). Indeed, when pairwise comparisons were conducted using only sites within each of the respective clusters, we detected no significant relationship between genetic and geographical distance (data not shown). Interestingly, this appears to be in contrast with many recent studies which have demonstrated stronger patterns of IBD at localized scales, but weaker or nonexistent evidence at larger spatial scales (Poissant *et al*. 2005; Bradbury *et al*. 2006; Crispo *et al*. 2006).

Weak, yet significant, genetic differentiation is typical for many populations of marine fish species (Ward *et al*. 1994; Waples 1998). More recent work employing microsatellite markers suggests that this trend is evident over a range of spatial scales. For example, Atlantic cod have been reported to exhibit weak, fine-scale population structuring (F_{ST} = 0.001; P = 0.021) between groups inhabiting neighbouring fjords (Jorde *et al*. 2007). Hierarchical analysis of pollack along the western coast of France demonstrated significant differentiation from populations of the western English Channel and the Bay of Biscay (F_{ST} = 0.004; P = 0.049); furthermore, even pairwise differences with a reference population from southern Norway ranged only from 0.001 to 0.008 (Charrier *et al*. 2006). Even at greater spatial scales, fixation indices ranging from 0.005 to 0.03 have been reported for tropical reef species inhabiting sites over 500 km apart (Fauvelot *et al*. 2007). Thus, the levels of divergence between groups of sticklebacks in the Saint Lawrence estuary are not atypical for a variety of marine species; however, they do appear substantially less divergent than other stickleback systems.

Globally, genetic differentiation among continental populations of sticklebacks based on microsatellites is relatively low (F_{ST} = 0.08), with twice as much variation among populations within continents (McKinnon *et al*. 2004). Similarly weak divergence is observed across the species' western European distribution, with significant structuring defined by drainage basin (Mäkinen *et al*. 2006). However, interpopulation divergence appears to be greatest at finer spatial scales. In Western Europe, genetic differentiation among studied lake, river and estuarine populations (F_{CT} = 0.11) was reported to be twice as great as that between their respective drainage basins (Reusch *et al*. 2001). Yet by far the

greatest reported divergence between stickleback populations occurs over a spatial scale of only 3 km (Hendry *et al*. 2002). Divergence between lake and inlet stream populations from the Misty Lake system in western Canada have been reported up to 70 times greater than that observed between freshwater and maritime landscapes within the Saint Lawrence estuary. However, nearly all research to date has focused on populations with clearly defined adaptive differentiation in morphology. As such, strong genetic divergence between locally adapted populations is expected. In contrast, based on preliminary morphological analyses (R. J. S. McCairns, unpublished data), we had no a priori evidence to suggest adaptive phenotypic divergence between stickleback populations within the estuary. Consequently, the question emerging from these results is whether the level of genetic differentiation exhibited in the Saint Lawrence estuary is indicative of nascent adaptation to the freshwater environment, or typical of weak divergence that might be predicted to accompany differentiation due to phenotypic plasticity or random drift.

Concordance between ecological and genetic landscapes

The geographical location of the genetic breakpoint between populations is indicative of a division into freshwater and saltwater habitats. The four sampling sites with the highest probability of FW cluster membership are found within the exclusively freshwater, fluvial portion of the estuary. Moreover, the 90% isocline of this cluster coincides with the downstream freshwater limits, whereas gradual isoclinic declines extend into the brackish water transition zone (Fig. 3). Interestingly, the same sites comprising the FW genetic landscape also group together based upon independent analyses of environmental data. Analyses employing GENELAND assumed no a priori population division. As such, the two genetic clusters identified may be viewed as an unbiased division of the genetic landscape. In contrast, clustering of sites based on environmental data was performed using a *k*-means technique specifying division into two clusters. The specification of two groups was based on the observation of two genetic landscapes, thus, defining the general conditions for a qualitative test of clustering similarity, and not suggestive of a dichotomous environment. Environmental clustering itself was defined strictly by minimized centroid differences of environmental data and was not influenced by results of the genetic clustering. Consequently, the similarity in site groupings based on environmental and genetic data provides the first hint of genetic differentiation as a function of environmental differences. This conjecture is further strengthened considering that environmental differences, independent of geographical distance, may account for up to 31.5% of the variance in genetic differentiation among sites (Table 4).

