Identifying the spatial scale of population structure in anadromous rainbow smelt (*Osmerus mordax*)

Adrienne I. Kovach [a,*], Timothy S. Breton [b], Claire Enterline [b,c], David L. Berlinsky [b]

[a] *Department of Natural Resources and the Environment, University of New Hampshire, Rudman Hall, 46 College Rd., Durham, NH 03824, USA*
[b] *Department of Biological Sciences, University of New Hampshire, Spaulding Hall, 38 College Rd., Durham, NH 03824, USA*
[c] *Maine Department of Marine Resources, Bureau of Marine Science, Sea Run Fisheries and Habitat, 21 State House Station, Augusta, ME 04333, USA*

ARTICLE INFO

ABSTRACT

Identifying conservation units based on the scale of biological processes is important in fisheries management. In anadromous fish, the spatial scale of population structure results from the opposing effects of homing and dispersal and varies among species by life history and geographic location. We used genetic data to evaluate the population structure of anadromous rainbow smelt (*Osmerus mordax*), a species of concern in U.S. waters. A total of 2211 smelt were genotyped at 10 microsatellite loci from 18 river systems and 11 bays along the northeastern Atlantic coast, spanning the entire U.S. range of the species. Across the study area gene flow was relatively high ($F_{ST} = 0.017$) and genetic variation followed an isolation by distance model, consistent with dispersal occurring most frequently among nearby rivers. Bayesian clustering approaches identified 4–6 genetically distinct population clusters, which varied in their geographic extent according to coastal circulation patterns. Genetically divergent populations were identified in topographically structured bays with local hydrography favoring larval retention. Conversely, gene flow was high across long stretches of topographically indistinct coastline where circulation patterns maximize passive larval dispersal. Overall, our results suggest that larval retention patterns, driven by hydrography, influence the population structure of rainbow smelt, and homing occurs to broad coastal zones of retention. The genetic structure identified in this study can be used in conjunction with ecological data to inform management at the appropriate spatial scale.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Understanding the scale at which natural populations are demographically structured and the causes of population divergence is necessary for successful conservation and management. In fisheries management, identifying conservation units based on the scale of biological processes is of critical importance (Reiss et al., 2009), because spatial structure affects how populations respond to fishing pressure as well as management actions (Cadrin and Secor, 2009). Genetic data can improve our understanding of biological structure and aid the objective of achieving congruence of spatial scales between population structure and management units (Reiss et al., 2009; Waples et al., 2008). Genetic approaches can also identify barriers to larval and adult dispersal and yield insight into drivers of spatial variability in recruitment patterns and the scale at which these processes occur (Selkoe et al., 2008).

The scale at which population genetic divergence and demographic independence occurs is reflective of a species' migratory or dispersal potential and varies across taxa (Bohonak, 1999; Avise, 2006). For fish, migration and gene flow are generally greater for marine species relative to anadromous and freshwater ones, such that population differentiation is generally stronger on finer spatial scales in freshwater species, intermediate in anadromous ones, and weakest in marine species (Ward et al., 1994). These trends are attributed to the relatively fewer and weaker barriers to dispersal found in the marine environment in comparison to geographical structuring provided by rivers, estuaries and the isolation of land-locked freshwater environments. Patterns vary among species, however, and fine-scale population structure has been found even in high migration marine species (reviewed in Hauser and Carvalho, 2008). The scale of demographic independence, therefore, must be determined specifically for each system of interest (species, population or geographic region).

Many anadromous species are philopatric, meaning that adults return to their natal site to reproduce after estuarine or marine feeding migrations. This behavior promotes population genetic structuring often on the scale of individual river systems (Allendorf and Waples, 1996; Spidle et al., 2003; Beacham et al., 2008). Not all individuals in a population are strictly philopatric, however, and straying between rivers is typical. Further, homing may not be to

* Corresponding author. Tel.: +1 603 862 1603; fax: +1 603 862 4976.
E-mail addresses: akovach@unh.edu, adrienne.kovach@unh.edu (A.I. Kovach).

a single river, but rather to a group of rivers that share a coastal area or inlet (McLean et al., 1999) or to specific tributaries within a river system (Dionne et al., 2009). Variability in patterns of philopatry might be explained by the member-vagrant hypothesis, which states that the population structure of a species is determined by the number of distinct larval retention sites or habitats rather than the number of spawning populations (Iles and Sinclair, 1982; Sinclair, 1988). In turn, the geographic scale of larval retention varies with species' life history strategies and geographic location and is influenced by factors including migration distance and duration of freshwater residence (McLean et al., 1999; Bradbury et al., 2006b). Population structure may also be influenced by divergent adaptation to local environmental, thermal or habitat characteristics (Taylor, 1991; Bradbury et al., 2006a, 2008b; Dionne et al., 2007, 2008) or by landscape features that function as dispersal barriers or promote migration (Costello et al., 2003; Meeuwig et al., 2010).

Anadromous rainbow smelt (*Osmerus mordax*) were historically distributed along the east coast of North America from Labrador south to the Delaware River (Buckley, 1989). The southern range of the species has contracted in recent decades, such that today it only extends to northern Massachusetts. Due to this population decline, smelt was listed as a Species of Concern in U.S. waters (NOAA, 2004). Anadromous smelt spawn in freshwater, coastal streams from February to early June in the northeastern U.S., depending on latitudinal gradients of water temperature (Buckley, 1989). Larvae are passively transported downstream after hatching and develop in estuaries and inner coastal waters. Adults return to estuaries and coastal waters after spawning and may move into slightly cooler shallow water within 2 km of the coast during the summer (Collette and Klein-MacPhee, 2002). Little is known of movements and habitat use during the adult phase of the life cycle, and regional differences may occur. Mark-recapture studies have documented straying between adjacent streams (Murawski et al., 1980), and it has been suggested that natal homing is rare when inter-river distances are small (Frechet et al., 1983). Genetic studies have shown that estuarine and bay-scale structuring occurs in some but not all portions of the species' range in Canadian waters (Bradbury et al., 2006b), but nothing is known about the spatial scale of population structure in U.S. waters.

Currently, a cohesive management plan does not exist for smelt in U.S. waters, although Maine, New Hampshire, and Massachusetts each have state-specific monitoring and management activities. To aid proactive management of the species in light of the recent population decline and because of the important coastal and estuarine sport-fishery that it supports, a primary recent focus has been on gathering data to develop a regional conservation plan. To this end, the purpose of this study was to investigate the population genetic structure of anadromous rainbow smelt across its range in the northeastern U.S. Our specific objectives were to (1) characterize genetic variation within and among multiple river systems; (2) determine the geographic scale of philopatry; and (3) provide insight into the appropriate spatial scale of management.

2. Methods

2.1. Sample collection

Adult rainbow smelt were sampled during spring spawning runs in 18 river systems from Cobscook Bay, Maine to Buzzard's Bay, Massachusetts between 2006 and 2010 (Fig. 1). Smelt were collected by fyke net above the head of tide, except in the Pleasant and Kennebec Rivers, where they were collected below the ice by gill and bag nets or hook and line. The sampling scheme included rivers from multiple separate estuaries and bays, as well as some rivers that shared common estuaries, such as the 5 rivers that feed

into Great Bay, New Hampshire (Squamscott, Lamprey, Bellamy, Oyster, and Salmon Falls). To facilitate testing for annual fluctuations in genetic variation, 11 of 18 rivers were sampled in multiple years (8 rivers in each of 2 years and 3 rivers in each of 3 years; Table 1). A fin clip sample (1 cm^2) was taken from each fish prior to release and preserved in 100% ethanol until genetic analysis.

2.2. Genetic analyses and descriptive statistics

Genomic DNA was extracted from fin clips using the Qiagen DNeasy Blood and Tissue kit (Qiagen, Valencia, CA) and assayed with 11 microsatellite markers (Omo1–5, Omo9, Omo11, Omo13–16; Coulson et al., 2006). Microsatellite analysis was conducted by polymerase chain reaction (PCR) in 12 µl volumes with 3.5 µl DNA template, 0.25–0.6 µM of each primer (labeled with FAM, HEX or NED fluorescent dyes), 1× GoTaq Flexi PCR buffer (Promega, Madison, WI), 0.2 mg/ml bovine serum albumin, 1.5 mM MgCl$_2$, 100 µM deoxynucleoside triphosphates (dNTPs), and 0.25 U GoTaq Flexi DNA polymerase (Promega). Multiplex PCRs were conducted in three sets: Omo1, 2, 4, and 14 with annealing at 62 °C; Omo3, 5, 9, and 11 with annealing at 60 °C; and Omo13, 15, and 16 with touchdown PCR at 64–60 °C. Cycling parameters followed those of Coulson et al. (2006). PCR products were electrophoresed in an ABI3130 automated capillary sequencer (Applied Biosystems, Carlsbad, CA) and manually scored using PeakScanner v.1.0 software (Applied Biosystems). Raw scores were sorted manually into allelic bins with the aid of positive controls (run on every PCR plate) to normalize interassay variation.

Tests of deviations from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) per locus and population were conducted in GENEPOP v.3.4 (Raymond and Rousset, 1995) using a Markov chain Monte Carlo method with 10,000 iterations and 10,000 batches. Significance level for multiple tests was adjusted to $\alpha = 0.05$ with the standard Bonferroni method (Rice, 1989). Significant LD was evident between Omo3 and Omo16 ($p < 0.0001$), therefore Omo3 was removed from the dataset; subsequent analyses were conducted with the remaining 10 loci.

