Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular Phylogenetics and Evolution 40 (2006) 148–154
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Mitochondrial DNA evidence of an early Holocene population expansion of threespine sticklebacks from Scotland

Ripan S. Malhi [a,b], Gillian Rhett [a], Alison M. Bell [c,*]

[a] *Trace Genetics, Inc (a DNAPrint Genomics Company), 4655 Meade St., Suite 300, Richmond, CA 94804, USA*
[b] *Department of Anthropology, University of California, Davis, One Shields Ave. 95616, USA*
[c] *Environmental and Evolutionary Biology, University of Glasgow, Glasgow, Scotland G12 8QQ, UK*

Received 15 August 2005; revised 19 February 2006; accepted 21 February 2006
Available online 17 April 2006

Abstract

In this study, we analyzed the cytochrome *b* gene in threespine stickleback (*Gasterosteus aculeatus*) populations from Scotland. We found evidence of a postglacial population expansion in Scotland and large differences in genetic diversity estimates among populations. Higher levels of genetic diversity are negatively correlated with distance from the ocean. In addition, distance from the ocean and predation risk both explain variation in plate count in Scottish populations. Overall, the mtDNA data support the racemic model of evolution in threespine stickleback.

Keywords: Genetic structure; Haplogroup; Migration; *Gasterosteus aculeatus*; Cytochrome *b*; Gene flow; Population expansion; Stickleback

1. Introduction

Following the retreat of the glaciers during the early Holocene (~12,000 yBP), some marine fishes independently colonized freshwater environments throughout the Northern Hemisphere. Threespine stickleback (*Gasterosteus aculeatus*) populations rapidly adapted to diverse freshwater environments resulting in unprecedented phenotypic diversity. Presently, some freshwater populations continue to experience gene flow from marine populations through interbreeding with anadromous sticklebacks (Taylor and McPhail, 1999).

Their evolutionary history and propensity for rapid, adaptive evolution (Bell and Foster, 1994; Bell et al., 2004) make the threespine stickleback an excellent system for studying parallel evolution. Such studies have become increasingly sophisticated with molecular tools that are allowing us to determine whether similar phenotypes evolved from the same changes in genetic mechanisms (Colosimo et al., 2005; Nachman, 2005). However, the study of parallel evolution requires independently evolving populations. To determine the degree of independence among different populations, we need to understand the history of stickleback populations. Specifically, the question of whether similar allelic patterns in noncontiguous regions are the result of common ancestry, migration or selection cannot be convincingly resolved without this information.

Several studies have examined the genetic diversity of threespine sticklebacks (Deagle et al., 1996; Haglund et al., 1992; Higuchi and Goto, 1996; Higuchi et al., 1996; Johnson and Taylor, 2004; Ortí et al., 1994; Raeymaekers et al., 2005; Reusch et al., 2001; Taylor and McPhail, 1999; Watanabe et al., 2003). Ortí et al. (1994) provided the first comprehensive study by investigating diversity within the mitochondrial cytochrome *b* gene in a global survey. Ortí et al. found two ancient clades, one in Japan and along the west coast of North America, the Trans North Pacific Clade, TNPC, originally named by (Thompson et al., 1997),

* Corresponding author. Present address: Environmental Science and Policy, University of California, Davis. One Shields Ave. 95616, USA. Fax: +011 44 141 330 5971.
E-mail addresses: ambell@ucdavis.edu, alison.bell@bio.gla.ac.uk (A.M. Bell).

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.02.019

and the other in populations from the west and east coast of North America and on the other side of the Atlantic basin in Europe (the Euro-American Clade (EAC)). Combined with fossil evidence, they concluded that the Atlantic clade was derived from the Pacific clade, and that this, in part, contributed to the reduced genetic variation among Atlantic populations. However, there is scope for additional mitochondrial DNA (mtDNA) information about Atlantic sticklebacks because unlike their freshwater Pacific counterparts (Taylor and McPhail, 1999), Atlantic populations were under-represented in population genetic surveys.

On a regional level, Raeymaekers et al. (2005) recently analyzed multilocus genotypes, STRs and allozymes to investigate the genetic structure of stickleback populations in Northern Europe. They identified data consistent with the 'racemic model' of evolution which posits that many freshwater populations independently emerged from marine/anadromous ancestors during the early Holocene, a period of postglacial retreat that resulted in a population expansion and colonization of freshwater habitats (Bell, 1976, 1986). However, the results from Raeymaekers et al. (2005) are also consistent with a model of differential gene flow (low gene flow among freshwater populations and high gene flow among anadromous populations) and a lower effective population size among freshwater populations.