Use of semipartial correlations permitted us to estimate the unique effects of each variable, simultaneously controlling for covariation with all other explanatory variables, including geographical distance. Inter-site differences in salinity exhibited the highest degree of correlation with pairwise genetic divergence compared to other variables comprising PC1 (Table 4). Moreover, no other variables, regardless of their importance in defining PC1, significantly improved model fit. These results all imply a strong relationship between environmental salinity and population divergence of sticklebacks within the Saint Lawrence estuary. Interestingly, similar observations have been reported in relation to population subdivision of Atlantic herring in the North Sea–Baltic Sea transition zone: although overall population divergence was weak, genetic differentiation correlated more with salinity differences than with geographical distance (Bekkevold *et al*. 2005). This population division in herring was also interpreted in the context of adaptive divergence in light of previously studied differences in fertilization success and larval development vis-à-vis salinity gradients. Although less is known about the relationship between salinity and early development in sticklebacks, recent evidence does suggest that salinity concentration may play an important role in fertilization success (Le Comber *et al*. 2004; Elofsson *et al*. 2006). Additionally, salinity gradients have been associated with both divergent morphological variation (Heuts 1947; Kristjánsson 2005; Leinonen *et al*. 2006), as well as population subdivision (Jones *et al*. 2006; Kitano *et al*. 2007; Ólafsdóttir *et al*. 2007). These studies, however, focused on stickleback populations with known, and substantial, morphological differentiation. Moreover, little attempt was made to quantify the relationship between environmental and genetic variation, nor, apart from Heuts's (1947) hypotheses regarding lateral plate polymorphism, to interpret the role of salinity as a selection gradient or as a source of environmental variation driving phenotypic plasticity.

Correspondence between ecological and genetic landscapes may be informative regarding the environmental variables potentially driving adaptive population divergence. Although stickleback range expansion into the freshwater environment may also have favoured increased genetic drift, thus, neutral divergence from the ancestral marine population (Excoffier & Ray 2008), we believe that this does not preclude a role for adaptive processes. First, modelling results precluded random differentiation as a likely scenario (Table 5). Moreover, genetic distance was best explained as a function of environmental differences, rather than geographical distance, suggesting population divergence is unlikely to be driven by strictly neutral processes. Thus, the emerging conclusion from our analyses is that any potentially adaptive divergence between stickleback populations of the Saint Lawrence estuary is most likely associated with the salinity gradient. However, landscape genetics results alone are merely correlative, not *de facto* evidence of adaptive differentiation. Nevertheless, the goal

of our analyses was hypothesis generation, not testing. The tentative hypothesis emerging from our observations is that of physiological adaptation to either the freshwater or the saltwater environment. However, at present we cannot discount the possibility of multicollinearity between salinity and another variable. Moreover, there is no a priori reason to suggest that colonization of freshwater necessitates physiological adaptation. Many species of euryhaline and anadromous fishes are able to switch between freshwater and saltwater environments following an acclimation period (Wootton 1990). It is, thus, possible that freshwater residency is facilitated by a plastic response, although under certain environmental conditions, plasticity itself may be adaptive (Gotthard & Nylin 1995; Ghalambor *et al*. 2007). Although recent hypotheses suggest a role for phenotypic plasticity in initiating evolutionary change (West-Eberhard 2005), at this time, any putative genetic divergence ultimately resulting from plastic phenotypic differentiation remains to be confirmed empirically. These considerations all underline the necessity for formal testing in studies of local adaptation (Kawecki & Ebert 2004). The power of the landscape genetics approach lies in its exploratory capacity to tease apart the most likely candidate populations and selective gradients, under natural conditions, to facilitate such formal testing.