2.3. Genetic variation within rivers

Measures of allelic diversity, including mean number of alleles, F_{IS}, and observed and expected heterozygosities (H_O and H_E) were calculated for each sampled river using GDA (Lewis and Zaykin, 2001). Mean allelic richness was calculated in FSTAT 2.9.3 (Goudet, 1995). To assess genetic variability among rivers, we compared allelic richness using a one-way analysis of variance (ANOVA), blocked by locus in JMP 8.0 (SAS Institute, Cary, NC). To assess annual fluctuation in genetic variation, a hierarchical analysis of molecular variance (AMOVA) with 10,000 permutations was conducted in ARLEQUIN 2.0 (Schneider et al., 2000) using samples collected from the same rivers in successive seasons. Genetic variation was partitioned among sample sites and years to evaluate whether greater variation was present among sites than among yearly samples from the same site, a requisite for identifying meaningful population structure (Waples, 1998). Temporal stability was also assessed using the F_{ST} estimator θ (Weir and Cockerham, 1984) in FSTAT. Pairwise population F_{ST}s were evaluated for pairs of yearly collections, and significance evaluated following standard Bonferroni correction for multiple tests ($\alpha = 0.05$). In the absence of significant differentiation, yearly samples from the same rivers were pooled for further analyses.

We tested for genetic effects of recent demographic bottlenecks, periods of severe population decline, in three sample locations (Cobscook, Jones and Weweantic) found at the northern and southern limits of our study area. Smelt samples from these locations were the most genetically divergent and had
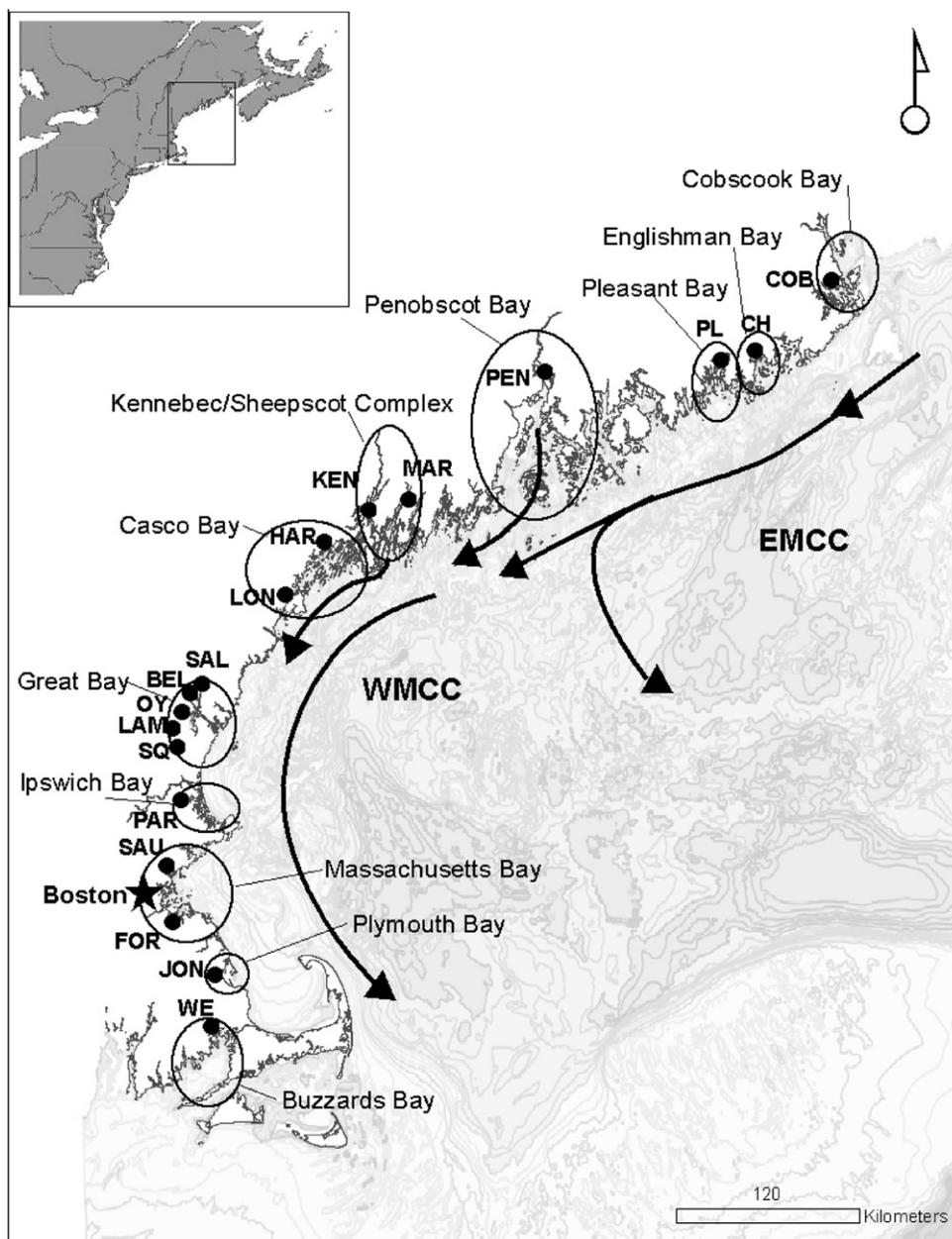
Fig. 1. Sampling locations of rainbow smelt from 18 river systems in the northeastern U.S. and their associated bays. Encircled sites share a common embayment or bay. Abbreviated site codes refer to Table 1. Principle features of the coastal circulation patterns in the Gulf of Maine are indicated with arrows, including the Eastern Maine Coastal Current (EMCC), Western Maine Coastal Current (WMCC), and freshwater inputs from the Penobscot and Kennebec Rivers.

the lowest allelic diversity (see Section 3.1), suggesting that they were the most likely to have experienced a recent local decline in population size. We used two methods, known to differ in their sensitivity for detecting bottlenecks of different time-scale and duration (Williamson-Nateson, 2005): BOTTLENECK 1.2.02 (Piry et al., 1999) is appropriate for detecting bottle-necks within a few dozen generations and the M-ratio method (Garza and Williamson, 2001) detects longer duration popula-tion declines that occurred in the more distance past. We ran 1000 iterations in BOTTLENECK with all mutation models, assum-ing an 88% stepwise mutation and 12% infinite allele mutation with variance among multiple steps set at 12. Results were assessed using the Wilcoxon signed-rank test of heterozygosity excess and the allele frequency mode-shift test (Luikart et al., 1998). We also used the M-ratio method (Garza and Williamson, 2001) with softwares M_P_Val.exe and Critical_M.exe (available at

http://swfsc.noaa.gov/textblock.aspx?Division=FED&id=3298). To calculate the M-ratio, we used an effective population size (N_e) of 1000 individuals (Hansen et al., 2002) and a standard microsatellite mutation rate (μ) of 5×10^{-4}. Mean size of single step mutations and percent mutations larger than single step were set to 0.88 and 2.8, respectively, following Garza and Williamson (2001).

2.4. Population genetic structure

To evaluate genetic variation among smelt from the 18 rivers, pairwise F_{ST}s were calculated in FSTAT and significance evaluated using Bonferroni correction, as above. A global F_{ST} was calcu-lated in GENEPOP. To evaluate an isolation by distance (IBD) model of genetic structure (i.e. whether genetic differentiation is a function of geographic distance), we used Mantel tests imple-mented in GENALEX 6.1 (Peakall and Smouse, 2006) to test for a

Table 1
Location information for rainbow smelt collected from 18 river systems that feed into 11 bays, including the abbreviated site name, associated embayment or river system, collection stream or river, and dates of sampling. n indicates sample size genotyped.

Bay	Name	Embayment or river system	Collection stream or river	Year	n
Cobscook Bay	COB	East Bay	East Bay Brook	May–Jun 08	42
				May–Jun 09	72
Englishman Bay	CH	Chandler River	Chandler River	Apr–Jun 09	31
			Schoppee Brook	Apr–May 10	85
Pleasant Bay	PL	Pleasant River	Pleasant River	Feb–May 10	79
Penobscot Bay	PEN	Penobscot River	Tannery Brook	Apr–May 08	74
				Apr–May 09	67
Sheepscot/Kennebec Complex	MAR	Sheepscot River, Marsh River	Deer Meadow Brook	Apr–May 08	62
				Apr 09	84
	KEN	Kennebec River, Androscoggin River	Kennebec River, Merrymeeting Bay	Jan–Feb 09	64
Casco Bay	HAR	Harraseeket River	Mill Stream at Mast Landing	Apr 08	60
				Apr 09	60
	LON	Fore River	Long Creek	Mar–Apr 09	73
Great Bay	SAL	Piscataqua River	Salmon Falls	Apr 08	46
	BEL	Little Bay	Bellamy River	Apr 07	67
				Apr 08	67
	OY	Little Bay	Oyster River	Feb–Apr 07	79
	LAM	Lamprey River	Lamprey River	Apr 08	76
	SQ	Squamscott River	Squamscott River	Mar 07	48
				Apr 08	89
				Mar–Apr 09	83
Ipswich Bay	PAR	Plum Island Sound	Parker River	Mar–May 08	77
				Mar–Apr 09	62
Massachusetts Bay	SAU	Broad Sound	Saugus River	Mar–May 06	29
				Mar–May 07	64
				Mar–May 08	68
	FOR	Hingham Bay	Fore River	Mar–May 06	83
				Mar–May 07	85
				Mar–May 08	70
Plymouth Bay	JON	Kingston Bay	Jones River	Mar–May 08	98
				Mar–May 09	89
Buzzards Bay	WE	–	Weweantic River	Mar–Apr 08	78

correlation between linearized genetic distance ($F_{ST}/(1 − F_{ST})$) and geographic distance. Shortest coastal geographic distances were calculated using Google Earth 4.2 (Google, Mountain View, CA) and followed shoreline contours (within 5 km of the coast) between sampling locations. IBD was assessed using geographic distances measured with and without inclusion of the Cape Cod Canal as a possible migration corridor between Buzzards and Cape Cod Bays.