The analysis of additional genetic systems can help determine if the genetic structure of stickleback populations in the North Atlantic were shaped by a population expansion. The use of mtDNA is a powerful tool for detecting past demographic events because the effective population is one fourth of that in the nuclear genome and therefore mtDNA is more sensitive to past demographic events. In addition, because mitochondrial genome is nonrecombining, we can trace the ancestry and relationships among haplotypes. A particularly useful tool for visualizing relationships among haplotypes are networks, which provide a meaningful representation of the genealogical relationships among individuals in recently diverged groups that is unattainable in the analysis of phylogenetic trees. Therefore, to investigate the genetic structure and population history of stickleback populations in Scotland, we sequenced the cytochrome *b* gene of 224 threespine sticklebacks from 1 marine and 11 freshwater populations. For comparative purposes, we also sequenced 11 individual ninespine sticklebacks (*Pungitius pungitius*).

Following the Loch Lomond Stadial at 11,000–10,000 rcybp, glaciers began to retreat in Scotland as temperatures rose dramatically. By around 7500 ybp the temperature in Scotland was about what it is today (Sissons, 1979). During the approx. 2500 years of glacial melt, a reasonable hypothesis is that more northern populations were covered in ice for longer (and are hence younger) than more southern populations.

Our results suggest two general conclusions. First, there was a postglacial population expansion in Scotland. Second, regional genetic structure in Scottish populations is primarily influenced by access to the ocean. These two lines of evidence add support to the 'racemic model' of stickleback evolution. Although here, we are primarily concerned with the pattern of genetic variation among populations, we also report data on lateral plate counts in the different populations. We report preliminary analyses on the relative influence of gene flow and local predation pressure, which has been identified as an important selective factor on variation in plate number (Bell and Richkind, 1981; Hagen and Gilbertson, 1972, 1973; Moodie et al., 1973; Reimchen, 1995).

2. Materials and methods

As part of a larger study, sticklebacks were collected from 11 different freshwater lochs in Scotland (Fig. 1). The intensity of predation by piscivorous fish (pike, perch, brown trout, and eels) was assessed using historical records and whole-lake seine surveys. Populations were classified as 'low risk' if they did not contain any piscivorous fishes, and were classified as 'high risk' if piscivorous fishes were detected (Table 2). Although sticklebacks are also subject to predation by birds and invertebrate larvae, those predators can move between water-bodies, making it difficult to determine predation risk by those predators. We classified populations greater than 14 km from the ocean as 'inland' and populations within 14 km of the ocean as 'coastal'. For comparison, one marine population was sampled near the island of Great Cumbrae, and one ninespine stickleback population was sampled on the Isle of Bute. Sticklebacks were collected between 11 August and 25 September 2004. Fish were transported to the laboratory in insulated buckets and maintained at Glasgow University Field Station. Fish were killed by a method approved by the home office. Samples for DNA were taken by fin clip from the caudal fin of dead fish. The number of lateral plates on alcohol-preserved specimens along the left side of the body was counted using a dissecting prod under a dissecting microscope.

DNA was extracted from 224 tissue samples from threespine sticklebacks and 11 samples from ninespine sticklebacks. Approximately 25 mg of each sample was used for the extraction with the Qiagen DNAEasy Tissue Kit®. A 1210 base pair (bp) region of the cytochrome *b* gene was amplified using primers Stkl14301 (CCTACCAGGACTTT AACCAGGACTA) and Stkl15511 (CCGGCGCTCTGG CGCTGAGCACTTT). The DNA was sequenced in both forward and reverse directions and a contig was created by combining forward and reverse sequences based on a large region of overlap. PCR amplification was performed in 20 μl volumes with 0.5 μM primers, 0.2 μM dNTPs, 2.5 mM $MgCl_2$, 1× Invitrogen PCR buffer (2 μl of 20× buffer, which consists of 200 mM Tris–HCl (pH 8.4) and 500 mM KCl), 0.3 U platinum Taq polymerase and 25–100 ng of template DNA. Amplification was carried out on Eppendorf Mastercycler thermal cyclers following an initial three minute step