Migration and the potential for adaptive divergence

Consideration of the degree of dispersal and gene flow between demes is fundamental to any discussion regarding the potential for local adaptation in an open system such as the Saint Lawrence estuary. Indeed, qualitative predictions regarding the conditions most likely to promote adaptive divergence were highly dependent on the effective migration rate (Table 5). Consequently, determining a critical rate of migration between groups is essential for any meaningful interpretation of evolutionary processes within the estuary. The highest estimate ($m = 0.11$), based on the upper 95% confidence limits obtained via MLNE, seemed an obvious upper boundary. However, this may be excessive given that it is nearly two times greater than even the highest dispersal estimate based on first-generation migrants identified with GENECLASS. Moreover, dispersal rate itself may represent an overestimate of the effective migration rate. For example, in the Misty Lake watershed, in which adaptive differentiation to fluvial and lacustrine environments has been well documented, approximately 37% of sticklebacks captured in the outlet stream were identified as probable migrants from the lake population, which in turn corresponded to estimates of actual effective migration rate ranging from 0.0002 to 0.02, depending on the estimator used (Hendry & Taylor 2004). In the same system, *m*-values into the inlet stream ranged only from 0.0001 to 0.0004, despite 4–7% of individuals being identified as probable lake migrants. As such, effective migration between landscapes is likely less than the rate of dispersal. Unfortunately, dispersal rate itself must be interpreted with extreme caution. Given the low levels of differentiation between FW and SW groups, the power of the assignment method was estimated to be only 0.2, thus, suggesting a high probability of type II error (Paetkau *et al*. 1997; Paetkau *et al*. 2004). Consequently, dispersal may be under estimated. Additionally, the actual number of first-generation migrants is similar to that expected by chance alone, assuming 0.5 for a critical alpha value. However, the average likelihood-ratio scores for comparisons (L_{HOME}/L_{MAX}) identifying first-generation migrants were over 15 times greater than those in which dispersal was rejected (LR = 0.008). As such, we may safely assume some level of gene flow within the estuary.

Most germane to the issue of gene flow impeding local adaptation is the rate of migration from SW into the FW landscape. Given that freshwater habitats have been colonized by ancestral marine populations (Bell & Foster 1994; Ortí *et al*. 1994), the FW landscape can be viewed as the novel environment, and that migration from the ancestral marine environment (SW) would impede any adaptations evolved in the novel habitat. Consequently, we may posit that the critical migration rate is 0.004, based on the MIGRATE estimate of *m* from SW to FW, with a possible upper bound of 0.01, as determined by the lower confidence limits of the MLNE derived estimate. At these levels of migration, we have qualitatively predicted the potential for adaptive differentiation between freshwater and saltwater landscapes, even under a moderate level of selection ($s = 0.01$); however, we cannot estimate the selection differential at this time. If selection is in fact weak ($s = 0.001$), the FW and SW demes may be relatively, genetically homogenous. Notwithstanding the caveats regarding estimated dispersal rates, if migration into the FW landscape is comparable to even the lowest estimate (0.02), then relative genetic homogeneity is predicted for all but the highest levels of selection. Moreover, at comparable levels of migration (0.0002–0.02) within the Misty Lake system, in which selection is presumed to be relatively strong, gene flow downstream from the lake population may have constrained adaptation in the outlet stream by 80–86% (Moore *et al*. 2007). As such, we cannot rule out a role for plasticity in explaining any putative interlandscape phenotypic differentiation to cope with these two contrasting environments. Common garden experiments to test these two alternative scenarios (adaptive differentiation vs. plasticity) are ongoing.

Conclusions

We have highlighted the utility of landscape genetic analyses to identify subtle population structure and concomitant environmental variation representative of a potential selection gradient. In the case of three-spined sticklebacks of the Saint Lawrence estuary, we found that genetic differentiation is

best explained by covariation with environmental salinity, suggesting that phenotypic variation may be at the level of environmental physiology. Until laboratory experiments are completed, we cannot rule out either alternative scenario, that this may represent cryptic adaptive variation to a heterogeneous environment, or a classic example of phenotypic plasticity. Although a landscape genetics analysis may represent but a first step in the study of local adaptation, it should be viewed as an efficient method of refining hypotheses reflective of actual environmental conditions experienced by focal populations. Moreover, this approach may be independent of the nature or magnitude of phenotypic differentiation between populations. As such, this may represent a powerful exploratory tool to identify natural populations amenable for research into the conditions favouring phenotypic plasticity, local adaptation or adaptive plasticity, an area of continued interest in evolutionary ecology.

Acknowledgements

For their assistance in the field and/or laboratory, we thank S. Bourget, C. Lippé, J. Roy and S. Uusi-Heikkilä. We are grateful to B. Pelchat of the Oceanographic Data Management section, Maurice Lamontagne Institute, Department of Fisheries and Oceans Canada for access to temperature and salinity data. This manuscript was improved considerably thanks to comments provided by M.M. Hansen and three anonymous reviewers. Financial support for this research was provided via a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to L.B. R.J.S.M. acknowledges the financial support of a Canadian Graduate Scholarship (NSERC) and Fonds de Soutien au Doctorat, from both Québec Océan and the Département de biologie, Université Laval.

References

Adkison MD (1995) Population differentiation in Pacific salmon: local adaptation, genetic drift, or the environment? *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 2762–2777.

Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, **25**, 4692–4693.

Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill AVS (2007) Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes*, **7**, 10–14.

Avise JC (2004) *Molecular Markers, Natural History, and Evolution*. Sinauer & Associates, Sunderland, Massachusetts.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences, USA*, **98**, 4563–4568.

Bekkevold D, Andre C, Dahlgren TG *et al.* (2005) Environmental correlates of population differentiation in Atlantic herring. *Evolution*, **59**, 2656–2668.

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

Bradbury IR, Coulson MW, Campana SE, Bentzen P (2006) Morphological and genetic differentiation in anadromous smelt *Osmerus mordax* (Mitchill): disentangling the effects of geography and morphology on gene flow. *Journal of Fish Biology*, **69**, 95–114.

Cavalli-Sforza LL (1998) The DNA revolution in population genetics. *Trends in Genetics*, **14**, 60–65.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, **21**, 550–570.

Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetical Research*, **70**, 155–174.

Charrier G, Durand JD, Quiniou L, Laroche J (2006) An investigation of the population genetic structure of pollack (*Pollachius pollachius*) based on microsatellite markers. *ICES Journal of Marine Science*, **63**, 1705–1709.

Cressie NAC (1993) *Statistics for Spatial Data*. John Wiley & Sons, New York.

Crispo E, Bentzen P, Reznick DN, Kinnison MT, Hendry AP (2006) The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, **15**, 49–62.

Crow JF (1987) Population genetics history: a personal view. *Annual Review of Genetics*, **21**, 1–22.

Douglas MR, Brunner PC, Bernatchez L (1999) Do assemblages of *Coregonus* (Teleostei: Salmoniformes) in the Central Alpine region of Europe represent species flocks? *Molecular Ecology*, **8**, 589–603.

Elofsson H, Van Look KJW, Sundell K, Sundh H, Borg B (2006) Stickleback sperm saved by salt in ovarian fluid. *Journal of Experimental Biology*, **209**, 4230–4237.

El-Sabh MI, Silverberg N (1990) The St. Lawrence Estuary: Introduction. In: *Oceanography of a Large-Scale Estuarine System: The St. Lawrence* (eds El-Sabh MI, Silverberg N), pp. 1–9. Springer-Verlag, Berlin, Germany.

Estoup A, Guillot G, Santos F (2007) GENELAND: Simulation and MCMC inference in landscape genetics. R package version 2.0.5. URL: http://cran.r-project.org/.

Excoffier L, Ray N (2008) Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution*, **23**, 347–351.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial-DNA restriction data. *Genetics*, **131**, 479–491.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Fauvelot C, Lemaire C, Planes S, Bonhomme F (2007) Inferring gene flow in coral reef fishes from different molecular markers: which loci to trust? *Heredity*, **99**, 331–339.

Felsenstein J (2005) PHYLIP *(Phylogeny Inference Package) version 3.6*. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle, Washington.

Fisher RA (1930) *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford, UK.

Funk DJ, Nosil P, Etges WJ (2006) Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proceedings of the National Academy of Sciences, USA*, **103**, 3209–3213.

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

Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, **21**, 394–407.

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 (1995) FSTAT (version 1.2): a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486.

Goudet J (2005) HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Notes*, **5**, 184–186.

Goudet J (2006) HIERFSTAT: Estimation and tests of hierarchical *F*-statistics. R package version 0.04-4. URL: http://cran.r-project.org/.

Guillot G, Estoup A, Mortier F, Cosson JF (2005a) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261–1280.

Guillot G, Mortier F, Estoup A (2005b) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, **4**, 9.

Hansen MM, Ruzzante DE, Nielsen EE, Bekkevold D, Mensberg KLD (2002) Long-term effective population sizes, temporal stability of genetic composition and potential for local adaptation in anadromous brown trout (*Salmo trutta*) populations. *Molecular Ecology*, **11**, 2523–2535.

Hendry AP, Taylor EB (2004) How much of the variation in adaptive divergence can be explained by gene flow? An evaluation using lake-stream stickleback pairs. *Evolution*, **58**, 2319–2331.

Hendry AP, Day T, Taylor EB (2001) Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. *Evolution*, **55**, 459–466.

Hendry AP, Taylor EB, McPhail JD (2002) Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the misty system. *Evolution*, **56**, 1199–1216.