We also used spatial autocorrelation analysis to evaluate fine-scale geographic structuring of genetic variation. This approach uses pair-wise genetic and geographic distance matrices to calculate an autocorrelation coefficient (r) for each of a series of predetermined distance classes (Smouse and Peakall, 1999). The resulting correlogram provides a means for inferring spatial patterns of gene flow based on the distances at which r (a measure of genetic similarity) is positive (Peakall et al., 2003). The distance where r intercepts the x-axis provides an estimate of the spatial extent of genetic structure, beyond which gene flow is no longer effective in connecting populations. Spatial autocorrelation analyses were performed in GENALEX with 999 permutations, 1000 bootstraps, and 50 km distance classes up to 700 km, both with and without inclusion of the Cape Cod Canal.

To identify genetically similar groupings of smelt without prior population assumptions, we used two Bayesian clustering approaches. First we used the LocPrior algorithm in STRUCTURE 2.3.2 (Pritchard et al., 2000; Hubisz et al., 2009), which incorporates sampling location information and is appropriate for detecting weak population structure (Hubisz et al., 2009). Five runs were performed at each a priori number of assumed populations (K) from 1 to 18, with a 300,000 burn-in and 200,000 iterations, using a no-admixture model with correlated allele frequencies (Falush et al., 2003). The most likely number of clusters (K) was determined using the ΔK method (Evanno et al., 2005), as well as by examining the plateau of the $\ln \Pr(X|K)$ (Pritchard et al., 2000) and evaluation of the bar plots. We conducted 25 additional runs for the most likely K and averaged results across runs using the greedy algorithm implemented in CLUMPP (Jakobsson and Rosenberg, 2007); results were plotted in DISTRUCT (Rosenberg, 2004). As different clustering methods may yield slightly different results (Latch et al., 2006; Francois and Durand, 2010), we also conducted an analysis using spatial clustering of groups in BAPS 5.3 (Corander et al., 2008).

To further evaluate the scale of population structure at hierarchical levels, we used assignment tests and AMOVA. Assignment tests were used to determine the probability of an individual originating from the location or population grouping in which it was sampled. We used the Bayesian approach of Rannala and Mountain (1997) in GENECLASS2 (Piry et al., 2004) to calculate the log likelihood of each individual's genotype originating from (1) each of the individual rivers, (2) each of the bays in Table 1, (3) the five population clusters identified by STRUCTURE, and (4) the four clusters identified by BAPS. The latter two tests were post hoc and used to evaluate the relative strength of the population groupings in comparison to structuring at the finer levels (bay and river). Samples from the five rivers of the Great Bay estuary were combined for this analysis due to high genetic similarity (see Section 3.2). Individuals were assigned to the river, bay or cluster of highest likelihood; percentage of correct assignments were compared among the different hierarchical groupings. Chi-square (χ^2) tests were performed to determine if the observed number of correct assignments was significantly greater than the number expected by chance. Expected numbers of correct assignments were calculated assuming equal assignment probabilities to each river, bay, or population cluster. We used AMOVA to compare the apportionment of genetic variation spatially among and within population groupings for the 11 bays, five STRUCTURE clusters, and four BAPS clusters.

Table 2
Analysis of molecular variance (AMOVA) partitioning genetic variation between individual rainbow smelt, within replicate collections (2 or 3 years, see Table 1 from 11 rivers and among yearly collections and rivers (temporal), as well as within 18 rivers, and within and among 3 different spatial population groupings: 4 genetically similar clusters identified by BAPS, 5 clusters identified by STRUCTURE, and 11 bays. d.f. = degrees of freedom.

Grouping	Source of variation	d.f.	Percent variation	Fixation index	p-Value
Temporal	Among rivers	10	1.26	0.01261	<0.0001
	Among yearly collections within a river	14	0.08	0.00083	<0.0001
	Within collections	3407	98.66	0.01343	<0.0001
4 BAPS clusters	Among clusters	3	1.83	0.02267	<0.0001
	Among rivers within clusters	14	0.44	0.00447	<0.0001
	Within rivers	4404	97.73	0.01827	<0.0001
5 STRUCTURE clusters	Among clusters	4	1.62	0.02004	<0.0001
	Among rivers within clusters	13	0.38	0.00389	<0.0001
	Within rivers	4404	98.00	0.01621	<0.0001
11 bays	Among bays	10	1.49	0.0155	<0.0001
	Among rivers within bays	7	0.06	0.0006	0.055
	Within rivers	4404	98.45	0.0149	<0.0001

3. Results

Mutlilocus genotypes using at least 9 of the 10 microsatellite markers were obtained for 2211 individuals. Only one collection contained >5% missing data at any one locus (10% missing data at *Omo*4 from Marsh River). No sample exhibited significant deviation from HWE at any locus. Accordingly, F_{IS} values were low and not significantly different from zero for all samples (−0.010 to 0.068; Table 3).

3.1. Genetic variation within rivers

Smelt samples collected from the same river system in separate years were found to be genetically homogenous by both F_{ST} and AMOVA. The within-river fixation index was negligible ($F_{SC} = 0.00083$) compared to that among rivers ($F_{CT} = 0.01261$), indicating that annual fluctuations in genetic variation were minimal compared to differences among sample locations (Waples, 1998; Table 2). Similarly, pairwise F_{ST}s for collections in successive years ranged from −0.0030 to 0.0034, with no significant comparisons following Bonferroni correction. Given this stability in genetic structure, yearly collections from the same rivers were pooled in further analyses.

Genetic diversity of population samples are shown in Table 3. Mean number of alleles ranged from 11.0 to 18.8. Observed heterozygosities were relatively high (mean $H_O = 0.859$), although

Table 3
Genetic diversity of rainbow smelt from 18 river systems in the northeastern U.S. Column values are means across 10 microsatellite markers for number of alleles, allelic richness (AR), expected (H_E) and observed (H_O) heterozygosities, and F_{IS} values. Abbreviated site names refer to Table 1 and n indicates sample sizes (pooled across yearly collections from the same site).

Name	n	Alleles	AR	H_E	H_O	F_{IS}
COB	114	14.2	12.2	0.840	0.833	0.009
CH	116	16.3	13.7	0.870	0.840	0.035
PL	79	15.1	13.6	0.866	0.842	0.028
PEN	141	17.1	13.9	0.857	0.846	0.012
MAR	146	17.5	14.0	0.869	0.863	0.008
KEN	64	14.9	13.8	0.873	0.859	0.016
HAR	120	16.7	13.9	0.868	0.871	−0.003
LON	73	15.5	13.9	0.868	0.810	0.068
SAL	46	14.8	14.7	0.870	0.845	0.030
BEL	134	15.8	14.2	0.873	0.837	0.042
OY	79	17.4	14.5	0.878	0.856	0.024
LAM	76	15.8	14.2	0.874	0.864	0.011
SQ	220	18.8	14.5	0.879	0.868	0.012
PAR	139	16.5	13.9	0.874	0.869	0.005
SAU	238	17.1	13.7	0.853	0.855	−0.002
FOR	161	18.3	14.1	0.857	0.838	0.022
JON	187	16.5	12.8	0.838	0.824	0.016
WE	78	11.0	10.2	0.763	0.770	−0.010

slightly reduced in Weweantic ($H_O = 0.765$). Allelic richness ranged from 10.2 in Weweantic to 14.7 in Salmon Falls, and differed significantly across the populations ($F = 5.25$, d.f. = 17, $p < 0.0001$). Allelic richness was significantly lower in the Weweantic sample than in all others except the sample from East Bay Brook in Cobscook Bay (hereafter Cobscook). Allelic richness of the Cobscook sample was also reduced relative to Salmon Falls. The remaining populations exhibited similar levels of allelic richness.

Bottleneck tests showed little evidence that smelt from Cobscook, Jones or Weweantic (sampling locations at the geographic extremes of our study area) had undergone a recent severe reduction in population size. Wilcoxon signed-rank tests detected significant heterozygote excess ($p < 0.01$) under the infinite alleles model for each of these populations. In seeming contradiction, however, they demonstrated significant heterozygote deficit under the step-wise mutation model and also for Weweantic under the two-phase mutation model. All three populations exhibited normal L-shaped allele frequency distributions (data not shown). No evidence of bottleneck effects were detected by the M-ratio tests, as estimated M values (0.7013, 0.8698, and 0.8122, for Cobscook, Jones and Weweantic, respectively) were well above their critical M values (0.5508–0.5812). Taken as a whole, these results suggest that these populations have not experienced a detectable genetic bottleneck in recent times.

3.2. Population genetic structure

Significant population structuring was detected across the 18 sampled rivers, with a global F_{ST} of 0.0144. A total of 108 significant pairwise comparisons were detected out of 153 tests, corresponding to F_{ST} values from 0.002 to 0.082 (see Appendix, Table A1). The highest levels of genetic differentiation occurred in comparisons of smelt from the geographic extremes of the sampled range (Jones and Weweantic Rivers in the south and Cobscook in the north) as well as in comparisons of these sites with all other sample locations. The five river samples from the Great Bay estuary had high genetic similarity (F_{ST}s ranged from −0.006 to 0.0024), as did several populations along the southern Maine coast.