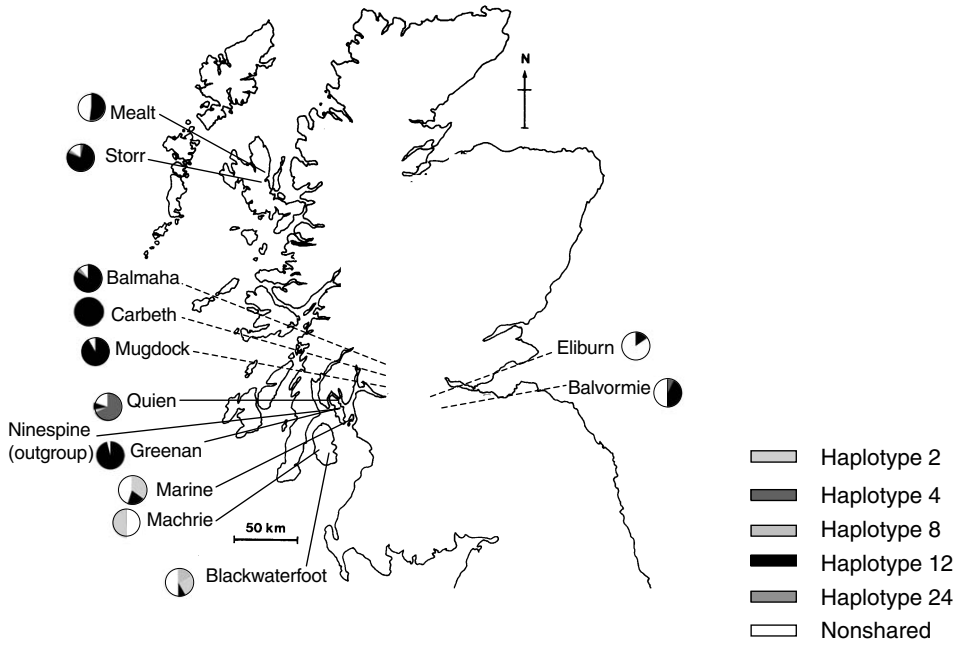

Fig. 1. Geographic location of populations and haplotypes.

at 94 °C, 35 cycles (10 s 94 °C, 15 s 56 °C, and 30 s 72 °C) were carried out, followed by a final one minute extension at 72 °C. Amplified DNA was purified by exonuclease I digestion, with 20μl PCR product, 0.1 μl 10× Invitrogen PCR buffer and 4 U exonuclease I in a 40 μl reaction. This reaction was incubated at 37 °C for 90 min followed by 80 °C for 20 min. The DNA was then cleaned up using the Eppendorf Perfect Prep PCR Cleanup 96 system with an Eppendorf vacuum manifold. The DNA was sequenced at the University of California, Davis Plant Genomics Facility.

DNA sequences were aligned and edited using BIOEDIT. All genetic diversity estimates, including calculations of mean pairwise difference and sequence diversity, were calculated using both ARLEQUIN (Schneider et al., 1997) and MEGA (Nei, 2000) software programs. We used two approaches for visualizing relationships among the populations. First, networks show relationships by depicting the number of mutational differences between haplotypes. This method accurately represents the reticulate structure of fine-scale genetic relationships (Posada and Crandall, 2001). Second, phylogenetic trees can show the broad ancestor-descendant relationships among haplotypes. Through the use of both approaches, we can learn about the fine-scale genetic structure of populations within the context of the broader evolutionary patterns of stickleback in Scotland. Networks were based on 221 samples and 1028 bp and were constructed using Median Joining (Bandelt et al., 1999) and Statistical Parsimony (Clement et al., 2000) algorithms. The phylogenetic tree was based on 28 haplotypes and constructed using the Neighbor-Joining algorithm and the Kimura 2 Parameter model. A bootstrap analysis was performed with a 1000 replicates. The phylogenetic tree was rooted with sequences from ninespine stickleback.