Heuts MJ (1947) Experimental studies on adaptive evolution in *Gasterosteus aculeatus* L. *Evolution*, **1**, 89–102.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898–1914.

Jones FC, Brown C, Pemberton JM, Braithwaite VA (2006) Reproductive isolation in a threespine stickleback hybrid zone. *Journal of Evolutionary Biology*, **19**, 1531–1544.

Jorde PE, Knutsen H, Espeland SH, Stenseth NC (2007) Spatial scale of genetic structuring in coastal cod *Gadus morhua* and geographic extent of local populations. *Marine Ecology Progress Series*, **343**, 229–237.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Kimura M (1968) Evolutionary rate at the molecular level. *Nature*, **217**, 624–626.

Kitano J, Mori S, Peichel CL (2007) Phenotypic divergence and reproductive isolation between sympatric forms of Japanese threespine sticklebacks. *Biological Journal of the Linnean Society*, **91**, 671–685.

Kristjánsson BK (2005) Rapid morphological changes in threespine stickleback, *Gasterosteus aculeatus*, in freshwater. *Environmental Biology of Fishes*, **74**, 357–363.

Laprise R, Dodson JJ (1994) Environmental variability as a factor controlling spatial patterns in distribution and species-diversity of zooplankton in the St. Lawrence estuary. *Marine Ecology Progress Series*, **107**, 67–81.

Largiadèr CR, Fries V, Kobler B, Bakker TCM (1999) Isolation and characterization of microsatellite loci from the three-spined stickleback (*Gasterosteus aculeatus* L.). *Molecular Ecology*, **8**, 342–344.

Latta RG (2006) Integrating patterns across multiple genetic markers to infer spatial processes. *Landscape Ecology*, **21**, 809–820.

Le Comber SC, Faulkes CG, Van Look KJW, Holt WV, Smith C (2004) Recovery of sperm activity after osmotic shock in the three-spined stickleback: implications for pre-oviposition ejaculation. *Behaviour*, **141**, 1555–1569.

Lecomte F, Dodson JJ (2004) Role of early life-history constraints and resource polymorphism in the segregation of sympatric populations of an estuarine fish. *Evolutionary Ecology Research*, **6**, 631–658.

Lecomte F, Dodson JJ (2005) Distinguishing trophic and habitat partitioning among sympatric populations of the estuarine fish *Osmerus mordax* Mitchill. *Journal of Fish Biology*, **66**, 1601–1623.

Leinonen T, Cano JM, Mäkinen H, Merilä J (2006) Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of Evolutionary Biology*, **19**, 1803–1812.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.

Mäkinen HS, Cano JM, Merilä J (2006) Genetic relationships among marine and freshwater populations of the European threespined stickleback (*Gasterosteus aculeatus*) revealed by microsatellites. *Molecular Ecology*, **15**, 1519–1534.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*, **17**, 285–291.

McKinnon JS, Rundle HD (2002) Speciation in nature: the threespine stickleback model systems. *Trends in Ecology & Evolution*, **17**, 480–488.

McKinnon JS, Mori S, Blackman BK *et al*. (2004) Evidence for ecology's role in speciation. *Nature*, **429**, 294–298.

Menozzi P, Piazza A, Cavalli-Sforza LL (1978) Synthetic maps of human gene frequencies in Europeans. *Science*, **201**, 786–792.

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

Nagylaki T, Lucier B (1980) Numerical analysis of random drift in a cline. *Genetics*, **94**, 497–517.

Nosil P, Egan SR, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: 'isolation by adaptation' and multiple roles for divergent selection. *Evolution*, **62**, 316–336.

Nychka D (2007) FIELDS: Tools for spatial data. R package version 3.5. URL: http://cran.r-project.org/.

Oksanen J, Kindt R, Legendre P, O'Hara RB (2007) VEGAN: Community Ecology Package. R package version 1.8–6. URL: http://cran.r-project.org/.

Ólafsdóttir GA, Snorrason SS, Ritchie MG (2007) Parallel evolution? Microsatellite variation of recently isolated marine and freshwater three-spined stickleback. *Journal of Fish Biology*, **70**, 125–131.

Ortí G, Bell MA, Reimchen TE, Meyer A (1994) Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. *Evolution*, **48**, 608–622.

Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics*, **147**, 1943–1957.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.

Peichel CL, Nereng KS, Ohgi KA *et al*. (2001) The genetic architecture of divergence between threespine stickleback species. *Nature*, **414**, 901–905.