In support of an isolation-by-distance (IBD) model, Mantel tests identified significant correlations between genetic and geographic distances, both with ($R^2 = 0.27$, $p = 0.0006$) and without ($R^2 = 0.47$, $p < 0.0001$) inclusion of the Cape Cod Canal as a migration corridor between Cape Cod and Buzzards Bays (Fig. 2). Spatial autocorrelation analyses detected greater fine-scale spatial genetic structure when distances were estimated using the canal: significant positive spatial genetic structure was detected in the 50, 100 and 150 km distance classes in the analysis with the canal, and in the 50 and 100 km distance classes in the analysis without the canal (Fig. 3). Further, in the analysis without the canal, elevated correlations
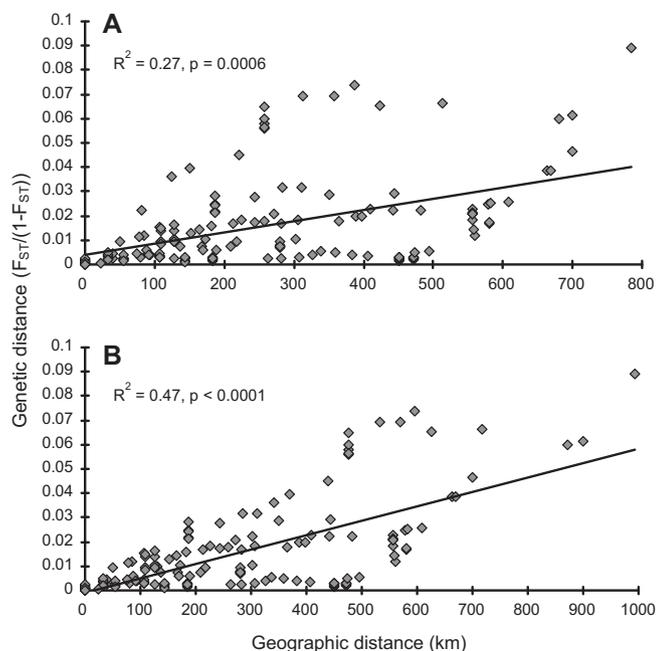
Fig. 2. Mantel tests of isolation by distance for rainbow smelt sampled from 18 river systems. Linearized genetic distance ($F_{ST}/(1 - F_{ST})$) is plotted vs. coastal geographic distance (5 km from shore) with (A) and without (B) inclusion of the Cape Cod Canal as a possible migration corridor.

Massachusetts Bays. Similar x intercepts (189 and 183 km) were identified with and without the canal.

Population clustering analyses with STRUCTURE suggested that $K = 5$ was optimal according to both the Pritchard et al. (2000) and Evanno et al. (2005) methods. Smelt sampling locations were found to be informative for identifying genetic structure using the LocPrior model, as evidenced by the low value for r (the parameter that estimates the informativeness of the sampling location data; $r = 0.46$ averaged across 25 runs at $K = 5$). Values of r close to or less than one indicate that the inclusion of sampling locations is informative, while values of $r \gg 1$ imply that location data is uninformative when inferring ancestry (Hubisz et al., 2009). The following 5 population clusters were identified, from North to South: (1) Cobscook; (2) Penobscot; (3) the Chandler, Pleasant, Marsh/Sheepscot, Kennebec, Harraseeket and Long samples of Maine, the 5 rivers of Great Bay NH, and Parker River (ME-NH-PAR cluster); (4) Saugus and Fore Rivers of Massachusetts Bay; and (5) Jones and Weweantic Rivers (Fig. 4A). Smelt from most sample locations were assigned to a single cluster with high proportion of membership (>90%), with the following exceptions: smelt from the Marsh and Sheepscot Rivers (MAR) shared membership with the ME-NH-PAR cluster (61%) and the Penobscot cluster; membership of Parker River smelt was split between ME-NH-PAR (59%) and the SAU-FOR cluster, and smelt from the Jones River had 88% membership to JON-WE and the remainder to the SAU-FOR cluster. Results of the spatial clustering analysis with BAPS (not shown) identified only 4 distinct population clusters as follows: (1) Cobscook; (2) the remaining rivers in Maine including the Penobscot, the 5 NH rivers, and Parker River; (3) Saugus, Fore and Jones rivers; (4) Weweantic River. For comparison with the BAPS results, we considered the result of the STRUCTURE analysis at $K = 4$ (representative run shown in Fig. 4B), which differed by grouping the Weweantic with the Jones and Fore rivers into a single cluster and retaining Penobscot as distinct. Due to difficulty in identifying population structure in the presence of isolation by distance, we also considered the results of analyses with $K = 6$ (representative run shown in Fig. 4C), for which the bar plots showed another potentially biologically meaningful pattern. This analysis identified the same 5 clusters as for $K = 5$, plus
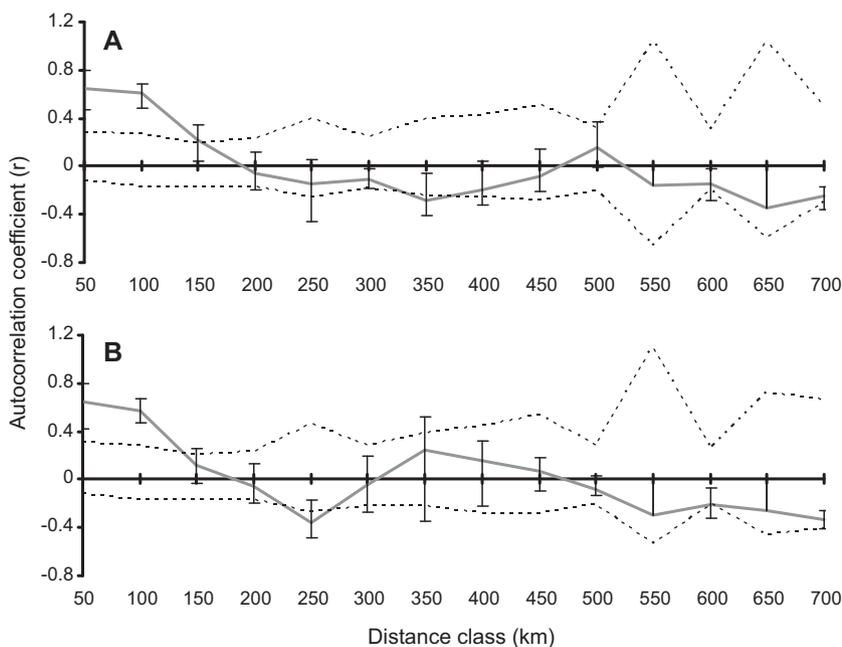
were evident at the 350 and 400 km distance classes, corresponding to the geographic distances separating the Weweantic from the Jones, Saugus and Fore Rivers following the coastline of Cape Cod. Higher genetic correlations among smelt in these rivers (separated by 80–150 km via the canal or 300–370 km around Cape Cod) relative to smelt separated by the 200–300 km or >400 km distances could drive the observed spatial patterns and indicate a role of the Cape Cod Canal in connecting smelt from Buzzards and

Fig. 3. Correlogram showing spatial correlation of individual genetic distance as a function of geographic distance for rainbow smelt sampled from 18 river systems with (A) and without (B) inclusion of the Cape Cod Canal as a possible migration corridor. Dotted lines indicate 95% confidence intervals about a null hypothesis of a random distribution of smelt; error bars determined by bootstrapping.
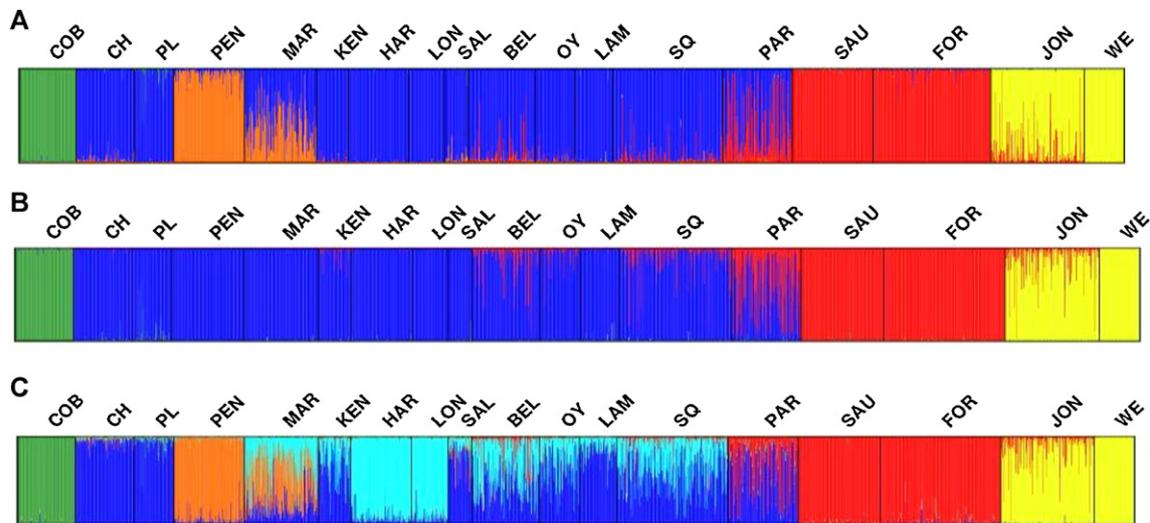
Fig. 4. Membership, as identified by STRUCTURE analysis, of rainbow smelt from 18 river systems to (A) $K = 5$, (B) $K = 4$, and (C) $K = 6$ genetically similar clusters. Each line represents the proportional assignment of an individual to the clusters, represented by the different colors. Abbreviated site codes refer to Table 1.

a 6th cluster that included the two sampling locations from Casco Bay (HAR and LON), with admixture with the neighboring Sheepscot/Kennebec Complex and Great Bay samples. While we see value in considering the results from $K = 4$–6, we limit further analyses to the optimal clusters defined by each of the two methods, BAPS and STRUCTURE.

Results of assignment tests were variable among rivers (10.1–84.6% correctly assigned; Table 4) and overall 31.2% individuals were correctly assigned to their river of origin (quality index of 26.7% for self-assignments). Self assignments to individual rivers were significantly greater than expected by chance ($\chi^2 = 2932$, d.f. = 13, $p < 0.0001$). The lowest percentage of correct assignments were for the rivers within the ME-NH-PAR cluster (10.1–26.0%), while the highest percentage of correct assignments were for Jones River, Cobscook Bay and Weweantic River (61%, 73.2% and 84.6%, respectively). Bay-level assignments also varied, but were considerably stronger; percentage of correct assignments ranged from 33% for the Great Bay rivers to 92.3% for Weweantic, with 38.2% correctly assigned overall (32.2% quality index), which was significantly greater than expected by chance ($\chi^2 = 6570$, d.f. = 10, $p < 0.0001$). Percentage of correct assignments to the population

Table 4
Percentage of correct assignments of rainbow smelt to four spatial groupings: 14 river systems (the 5 New Hampshire rivers that feed Great Bay estuary are combined into a single NH sample), 11 bays, the 5 STRUCTURE clusters, and 4 BAPS clusters. Abbreviated sample names refer to Table 1 ME-NH-PAR refers to the cluster consisting of most Maine sampling sites except Cobscook and Penobscot, the 5 New Hampshire rivers and Parker River (see text).