Tajima's test of neutrality (Tajima, 1989) was calculated using MEGA. During an expansion, the population deviates from mutation-drift equilibrium as new mutations are created faster than can be removed by genetic drift. Tajima's test compares the number of segregating sites per site with the nucleotide diversity. A significantly negative result is indicative of a population expansion and is often visualized by the presence of a large 'star-like' network in the network analysis. Differential gene flow among populations cannot result in a similar genetic pattern. Estimates of divergence time were based on a rough calibration of the time separation between *G. aculeatus* and *P. pungitius* at 10 million years ago (Bell, 1994). This crude calibration results in rough time estimates that should be viewed as a method to identify relative separation time between events and should not be interpreted in strict calendar time.

Net intermatch difference between *G. aculeatus* and *P. pungitius* was calculated in MEGA. We calculated a sequence diversity estimate of 15.4%. Based on the fossil evidence, this places the accumulation of sequence diversity at 1.54% per million years. This is similar to the estimates calculated by (Watanabe et al., 2003) at 1.6% per million years. Using this estimate, we calculated a substitution in the mtDNA cytochrome *b* gene occurring approximately every 60,000 yBP. Estimates of divergence time within networks were based on calculations of rho (r) (Forster et al., 1996; Saillard et al., 2000). F_{st} estimates were calculated using ARLEQUIN (Schneider et al., 1997). We used linear regression to test for the influence of distance from the ocean and elevation on the number of segregating sites, number of haplotypes, mean pairwise difference, and nucleotide diversity.

3. Results

We identified 28 haplotypes and 32 polymorphic sites in Scotland (Table 1). Five haplotypes were shared among some populations in Scotland and none of these shared haplotypes was exclusive to inland populations (Table 2). That is, the only haplotype shared among inland populations is haplotype 12, which is probably due to a common source of origin for these populations. Haplotype 4 was found in high frequency in Quien and was also present in Balvormie and Storr. It is likely that the presence of this haplotype in Balvormie and Storr is a result of a similar colonizing population rather than gene flow because of the large geographic distance between these two populations. Haplotypes 2 and 8 are shared among neighboring coastal (Machrie and Blackwaterfoot) populations and could be the result of gene flow or a similar colonizing population. Finally, haplotype 12 was found in all populations except Machrie.

The phylogenetic tree (Fig. 2) shows that haplotypes found near the ocean are ancestral to haplotypes found among inland populations. The bootstrap values for the derived haplotype clade are much lower than for the ancestral haplotype clade likely because of the low level of divergence among the derived haplotypes.

Within Scotland, populations exhibited a clear trend in genetic diversity. Populations further from the ocean had lower genetic diversity (segregating sites, nucleotide diversity, and mean pairwise difference) than populations closer to the ocean ($R^2 = 0.577$ or greater, $P =$ or less than 0.021; Table 2, see also Vrijenhoek et al., 1977). The plate data exhibit a parallel trend: coastal populations had more lateral plates than inland populations (Table 3, inland: 5.4 ± 0.26, $n = 90$; coastal: 8.0 ± 0.44, $n = 55$, $F_{1,134} = 37.1$, and $P < 0.0001$). We also detected an effect of predation risk on variation in plate number (low risk: 6.5 ± 0.27, $n = 61$; high risk: 7.13 ± 0.45, $n = 62$, $F_{1,113} = 10.27$, and $P = 0.002$). However, inland populations were predominately 'low risk'

Table 1
Nucleotide positions of polymorphic sites

	1440	1444	1446	1458	1461	1536	1560	1563	1575	1591	1599	1610	1653	1662	1722	1723	1743	1744	1749	1753	1758	1764	1770	1782	1791	1795	1803	1869	1895	1914	1940	1950	1985	1988	2012	2023	2241	2254	2265	2295	2300	2343	2370	2388	2394	2461	Scotland "Coastal"	Scotland "Inland"
	C	G	A	A	A	C	T	C	T	A	T	G	C	C	A	A	A	A	A	T	T	T	T	A	C	G	A	G	T	T	C	C	C	G	A	G	C	C	C	C	C	C	T	C	A	C	C	A	G		
Haplotype 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	4	
Haplotype 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	7	
Haplotype 3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	1	
Haplotype 4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	17	2
Haplotype 5		A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	6	
Haplotype 6		A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	
Haplotype 7		A	.	.	.	.	.	.	.	.	.	T	.	.	.	.	G	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	
Haplotype 8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	17	
Haplotype 9	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	1	
Haplotype 10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	T	.	.	.	1	
Haplotype 11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	G	.	.	A	T	.	.	.	.	1	
Haplotype 12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	68	59
Haplotype 13	T	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	3	
Haplotype 14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	C	.	.	.	G	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	1	
Haplotype 15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	1	
Haplotype 16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	1	
Haplotype 17	.	.	.	.	.	.	.	.	.	G	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	1	
Haplotype 18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	2	
Haplotype 19	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	1	
Haplotype 20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	T	.	.	.	.		2
Haplotype 21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	10	
Haplotype 22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.	.	.	.	.	T	G	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	1	
Haplotype 23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	T	.	.	.	.	1	
Haplotype 24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	.		1
Haplotype 25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	A	T	.	.	.	.		1
Haplotype 26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.		6
Haplotype 27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	C	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.		5
Haplotype 28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	C	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	1	