Picard P, Dodson JJ, FitzGerald GJ (1990) Habitat segregation among the age-groups of *Gasterosteus aculeatus* (Pisces, Gasterosteidae) in the middle St-Lawrence estuary, Canada. *Canadian Journal of Zoology*, **68**, 1202–1208.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS 2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, **95**, 536–539.

Poissant J, Knight TW, Ferguson MM (2005) Nonequilibrium conditions following landscape rearrangement: the relative contribution of past and current hydrological landscapes on the genetic structure of a stream-dwelling fish. *Molecular Ecology*, **14**, 1321–1331.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, **6**, 847–859.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Raeymaekers JAM, Maes GE, Geldof S, Honis I, Nackaerts K, Volckaert FAM (2008) Modeling genetic connectivity in sticklebacks as a guideline for river restoration. *Evolutionary Applications*, online early.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences, USA*, **94**, 9197–9201.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Reebs SG, Whoriskey FG, FitzGerald GJ (1984) Diel patterns of fanning activity, egg respiration, and the nocturnal behavior of male threespined sticklebacks, *Gasterosteus aculeatus* L. (F-*trachurus*). *Canadian Journal of Zoology*, **62**, 329–334.

Reusch TBH, Wegner KM, Kalbe M (2001) Rapid genetic divergence in postglacial populations of threespine stickleback (*Gasterosteus aculeatus*): the role of habitat type, drainage and geographical proximity. *Molecular Ecology*, **10**, 2435–2445.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Schluter D (1996) Ecological speciation in postglacial fishes. *Philosophical Transactions of the Royal Society of London. Series B-Biological Sciences*, **351**, 807–814.

Shimoda N, Knapik EW, Ziniti J *et al*. (1999) Zebrafish genetic map with 2000 microsatellite markers. *Genomics*, **58**, 219–232.

Slatkin M (1973) Gene flow and selection in a cline. *Genetics*, **75**, 733–756.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Storfer A, Murphy MA, Evans JS *et al*. (2007) Putting the 'landscape' in landscape genetics. *Heredity*, **98**, 128–142.

Team RDC (2007) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Tita G, Vincx M, Desrosiers G (1999) Size spectra, body width and morphotypes of intertidal nematodes: an ecological interpretation. *Journal of the Marine Biological Association, UK*, **79**, 1007–1015.

Venables WN, Ripley B (2002) *Modern Applied Statistics with S*, 4th edn. Springer-Verlag, New York.

Vincent WF, Dodson JJ (1999) The St. Lawrence River, Canada-USA: the need for an ecosystem-level understanding of large rivers. *Japanese Journal of Limnology*, **60**, 29–50.

Vincent WF, Dodson JJ, Bertrand N, Frenette JJ (1996) Photosynthetic and bacterial production gradients in a larval fish nursery: the St Lawrence river transition zone. *Marine Ecology Progress Series*, **139**, 227–238.

Wang JL, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, **163**, 429–446.

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438–450.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419–1439.

Ward RD, Woodwark M, Skibinski DOF (1994) A comparison of genetic diversity levels in marine, fresh-water, and anadromous fishes. *Journal of Fish Biology*, **44**, 213–232.

West-Eberhard MJ (2005) Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences, USA*, **102**, 6543–6549.

Winkler G, Dodson JJ, Bertrand N, Thivierge D, Vincent WF (2003) Trophic coupling across the St. Lawrence River estuarine transition zone. *Marine Ecology Progress Series*, **251**, 59–73.

Wootton RJ (1990) *The Ecology of Teleost Fishes*. Chapman & Hall, London, UK.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 0097–0159.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Yang RC (1998) Estimating hierarchical *F*-statistics. *Evolution*, **52**, 950–956.

Yue GH, David L, Orban L (2007) Mutation rate and pattern of microsatellites in common carp (*Cyprinus carpio* L.). *Genetica*, **129**, 329–331.

This research is part of Scott McCairns's PhD thesis on environmental and genetic variance underlying morphological, physiological and behavioural traits of three-spined sticklebacks inhabiting the Saint Lawrence estuary. Louis Bernatchez is Scott McCairns's thesis supervisor. His research interests focus on the understanding of patterns and processes of molecular and organismal evolution, and their relevance to conservation.

Supplementary material

The following supplementary material is available for this article:

Table S1 Sampling sites, as per Fig. 1, and the number of individuals collected (*N*) during each year of sampling

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-294X.2008.03884.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.