River system		Bay level		STRUCTURE ($K = 5$)		BAPS ($K = 4$)	
Pop	%	Pop	%	Pop	%	Pop	%
COB	73.2	COB	86.8	COB	84.2	COB	86.0
CH	18.8	CH	52.6	PEN	57.4	ME-NH-PAR	79.8
PL	10.1	PL	58.2	ME-NH-PAR	60.8	SAU-FOR-JON	78.0
PEN	38.6	PEN	71.0	SAU-FOR	71.1	WE	85.9
MAR	22.6	MAR-KEN	48.6	JON-WE	76.6		
KEN	14.1	HAR-LON	60.6				
HAR	20.8	NH	33.0				
LON	17.8	PAR	49.6				
NH	18.2	SAU-FOR	68.4				
PAR	18.7	JON	80.0				
SAU	35.4	WE	92.3				
FOR	33.2						
JON	61.0						
WE	84.6						

clusters identified by STRUCTURE (5 clusters) and BAPS (4 clusters) ranged from 57.4% to 84.2% for the STRUCTURE groupings and 78–86% for the BAPS groupings; overall 67.8% and 79.8% of individuals were correctly assigned for the two groupings, respectively (quality indices of 61.7% and 68.1%). Correct assignments to population cluster of origin were significantly greater than expected by chance ($\chi^2 = 2371$, d.f. = 4, $p < 0.0001$) and ($\chi^2 = 2665$, d.f. = 3, $p < 0.0001$ for the two groupings, respectively).

AMOVA detected significant among group variation (1.49–1.83%) relative to within group variation (0.06–0.44%; Table 2) for the three hierarchical population groupings (bay-level and cluster-level). Highest F_{CT} (among group differentiation) and percent variation were explained by the 4-cluster BAPS grouping.

4. Discussion

Defining population structure and units of management should be relatively straight forward in philopatric species, such as many anadromous fish, but may be challenging when gene flow is high (McLean and Taylor, 2001). We found that genetic variation in rainbow smelt was neither structured at the river-level nor consistently at the level of individual estuaries nor bays. Rather, smelt from 18 rivers draining to 11 bays in the northeastern U.S. were structured broadly into four to six genetically distinct groupings that varied in geographic extent. Some rivers and bays had distinct local genetic signatures, while others were connected by gene flow across large portions of the study area. Our findings highlight that variation exists for anadromous species in the spatial scale of philopatry and that knowledge of population genetic structure facilitates defining population units.

4.1. Spatial scale of population structure

At the bay level, genetically divergent populations of smelt were found in Cobscook, Massachusetts and Buzzards Bays, with smelt from Plymouth Bay admixed with those from Massachusetts and Buzzards Bays. Gene flow was high among smelt sampled in most locations from Englishman Bay to Plum Island Sound, with the exception of Penobscot and to a lesser degree Casco Bay, which each showed divergence within this long stretch of connected coastline. Although there was not complete congruence in the results of the two clustering methods we used, they were largely consistent and differed only with respect to the genetic distinctiveness

of Penobscot Bay and the level of admixture between Plymouth and Buzzards Bays. Assignment tests and AMOVA supported both spatial groupings and the strong genetic distinctiveness of Buzzards Bay (Weweantic), while indicating that Penobscot was less divergent than the other four clusters. AMOVA also provided some support for bay-level structuring, although the variation among bays was lower than that among clusters. Bay-level assignment tests gave further support to population structuring at the level of Cobscook, Penobscot, Casco, Massachusetts, Plymouth and Buzzards Bays and confirmed high connectivity among Englishman, Pleasant, Sheepscot/Kennebec, Great, and Ipswich Bays.

At the level of individual rivers, variation in genetic heterogeneity was also apparent. Some rivers were widely connected to others across a few hundred kilometers (e.g. Parker, Pleasant, Chandler and the rivers of the Great Bay estuary) and others were differentiated from even their nearest neighbors (e.g. Weweantic, Jones, Cobscook, and Penobscot). Both AMOVA and assignment tests indicated weak, but significant and temporally stable, structuring at the river level. While individual fish were assigned to their river of origin more often than by chance, the percentage of correct assignments was <30% for many rivers, and only Weweantic, Jones, and Cobscook had strong river-specific signatures (>60% correct self-assignments). Each of these rivers was the only river sampled in its respective bay and, therefore, the patterns of genetic distinctiveness are likely an attribute of these bays rather than any single river. Misassignments of individuals were typically highest to neighboring rivers (data not shown) and the percentage of correct assignments was much higher to individual bays than rivers, further supporting structuring at a bay-scale or higher level.

Despite concerns about declining smelt populations (Chase and Childs, 2001; Chase, 2009), we did not find evidence of significant genetic consequences of population reduction. Overall, genetic diversity was high for most rivers, but reduced for smelt sampled from Weweantic and Cobscook relative to other locations. Reduced diversity in Weweantic is consistent with its location at the most southern extent of the species' current range. Populations at the periphery of a species' geographical range often have lower genetic variation, resulting from reduced gene flow and lower effective population sizes (Schwartz et al., 2003). Reduced genetic variation in smelt from Cobscook Bay may result from demographic rather than geographic isolation of this population, which is located within the continuous distribution of smelt along coastal, mainland North America. The highly structured coastal geography and enclosed position of Cobscook Bay may limit connectivity with other populations (see Section 4.2). Despite these modest reductions in genetic diversity, genetic bottleneck tests found no consistent evidence that any population had experienced a recent severe decline. Nonetheless, we cannot rule out the possibility that the negative findings are an artifact of the sensitivity of the methods we used. These classical bottleneck tests are based on the patterns of allele loss during demographic bottlenecks, but due largely to simplifying assumptions, they have limitations in their power to accurately reveal recent population declines (Hoffman et al., 2011). The genetic distinctiveness of the smelt from these locations, combined with their reduced diversity, may warrant consideration and further monitoring to track population persistence.

Across the study area, genetic variation in smelt followed an IBD pattern (Wright, 1943), such that in general genetic differentiation increased with increasing geographic distance. An IBD pattern signifies that gene flow is sufficient to connect adjacent populations and prevent the formation of isolated demes, but that long-distance dispersal is rare enough that complete genetic homogenization does not occur (Slatkin, 1993). An IBD pattern of population structure results from a stepping stone model of dispersal (Kimura, 1953), in which gene flow occurs most prevalently among neighboring populations, consistent with previous reports

of adult smelt straying frequently and most often to nearby rivers (Murawski et al., 1980; Frechet et al., 1983). Spatial autocorrelation analyses indicated that the spatial extent of positive genetic structure is approximately 180 km. This distance represents the maximal extent of population connectivity; beyond 180 km, gene flow is limited and genetic drift becomes more influential. Spatial autocorrelation results are often used to infer the effective spatial extent beyond which continuous populations are independent (Diniz-Filho and Telles, 2002). The observed patterns of gene flow in this study suggest that the spatial extent of population connectivity for smelt varies across the study area and, therefore, no single geographic distance can be used to approximate the scale of demographic independence.

The level of genetic differentiation found in rainbow smelt in this study (mean $F_{ST} = 0.015$; range 0–0.08) is similar to that found in smelt from mainland Canada (mean $F_{ST} = 0.017$ across New Brunswick and Nova Scotia), but an order of magnitude lower than the differentiation in smelt from Newfoundland (mean $F_{ST} = 0.11$; Bradbury et al., 2006a). The weak genetic differentiation found in smelt from mainland Canada and this study suggests that gene flow is high in these regions and philopatry is not confined to a local (river-level) scale. This level of gene flow is also higher than that found in Atlantic salmon (*Salmo salar*) in this region, for which population differentiation occurs among and within individual rivers (Spidle et al., 2003; Dionne et al., 2009). Comparably low levels of differentiation have been found in a few other anadromous species, indicative of broad-scale structuring, e.g. pink salmon (*Oncorhynchus gorbusha*; Olsen et al., 1998, 2000), eulachon (*Thaleichthys pacificus*; McLean and Taylor, 2001; Beacham et al., 2005), and American shad (*Alosa sapidissima*; Waters et al., 2000).

Many factors may account for the differences in population structure among anadromous species, including environmental and life history differences, and the migration distance and duration of residence in freshwater. Spawning sites located near to the coast and a relatively short freshwater duration may be important in contributing to patterns of broad-scale structure. Pink salmon, for example, generally spawn in freshwater close to the sea or in the intertidal zone, and fry migrate to sea soon after emergence (Bonar et al., 1989). Accordingly, pink salmon populations are less structured than are populations of several other salmonids that spawn farther inland and spend a longer time in freshwater (Allendorf and Waples, 1996; Quinn, 2005; Avise, 2006). A direct comparison can be made between smelt and eulachon, a Pacific osmerid with a similar life history to rainbow smelt (McLean and Taylor, 2001; Beacham et al., 2005). Both smelt and eulachon spawn within tidally influenced portions of the river and, upon hatching, their larvae are quickly washed to estuarine or coastal waters. McLean et al. (1999) hypothesized that, for eulachon, the brief period of time spent in freshwater precludes imprinting to individual rivers. Instead, imprinting and subsequent homing may occur to a coastal area (an estuary, bay or coastal retention zone) that is shared by early life stage fish from many spawning sources (nearby river systems). The high levels of gene flow we observed across the majority of our study area suggest that rainbow smelt also home to larger coastal areas rather than individual river systems.