Table 2
Diversity estimates in Scotland

Population	Distance from ocean (km)	Location	Predation risk	Elevation	N	Haplotypes	Segregating sites	Mean pairwise difference
Scotland								
Balmaha	12	Inland	Low	112	19	3	2	0.30409
Balvormie	7.5	Inland	Low	706	14	2	3	0.42857
Blackwaterfoot	0.05	Coastal	High	34	12	6	10	3.84848
Carbeth	12	Inland	High	403	19	1	0	0
Eliburn	12.5	Inland	High	459	12	4	2	0.83333
Greenan	4.8	Coastal	Low	205	26	2	2	0.15385
Machrie	0.05	Coastal	Low	26	11	2	6	3.2727
Marine	0	n/a	n/a	17	19	6	12	3.4269
Mealt	0	Coastal	High	158	26	4	4	0.74769
Mugdock	12	Inland	Low	100	12	2	2	0.33333
Quien	0.75	Coastal	High	162	25	6	7	1.84667
Storr	0.75	Coastal	High	490	28	4	7	0.5468
Regional total								
Scotland					224	28	32	4.413

Bold indicates inland populations.

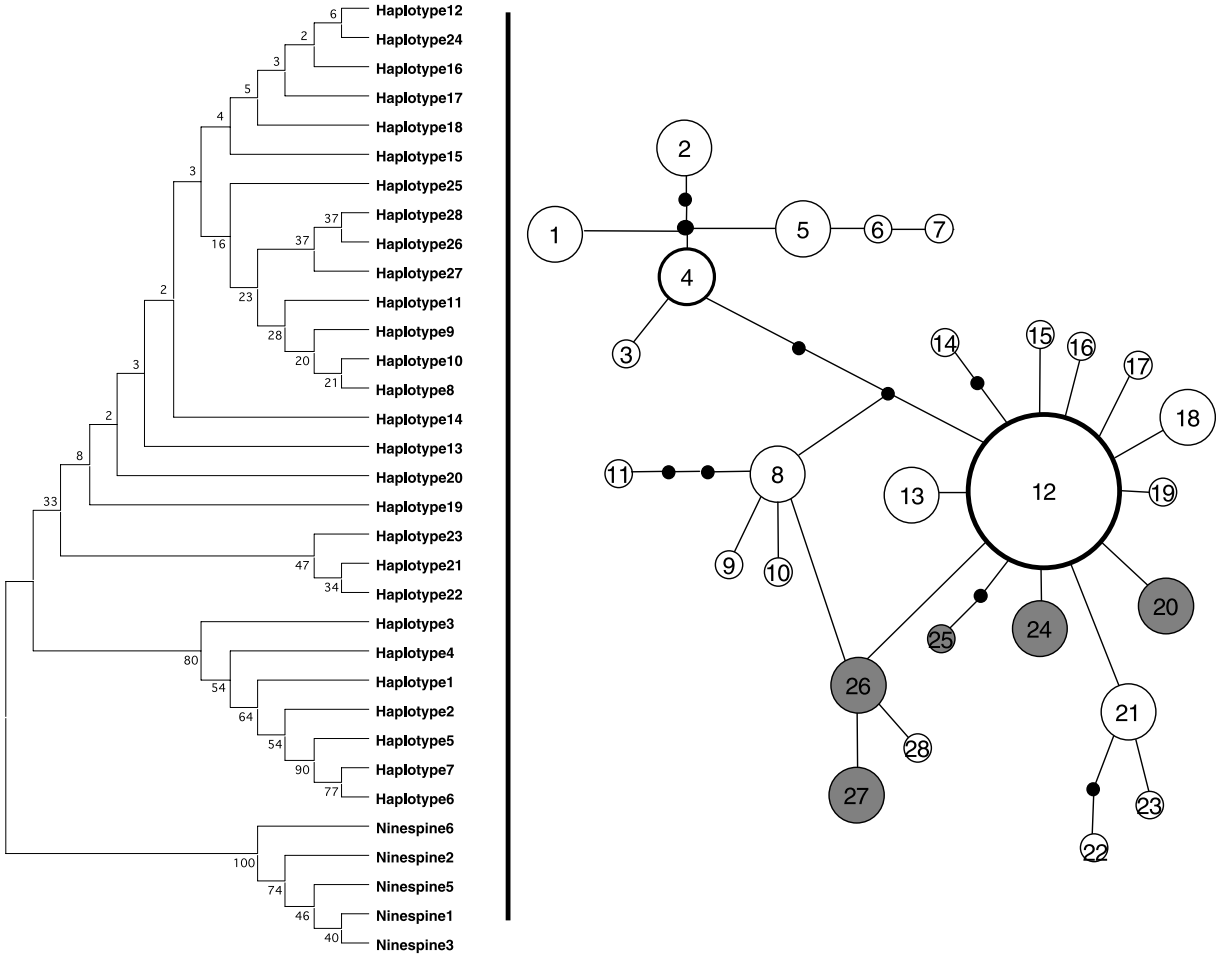

Fig. 2. Network and phylogenetic tree constructed to view genetic relationships among haplotypes. Shaded haplotypes are specific to "inland" populations, unshaded haplotypes are specific to "coastal" populations and unshaded haplotypes with thick borders are found in both "inland" and "coastal" populations.

Table 3
Lateral plate counts

Population	Location	Risk	*N*	Mean	Standard error of mean	Minimum	Maximum
Balmaha	Inland	Low	14	5.29	0.22	4	7
Balvormie	Inland	Low	19	4.74	0.46	2	10
Carbeth	Inland	High	22	4.45	0.16	4	6
Eliburn	Inland	High	13	4.77	0.93	3	15
Mugdock	Inland	Low	22	7.55	0.59	5	16
Blackwaterfoot	Coastal	High	10	8.9	1.14	6	18
Greenan	Coastal	Low	5	6.8	0.24	6	7
Machine	Coastal	Low	10	6.4	0.4	5	9
Mealt	Coastal	High	10	5.4	0.34	4	8
Quien	Coastal	High	10	9.1	1.35	6	18
Storr	Coastal	High	10	11	0.76	8	15