4.2. Influence of hydrographic conditions

Population structure results from the interplay between philopatry and dispersal: homing promotes philopatry and reproductive isolation (Horrall, 1981), resulting in limited gene flow and the accumulation of genetic differences among populations, while dispersal (of all life stages) homogenizes gene flow. In smelt, population structure may be more constrained by the processes that influence the dispersal of the early life stages than by the movement abilities of the adults (Baby et al., 1991; Bernatchez

and Martin, 1996; Bradbury et al., 2006a,b), consistent with the member-vagrant hypothesis (Iles and Sinclair, 1982). This hypothesis leads to the prediction that the number of genetically distinct populations directly follows from the number of distinct larval retention zones (geographic, physical or oceanographically distinct areas for larval retention; Baby et al., 1991). Variation in gene flow in smelt observed in this study may be explained by patterns of larval dispersal and retention across the study area.

Smelt larval dispersal is determined by an interaction of passive (primarily hydrographic) and active (swimming behavior) processes (Bradbury et al., 2006b). Seasonal circulation patterns in the estuary are influenced by the amount of freshwater influx and surface water export, which in turn vary with wind patterns and river inflow. Topography also influences circulation through its effect on the location of frontal formation within the estuary (O'Donnell, 1993). Smelt larvae use active vertical migration, by which they change their swimming depth in response to tidal flow, thereby enhancing their retention within the favorable environment of the high turbidity zone of the estuary (Laprise and Dodson, 1989; Dauvin and Dodson, 1990; Sirois and Dodson, 2000). Nonetheless, the combination of the passive forces, which vary seasonally and geographically, largely determine whether young larvae are retained locally or washed out to coastal waters beyond the outer reaches of the estuary. For example, onshore winds and coastal topography effectively retain smelt larvae within estuaries in Newfoundland, while a topographically less complex coastline and lack of consistent onshore winds in mainland Canada promote widespread larval dispersal in this region (Bradbury et al., 2006a,b). These divergent larval dispersal patterns are manifest in differences in gene flow (Bradbury et al., 2006a). Regional differences in hydrographic patterns, therefore, likely drive regional differences in smelt population structure.

The geography of the northern New England coast is more similar to that of mainland Canada than Newfoundland, with large stretches of continuous and topographically simple coastline and few well-defined inlets or bays. Coastal circulation within this region is driven by the Gulf of Maine Coastal Current (GMCC), which has a cyclonic (counter-clockwise) circulation that is prominent in the summer months when smelt larvae are transported into estuaries and coastal waters (Pettigrew et al., 1998, 2005). The Eastern Maine Coastal Current (EMCC) flows from the Bay of Fundy southwest along the coast and, in the area of Penobscot Bay, often splits southward toward Jordan Basin depending on annual climatic conditions. The remaining portion of the EMCC combines with outflow from Penobscot Bay and continues southwestward toward coastal Massachusetts, creating the Western Maine Coastal Current (WMCC; Pettigrew et al., 1998, 2005). This circulation pattern favors downstream, southwest transport of larvae, rather than local retention in most areas of the coast (Huret et al., 2007; Xue et al., 2008). Models estimating transport of lobster (*Homerus americanus*) larvae found that areas west of Penobscot Bay down to Cape Anne received large percentages of larvae from eastern Maine, whereas Penobscot Bay, the Bay of Fundy, and the area south of Cape Cod had relatively high retention and little reception of larvae from other areas (Xue et al., 2008). Similarly, the WMCC transports Atlantic cod (*Gadus morhua*) larvae from coastal Maine to Ipswich and Massachusetts Bays (Huret et al., 2007). Massachusetts Bay maintains high retention as the strength of the coastal circulation pattern has largely diminished by this point (Incze et al., 2010).

The patterns of genetic variation we observed in smelt are largely consistent with the circulation patterns of the Gulf of Maine. The most strongly differentiated sites were associated with areas where circulation favors high larval retention and low reception from other regions, while the most broadly connected sites were influenced largely by the southwest transport patterns. For example, the mouth of Cobscook Bay remains relatively outside of the GMCC circulation pattern, enabling high local larval retention within the bay, consistent with its strong genetic distinctiveness. Conversely, larvae flushing out of the Chandler and Pleasant Rivers are swept southwest along the coast with the EMCC, consistent with the high gene flow observed from Englishman Bay to Plum Island Sound. High outflow from the Penobscot River may minimize reception of larvae from easterly sites into Penobscot Bay. This low reception, combined with a more complex topography in the bay that promotes high larval retention, is consistent with the genetic distinctiveness of the Penobscot sample. In the same way, high similarity between the sites within Casco Bay (Harraseeket and Long) and their reduced connectivity with most other sites except neighboring rivers of the Great Bay and Sheepscot/Kennebec Complex is likely due to the local hydrographic and topographic features. High outflow from the Kennebec and Androscoggin Rivers deflects the WMCC and thereby minimizes transport of eastern larvae offshore into Casco Bay; additionally, the partially enclosed structure of this bay likely facilitates some level of larval retention. Surface-drifter studies by Janzen et al. (2005) suggested a role of circulation patterns in promoting retention of surface-dwelling organisms in Casco Bay. Lastly, genetic distinctiveness of the smelt from Massachusetts Bay (Saugus and Fore), Plymouth Bay (Jones) and Buzzards Bay (Weweantic) is consistent with the predictions of high local retention from the lobster larval transport models described above.

While larval retention occurs in the bays described above, it is incomplete, and observed patterns of admixture are likely a result of additional local hydrographic influences. The admixture of the Marsh and Sheepscot River sample with the Penobscot sample may be the result of westward transport of Penobscot larvae as outflow from the Penobscot River combines with the larger near-shore current. The portion of the WMCC contributed by outflow from Penobscot River is redirected southward, away from the coast, at the location of the Sheepscot/Kennebec Complex by high outflow from the Kennebec and Androscoggin Rivers (Pettigrew et al., 1998). This circulation pattern likely minimizes further westward larval transport, consistent with a lack of genetic connectivity of the Penobscot samples beyond the Sheepscot/Kennebec Complex. Admixture of the Parker River sample, and to a lesser extent a few of the Great Bay samples with the Saugus-Fore grouping, may be a function of northeasterly larval transport out of Massachusetts Bay against the dominant current in near-shore back-flow eddies associated with river plumes (Brooks, 1994). Connectivity between smelt in the Jones and Weweantic Rivers, despite a strong divergence of the Weweantic smelt and the significant spatial barrier posed by Cape Cod, suggest that the Cape Cod Canal may function as a corridor that facilitates infrequent migration of smelt between Plymouth and Buzzards Bays. Other marine fish use the canal, and smelt may enter it if they encounter the strong tidal currents at either canal entrance (B. Chase, pers. comm.). Finally, admixture between rivers could also result from historical stocking efforts by fishermen. Any such efforts are thought to have been small scaled, however, although no records exist documenting their extent (G. Wippelhauser, pers. comm.).

Overall, the population genetic structure of smelt in this study is consistent with local hydrographic and coastal circulation patterns and our findings lend support to the hypothesized greater importance of larval retention processes than homing (Bradbury et al., 2008a). Additional information about adult movements from tagging and mark-recapture studies is needed to test this hypothesis, and this should be a focus of future research to more thoroughly evaluate the role of adults in influencing population structure. Recently, Banks et al. (2007) suggested that patterns of oceanography and coastal geography might be useful in defining the optimal scale of fisheries management at a regional level. To this end, additional data are needed to evaluate whether there is a common

influence of the hydrographic patterns in this region on the structure of other anadromous and marine fish species with near coastal spawning locations. Kovach et al. (2010) found evidence that gene flow in Atlantic cod in the western Gulf of Maine is consistent with ocean circulation patterns, although genetic data from the eastern Gulf of Maine were lacking in their study. Future efforts should consider hydrographic patterns in evaluating population genetic structure of other species in this region.

4.3. Management implications

Dionne et al. (2009) summarized the consequences of management at the inappropriate spatial scale for Atlantic salmon (*Salmo salar*); these considerations are equally relevant to rainbow smelt. In brief, management at too large a scale assumes that gene flow and effective population sizes are higher than they are in the presence of population structure. Persistence of small and genetically distinct populations may be at greater risk due to stochastic processes, and these populations may be more susceptible to loss of genetic variation. Management at too large a scale also ignores the potential importance of local adaptation and underestimates the impact of fishing. Conversely, management at too fine of a spatial scale, e.g. at the level of individual rivers that are in fact demographically connected, ignores the role of gene flow in promoting population persistence and may result in inefficient resource use and superfluous management activities. Management plans tailored to the pattern of gene flow can capitalize on the understanding that management within one component of the population unit may benefit the other components, thereby reducing the need for activities at each individual river.

Given the above considerations, the findings of our study can inform management activities, including stocking efforts, for rainbow smelt. Our results show that population connectivity is consistent with the coastal circulation patterns and that the scale of demographic independence exceeds the scale of most individual rivers and some bays. These findings suggest that the appropriate scale of management may be defined by local hydrographic processes, and may vary spatially, including multiple neighboring river systems feeding one or more bays. Conservatively, we recommend a scale not broader than the bay-level, with priority focus on the most hydrographically distinct bays in the region. Although we found evidence of weak structuring at the level of individual rivers, a river-specific approach to management might neither be practical, nor necessary in all cases, especially where several geographically proximate rivers feed the same estuary. River-specific management might be warranted in special circumstances, such as where concerns exist over local declines – for example, the Weweantic River, which, as the southern most population within the species' current range, has limited gene flow and a potentially low local effective population size. Additional data on adult movements and straying rates are needed, however, before firm conclusions can be made about management at the level of individual rivers. Our study illustrates the utility of genetic monitoring approaches for defining management units and conservation priorities (Schwartz et al., 2006; Dionne et al., 2009). On the other hand, neutral genetic data alone may be insufficient in setting management guidelines in the absence of data on ecological, life history, and adaptive processes (Crandall et al., 2000; Waples, 1998). Local adaptation, for example, may occur even in the face of moderate ongoing gene flow (Garant et al., 2007). For this reason, future management plans should incorporate our genetic findings with forthcoming information on population characteristics, including data currently being collected on spawning time, growth curves, population dynamics and survival, as well as additional data on the ecology of smelt, including especially information on adult movements.