so the two factors are confounded. A variance components analysis suggests that both factors are important, but 'inland' explains more of the variation than predation risk does (inland: 43%, risk: 39%).

Pairwise F_{st} estimates for populations in Scotland were nearly all statistically significantly different (see Table 4). There was no relationship between F_{st} values and geographic distance among populations ($r = 0.08$, $P = 0.53$, and $n = 63$).

The statistical parsimony network provides multiple insights into Scottish stickleback population history (Fig. 2). The network is dominated by a 'star-like' pattern indicative of a range expansion. The results of the Tajima *D* test are consistent with evidence for a recent range expansion ($D = -1.747$). Over half of the tissue samples in this study belong to the haplotypes associated with the major range expansion and nearly all populations contained this haplotype. In particular, inland populations primarily

Table 4
Pairwise F_{st} values

	Balmaha	Balvormie	Greenan	Quien	Blackwaterfoot	**Carbeth**	**Eliburn**	Machrie	Marine	Mealt	**Mugdock**	Storr
Balmaha	0											
Balvormie	0.66633	0										
Greenan	0.03101	0.73806	0									
Quien	0.55826	0.58386	0.60375	0								
Blackwaterfoot	0.49838	0.49873	0.56424	0.16559	0							
Carbeth	0.03704	0.77623	0	0.58507	0.5286	0						
Eliburn	0.62816	0.7043	0.70048	0.60288	0.45133	0.7262	0					
Machrie	0.62216	0.60164	0.68439	0.29883	0.01605	0.65779	0.52344	0				
Marine	0.4307	0.46465	0.48266	0.18777	0.03513	0.45049	0.39396	0.11622	0			
Mealt	0.22651	0.43643	0.26292	0.55634	0.50627	0.2575	0.52803	0.60763	0.45106	0		
Mugdock	0.01963	0.63756	0.01945	0.51832	0.42975	0.04018	0.59163	0.55863	0.37751	0.19556	0	
Storr	0.00747	0.5918	0.01033	0.52622	0.48658	0	0.5703	0.60856	0.4197	0.20843	0	0

Inland populations are in bold.

exhibited haplotypes only one or two mutational steps away from haplotype 12 (Haplotypes 13–25, Fig. 1). Haplotypes derived from haplotype 12 tended to be population specific, which suggests that the inland populations were isolated after the expansion event. In contrast, coastal populations contained haplotypes associated with the expansion and ancestral haplotypes found in marine sticklebacks.

When we defined haplotype 12 as the ancestral type in this expansion and all others associated with the expansion as derived from haplotype 12, we obtained a rho-value (r) of 0.25 with a standard deviation of 0.0916. This places the major expansion in Scotland at $15{,}000 \pm 5498$ years before present (yBP).

4. Discussion

We detected the genetic signature of a major population expansion in Scotland approximately 15,000 ybp $\pm$ 5498 yBP, which is consistent with the time when the glaciers retreated in the Atlantic, allowing marine sticklebacks to gain access to freshwater environments. Previously, Raeymaekers et al. (2005) identified STR evidence consistent with a recent population expansion in Belgium and the Netherlands. The data from these two studies using two independent genetic systems from different geographic locations demonstrates that early Holocene population expansions were widespread in the North Atlantic.

Levels of genetic diversity in coastal populations in Scotland were comparable to levels of genetic diversity within Pacific populations Ortí et al. (1994). The analysis of additional genetic regions along with additional sampling in the Atlantic and Pacific are needed to understand the cause of the similarity in levels of diversity among these populations.

Inland populations harbor predominantly haplotypes associated with the major population expansion. Following the range expansion, these inland populations evolved in isolation, as evidenced by the population specific postglacial haplotypes. In contrast, in addition to 'expansion' haplotypes, coastal populations also contain haplotypes that are common in the marine population. The influence of both processes (the population expansion and access to the ocean) is illustrated by the neighboring Greenan and Quien populations, which both occur on the Isle of Bute. Greenan exhibited one of the lowest levels of genetic diversity in this dataset, while Quien exhibited one of the highest levels of genetic diversity. While Quien is connected to the ocean by a short stream, Greenan is in a valley in the center of the island. The topography of these two lochs suggests that Quien has experienced frequent gene flow with marine forms during its history, whereas the Greenan population has not. Therefore, we hypothesize that the general haplotype structure within the Greenan population probably reflects initial colonization events, subsequent genetic drift and natural selection that has occurred since initial colonization, while the haplotype structure of the Quien population is probably more influenced by gene flow with marine forms.

It has been proposed that sticklebacks conform to a 'racemic model' of evolution (Bell, 1987). This model makes several predictions, two of which are tested here. First, it predicts that the haplotypes in freshwater populations are a random sample of haplotypes in marine populations. Second, it predicts that derived freshwater populations should be less genetically diverse than marine populations. The pattern of mtDNA diversity among Scottish populations generally conforms to these predictions.

First, the most common haplotypes in the marine population that we sampled were haplotypes 5, 8, and 12. While haplotype 12 was the most common freshwater haplotype, it was among, but not the most, common of the marine haplotypes. It is possible that this is a result of genetic substructure among marine populations. Additional marine populations need to be sampled to determine the form and extent of this substructure within marine populations.

Second, we showed that inland populations were less genetically diverse than coastal populations. In addition, distance to the ocean was negatively related to genetic diversity. The presence of 'ancestral' marine haplotypes in coastal freshwater populations is most likely a result of gene flow from marine populations to freshwater ones. Johnson and Taylor (2004) identified a similar pattern in the distribution of TNPC haplotypes in populations from the Pacific Northwest. They showed that TNPC haplotypes were absent from

high altitude populations and concluded that 'accessibility' to the ocean was an important factor in determining the distribution of mtDNA haplotypes. Therefore, our results together with Johnson and Taylor (2004) suggest that access to the ocean is a key predictor of freshwater stickleback genetic diversity in both the Pacific and Atlantic.