Table A1
Genetic differentiation of adult rainbow smelt from 18 river systems in the northeastern U.S. Pairwise F_{ST} values are displayed below the diagonal with **bold** indicating significance following standard Bonferroni correction ($p \leq 0.0003$). Symbols *, **, and *** indicate significance at the 0.05, 0.01, and 0.001 levels, respectively. Abbreviated site names refer to Table 1.

	COB	CH	PL	PEN	MAR	KEN	HAR	LON	SAL	BEL	OY	LAM	SQ	PAR	SAU	FOR	JON	WE
COB	–																	
CH	0.0118	–																
PL	0.0149	-0.0004	–															
PEN	0.0204	0.0057	0.0072	–														
MAR	0.0176	0.0025	0.0025	0.0037	–													
KEN	0.0222	0.0039	0.0029**	0.0096	0.0044	–												
HAR	0.0218	0.0051	0.0055	0.0102	0.0030	-0.0003	–											
LON	0.0219	0.0036***	0.0037***	0.0095	0.0025**	0.0045***	-0.0003	–										
SAL	0.0224	0.0028	0.0014	0.0072	0.0030	0.0026	0.0036	0.0019	–									
BEL	0.0208	0.0027***	0.0028***	0.0069	0.0021***	0.0017**	0.0033***	0.0025**	0.0014	–								
OY	0.0178	0.002***	0.0018***	0.0078	0.0024**	0.0024**	0.0033***	0.0023*	0.0018	0.0012*	–							
LAM	0.0204	0.0025***	0.0024*	0.0095	0.0023*	0.0009*	0.0042	0.0037**	0.0024	0.0005	0.0001	–						
SQ	0.0203	0.0030	0.0028***	0.0072	0.0025	0.0032***	0.0030	0.0016*	0.0012	-0.0006	0.0001	0.0002	–					
PAR	0.0240	0.0054	0.0048	0.0103	0.0073	0.0059	0.0074	0.0059	0.0033***	0.0025	0.0038	0.0048**	0.0020	–				
SAU	0.0373	0.0165	0.0119	0.0194	0.0165	0.0169	0.0167	0.0140	0.0092	0.0093	0.0135	0.0144	0.0083	0.0042	–			
FOR	0.0370	0.0168	0.0143	0.0196	0.0180	0.0176	0.0182	0.0154	0.0092	0.0102	0.0137	0.0161	0.0095	0.0039	0.0004*	–		
JON	0.0443	0.0253	0.0245	0.0286	0.0280	0.0309	0.0305	0.0270	0.0242	0.0213	0.0239	0.0274	0.0208	0.0126	0.0112	0.0095	–	
WE	0.0818	0.0577	0.0567	0.0623	0.0612	0.0686	0.0648	0.0649	0.0549	0.0531	0.0565	0.0607	0.0533	0.0429	0.0379	0.0350	0.0219	–

Acknowledgements

Appendix A.

Table A1.

References

Allendorf, F.W., Waples, R.S., 1996. Conservation and genetics of salmonid fishes. In: Avise, J.C., Hamrick, J.L. (Eds.), Conservation Genetics: Case Histories From Nature. Chapman and Hall, New York, pp. 238–275.

Avise, J.C., 2006. Molecular Markers, Natural History, and Evolution, 2nd ed. Sinauer Associates, Inc., Sunderland, MA.

Baby, M.C., Bernatchez, L., Dodson, J.J., 1991. Genetic structure and relationships among anadromous and landlocked populations of rainbow smelt, Osmerus mordax, Mitchill, as revealed by mtDNA restriction analysis. J. Fish Biol. 39, 61–68.

Banks, S.C., Piggott, M.P., Williamson, J.E., Bove, U., Holbrook, N.J., Beheregaray, L.B., 2007. Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. Ecology 88, 3055–3064.

Beacham, T.D., Hay, D.E., Le, K.D., 2005. Population structure and stock identification of eulachon (Thaleichthys pacificus), an anadromous smelt, in the Pacific Northwest. Mar. Biotechnol. 7, 363–372.

Beacham, T.D., Spilsted, B., Le, K.D., Wetklo, M., 2008. Population structure and stock identification of chum salmon (Oncorhynchus keta) from British Columbia determined with microsatellite DNA variation. Can. J. Zool. 86, 1002–1014.

Bernatchez, L., Martin, S., 1996. Mitochondrial DNA diversity in anadromous rainbow smelt, Osmerus mordax: a genetic assessment of the member-vagrant hypothesis. Can. J. Fish. Aquat. Sci. 53, 424–433.

Bohonak, A.J., 1999. Dispersal, gene flow and population structure. Q. Rev. Biol. 74, 21–45.

Bonar, S.A., Pauley, G.B., Thomas, G.L., 1989. Species profiles: life histories and environmental requirements of coastal fishes and invertebrates (Pacific Northwest) – pink salmon. U.S. Fish Wildl. Serv. Biol. Rep. 82 (11.88). U.S. Army Corpos of Engineers, TR EL-82-4, 18 pp.

Bradbury, I.R., Coulson, M.W., Campana, S.E., Bentzen, P., 2006a. Morphological and genetic differentiation in anadromous smelt Osmerus mordax (Mitchill): disentangling the effects of geography and morphology on gene flow. J. Fish Biol. 69 (Suppl. C), 95–114.

Bradbury, I.R., Gardiner, K., Snelgrove, P.V.R., Campana, S.E., Bentzen, P., Guan, L., 2006b. Larval transport, vertical distribution, and localized recruitment in anadromous rainbow smelt (Osmerus mordax). Can. J. Fish. Aquat. Sci. 63, 2822–2836.

Bradbury, I.R., Campana, S.E., Bentzen, P., 2008a. Low genetic connectivity in an estuarine fish with pelagic larvae. Can. J. Fish. Aquat. Sci. 65, 147–158.

Bradbury, I.R., Campana, S.E., Bentzen, P., 2008b. Estimating contemporary early life-history dispersal in an estuarine fish: integrating molecular and otolith elemental approaches. Mol. Ecol. 17, 1438–1450.

Brooks, D.A., 1994. A model study of the buoyancy-driven circulation in the Gulf of Maine. J. Phys. Oceanogr. 24, 2387–2412.

Buckley, J., 1989. Species profile: life histories and environmental requirements of coastal fishes and invertebrates. Rainbow Smelt. USFWS Biol. Rept. U.S. Army Corps of Engineers, TR EL-82-4, 11 pp.

Cadrin, S.X., Secor, D.H., 2009. Accounting for spatial population structure in stock assessment: past, present and future. In: Beamish, R.J., Rothschild, B.J. (Eds.), The Future of Fisheries Science in North America. Fish & Fisheries Series, vol. 31. Springer, pp. 405–426.

Chase, B.C., Childs, A.R., 2001. Rainbow smelt (Osmerus mordax) spawning habitat in the Weymouth-Fore River. Massachusetts Division of Marine Fisheries Technical Report TR-5.

Chase, B.C., 2009. The spawning habitat of anadromous rainbow smelt: trouble at the tidal interface. Am. Fish. Soc. Symp. 69, 859–862.

Collette, B.B., Klein-MacPhee, G., 2002. Fishes of the Gulf of Maine. Smithsonian Institution Press, Washington, 748 pp.

Corander, J., Sirén, J., Arjas, E., 2008. Bayesian spatial modeling of genetic population structure. Comput. Stat. 23, 111–129.

Costello, A.B., Down, T.E., Pollard, S.M., Pacas, S.J., Taylor, E.B., 2003. The influence of history and contemporary stream hydrology on the evolution of genetic diversity within species: an examination of microsatellite DNA variation in bull trout, Salvelinus confluentus (Pisces: Salmonidae). Evolution 57, 328–344.

Coulson, M.W., Patterson, I.G., Green, A., Kepkay, R., Bentzen, P., 2006. Characterization of di- and tetranucleotide microsatellite markers in rainbow smelt (Osmerus mordax). Mol. Ecol. Notes 6, 942–944.

Crandall, K.A., Bininda-Emonds, O.R.P., Mace, G.M., Wayne, R.K., 2000. Considering evolutionary processes in conservation biology. Trends Ecol. Evol. 15, 290–295.

Dauvin, J.C., Dodson, J.J., 1990. Relationship between feeding incidence and vertical and longitudinal distribution of rainbow smelt larvae (Osmerus mordax) in a turbid well-mixed estuary. Mar. Ecol. Prog. Ser. 60, 1–12.

Diniz-Filho, J.A.F., Telles, M.P.D., 2002. Spatial autocorrelation analysis and the identification of operational units for conservation in continuous populations. Conserv. Biol. 16, 924–935.

Dionne, M., Miller, K.M., Dodson, J.J., Bernatchez, L., 2007. Clinal variation in MHC diversity with temperature: evidence for the role of host–pathogen interaction on local adaptation in Atlantic salmon. Evolution 61, 2154–2164.

Dionne, M., Caron, F., Dodson, J.J., Bernatchez, L., 2008. Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. Mol. Ecol. 17, 2832–2896.

Dionne, M., Caron, F., Dodson, J.J., Bernatchez, L., 2009. Comparative survey of within-river genetic structure in Atlantic salmon; relevance for management and conservation. Conserv. Genet. 10, 869–879.

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

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164, 1567–1587.

Francois, O., Durand, E., 2010. Spatially explicit Bayesian clustering models in population genetics. Mol. Ecol. Resour. 10, 773–784.

Frechet, A., Dodson, J.J., Powles, H., 1983. Use of variation in biological characters for the classification of anadromous rainbow smelt (Osmerus mordax) groups. Can. J. Fish. Aquat. Sci. 40, 718–727.

Garant, D., Forde, S.E., Hendry, A.P., 2007. The multifarious effects of dispersal and gene flow on contemporary adaptation. Funct. Ecol. 21, 434–443.

Garza, J.C., Williamson, E.G., 2001. Detection of reduction in population size using data from microsatellite loci. Mol. Ecol. 10, 305–318.

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

Hansen, M.M., Ruzzante, D.E., Nielsen, E., Bekkevold, D., Mensberg, K.L.D., 2002. Long-term effective population sizes, temporal stability of genetic composition and potential for local adaptation in anadromous brown trout (Salmo trutta) populations. Mol. Ecol. 11, 2523–2535.

Hauser, L., Carvalho, G.R., 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish Fish. 9, 333–362.

Hoffman, J.I., Grant, S.M., Forcada, J., Phillips, C.D., 2011. Bayesian inference of a historical bottleneck in a heavily exploited marine mammal. Mol. Ecol. 20, 3989–4008.

Horrall, R.M., 1981. Behavioral stock-isolating mechanisms in Great Lakes fishes with special reference to homing and site imprinting. Can. J. Fish. Aquat. Sci. 38, 1481–1496.

Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K., 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. 9, 1322–1332.

Huret, M., Runge, J.A., Chen, C., Cowles, G., Xu, Q., Pringle, J.M., 2007. Dispersal modeling of fish early life stages: sensitivity with application to Atlantic cod in the western Gulf of Maine. Mar. Ecol. Prog. Ser. 347, 261–274.

Incze, L., Xue, H., Wolff, N., Xu, D., Wilson, C., Steneck, R., Wahle, R., Lawton, P., Pettigrew, N., Chen, Y., 2010. Connectivity of lobster (Homarus americanus) populations in the coastal Gulf of Maine. Part II. Coupled biophysical dynamics. Fish. Oceanogr. 19 (1), 1–20.

Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806.

Janzen, C.D., Churchill, J.H., Pettigrew, N.R., 2005. Observations of exchange between eastern Casco Bay and the western Gulf of Maine. Deep Sea Res. 52, 2411–2429.

Kimura, M., 1953. Stepping-stone model of population. Annu. Rep. Natl. Inst. Genet. 3, 62–63.

Kovach, A.I., Breton, T.S., Berlinsky, D.L., Maceda, L., Wirgin, I., 2010. Fine-scale spatial and temporal genetic structure of cod off the Atlantic coast of the USA. Mar. Ecol. Prog. Ser. 410, 177–195.

Iles, T.D., Sinclair, M., 1982. Atlantic herring: stock discreteness and abundance. Science 215, 627–633.

Laprise, R., Dodson, J.J., 1989. Ontogeny and importance of tidal vertical migrations in the retention of larval smelt Osmerus mordax in a well-mixed estuary. Mar. Ecol. Prog. Ser. 55, 101–111.

Latch, E., Dharmarajan, G., Glaubitz, J., Rhodes, O.E., 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conserv. Genet. 7, 295–302.

Lewis, P.O., Zaykin, D., 2001. Genetic data analysis: computer program for analysis of allelic data. Version 1.0 (d16c). http://lewis.eeb.uconn.edu/lewishome/software.html.

Luikart, G., Allendorf, F.W., Cornuet, J.-M., Sherwin, W.B., 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. J. Hered. 89, 238–247.

McLean, J.E., Hay, D.E., Taylor, E.G., 1999. Marine population structure in an anadromous fish: life-history influences patterns of mitochondrial DNA variation in the eulachon, Thaleichthys pacificus. Mol. Ecol. 8, S143–S158.

McLean, J.E., Taylor, E.B., 2001. Resolution of population structure in a species with high gene flow: microsatellite variation in the eulachon (Osmeridae: Thaleichthys pacificus). Mar. Biol. 139, 411–420.

Meeuwig, M.H., Guy, C.S., Kalinowski, S.T., Fredenberg, W.A., 2010. Landscape influences on genetic differentiation among bull trout populations in a stream-lake network. Mol. Ecol. 19, 3620–3633.

Murawski, S.A., Clayton, G.R., Reed, R.J., Cole, C.H., 1980. Movements of spawning rainbow smelt, Osmerus mordax, in a Massachusetts Estuary. Estuaries 3 (4), 308–314.

NOAA (National Oceanic and Atmospheric Administration), 2004. Species of concern in the northeast region. National Marine Fisheries Service, Proactive Conservation Program, Silver Spring, MD.

O'Donnell, J.O., 1993. Surface fronts in estuaries: a review. Estuaries 16, 12–39.

Olsen, J.B., Bentzen, P., Banks, M.A., Shaklee, J.B., Young, S., 2000. Microsatellites reveal population identity of individual pink salmon to allow supportive breeding of a population at risk of extinction. T. Am. Fish. Soc. 129, 232–242.

Olsen, J.B., Seeb, L.W., Bentzen, P., Seeb, J.E., 1998. Genetic interpretation of broadscale microsatellite polymorphism in odd-year pink salmon. T. Am. Fish. Soc. 127, 535–550.

Peakall, R., Ruibal, M., Lindenmayer, D.B., 2003. Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat Rattus fuscipes. Evolution 57, 1182–1195.

Peakall, R., Smouse, P.E., 2006. GENELEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.

Pettigrew, N.R., Townsend, D.W., Xue, H., Wallinsa, J.P., Brickley, P.J., Hetland, R.D., 1998. Observations of the Eastern Maine Coastal Current and its offshore extensions in 1994. J. Geophys. Res. 103, 30623–30640.

Pettigrew, N.R., Churchill, J.H., Janzen, C.D., Mangum, L.J., Signell, R.P., Thomas, A.C., Townsend, D.W., Wallinga, J.P., Xue, H., 2005. The kinematic and hydrographic structure of the Gulf of Maine Coastal Current. Deep Sea Res. II 52, 2369–2391.

Piry, S., Luikart, G., Cornuet, J.-M., 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J. Hered. 90, 502–503.

Piry, S., Alapetite, A., Cornuet, J.-M., Paetkau, D., Baudouin, L., Estoup, A., 2004. GeneClass2: a software for genetic assignment and first-generation migrant detection. J. Hered. 95, 536–539.

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

Quinn, T.P., 2005. The Behavior and Ecology of Pacific Salmon and Trout. University of Washington Press, Seattle, WA, USA.

Rannala, B., Mountain, J.L., 1997. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. U.S.A. 94, 9197–9201.

Raymond, M., Rousset, F., 1995. GENEPOP (version 1.2) population genetic software for exact tests and ecumenicism. J. Hered. 86, 248–249.

Reiss, H., Haloarau, G., Dickey-Collas, M., Wolff, W.J., 2009. Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish Fish. 10, 361–395.

Rice, W.R., 1989. Analyzing tables of statistical tests. Evolution 43, 223–225.

Rosenberg, N.A., 2004. DISTRUCT: a program for graphical display of population structure. Mol. Ecol. Notes 4, 137–138.

Schneider, S., Roessli, D., Excoffier, L., 2000. Arlequin: a software for population genetic data analysis. User manual ver. 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Geneva.

Schwartz, M.K., Mills, L.S., Ortega, Y., Ruggiero, L.F., Allendorf, F.W., 2003. Landscape location affects genetic variation of Canada lynx. Mol. Ecol. 12, 1807–1816.

Schwartz, M.K., Luikart, G., Waples, R.S., 2006. Genetic monitoring as a promising tool for conservation and management. Trends Ecol. Evol. 22, 25–33.

Selkoe, K.A., Henzler, C.M., Gaines, S.D., 2008. Seascape genetics and the spatial ecology of marine populations. Fish Fish. 9, 363–377.

Sinclair, M., 1988. Marine Populations. An Essay on Population Regulation and Speciation. Books in Recruitment Fishery Oceanography. University of Washington Press, Seattle, WA.

Sirois, P., Dodson, J.J., 2000. Critical periods and growth-dependent survival of larvae of an estuarine fish, the rainbow smelt Osmerus mordax. Mar. Ecol. Prog. Ser. 203, 233–245.

Slatkin, M., 1993. Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47, 264–279.

Smouse, P.E., Peakall, R., 1999. Spatial autocorrelation analysis of multi-allele and multi-locus genetic microstructure. Heredity 82, 561–573.

Spidle, A.P., Kalinowski, S.T., Lubinski, B.A., Perkins, D.L., Beland, K.F., Kocik, J.F., King, T.L., 2003. Population structure of Atlantic salmon in Maine with reference to populations from Atlantic Canada. Trans. Am. Fish. Soc. 132, 196–209.

Taylor, E.B., 1991. A review of local adaptation in salmonidae, with particular reference to Pacific and Atlantic salmon. Aquaculture 98, 185–207.

Waples, R.S., 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J. Hered. 89, 438–450.

Waples, R.S., Punt, A.E., Cope, J.M., 2008. Integrating genetic data into management of marine resources: how can we do it better? Fish Fish. 9, 423–449.

Ward, R.D., Woodward, M., Sibinski, D.O.F., 1994. A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. J. Fish Biol. 44, 213–232.

Waters, J.M., Epifanio, J.M., Gunter, T., Brown, B.L., 2000. Homing behavior facilitates subtle genetic differentiation among river populations of Alosa sapidissima: microsatellites and mtDNA. J. Fish Biol. 56, 622–636.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370.

Williamson-Nateson, E.G., 2005. Comparison of methods for detecting bottlenecks from microsatellite loci. Conserv. Genet. 6, 551–562.

Wright, S., 1943. Isolation by distance. Genetics 28, 1114–1138.

Xue, H., Incze, L., Xu, D., Wolff, N., Pettigrew, N., 2008. Connectivity of lobster populations in the coastal Gulf of Maine. Part I. Circulation and larval transport potential. Ecol. Model. 210, 193–211.