The plate data are also consistent with the hypothesis that coastal populations are receiving migrants from the ocean. Marine sticklebacks generally have a full complement of armor, with a rapid reduction in the number of lateral plates moving inland (Bell and Foster, 1994). We found that inland populations had fewer lateral plates than coastal populations, but predation risk is also important (as in Hagen and Gilbertson, 1972). Future studies should be designed to explicitly test the relative importance of these two factors.

The regional distribution of mtDNA haplotypes and plate counts in Scotland and the Pacific Northwest (Johnson and Taylor, 2004) fits well with the model that low-plated alleles, present at detectable frequencies in completely plated stickleback populations, are maintained by gene flow with freshwater populations (Colosimo et al., 2005; Thompson et al., 1997). It appears that freshwater populations with access to the ocean act as a reservoir to keep alleles associated with freshwater populations in the marine gene pool. Standing variation for freshwater alleles in marine populations might help explain sticklebacks' ability to evolve rapidly in novel freshwater environments.

Acknowledgment

This research was supported by the NSF-IRFP to A.M.B.

References

Bandelt, H., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16, 37–48.

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

Bell, M.A., 1986. Stickleback fishes: bridging the gap between population biology and paleobiology. Trends in Ecology and Evolution 3, 320–324.

Bell, M.A., 1987. Interacting evolutionary constraints in pelvic reduction of threespined sticklebacks. Biological Journal of the Linnean Society 31, 347–382.

Bell, M.A., 1994. Paleobiology and evolution of threespine stickleback. In: Bell, M.A., Foster, S.A. (Eds.), The Evolutionary Biology of the Threespine Stickleback. Oxford University Press, Oxford, pp. 439–471.

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

Bell, M.A., Foster, S.A., 1994. The Evolutionary Biology of the Threespine Stickleback. Oxford University Press, Oxford.

Bell, M.A., Richkind, K.E., 1981. Clinal variation of lateral plates in threespined stickleback fish. American Naturalist 117, 113–132.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer programto estimate gene genealogies. Molecular Evolution 9, 1657–1659.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., et al., 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307, 1928–1933.

Deagle, B.E., Reimchen, T.E., Levin, D.B., 1996. Origins of endemic stickleback from the Queen Charlotte Islands: mitochondrial and morphological evidence. Canadian Journal of Zoology 74, 1045–1056.

Forster, P., Harding, R., Torroni, A., Bandelt, H.J., 1996. Origin and evolution of native American mtDNA variation: a reappraisal. American Journal of Human Genetics 59, 935–945.

Hagen, D.W., Gilbertson, L.G., 1972. Geographic variation and environmental selection in *Gasterosteus aculeatus* in the Pacific Northwest, America. Evolution 26, 32–51.

Hagen, D.W., Gilbertson, L.G., 1973. Selective predation and the intensity of selection acting upon the lateral plates of threespined sticklebacks. Heredity 30, 273–287.

Haglund, T.R., Buth, D.G., Lawson, R., 1992. Allozyme variation and phylogenetic-relationships of Asian, North-American, and European populations of the threespine stickleback, *Gasterosteus aculeatus*. Copeia, 432–443.

Higuchi, M., Goto, A., 1996. Genetic evidence supporting the existence of two distinct species in the genus *Gasterosteus* around Japan. Environmental Biology of Fishes 47, 1–16.

Higuchi, M., Goto, A., Yamazaki, F., 1996. Genetic structure of threespine stickleback, *Gasterosteus aculeatus*, in Lake Harutori, Japan, with reference to coexisting anadromous and freshwater forms. Ichthyological Research 43, 349–358.

Johnson, L.S., Taylor, E.B., 2004. The distribution of divergent mitochondrial DNA lineages of threespine stickleback (*Gasterosteus aculeatus*) in the northeastern Pacific Basin: post-glacial dispersal and lake accessibility. Journal of Biogeography 31, 1073–1083.

Moodie, G.E., McPhail, J.D., Hagen, D.W., 1973. Experimental demonstration of selective predation in *Gasterosteus aculeatus*. Behaviour 47, 95–105.

Nachman, M.W., 2005. The genetic basis of adaptation: lessons from concealing coloration in pocket mice. Genetica 123, 125–136.

Nei, M., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Ortí, G., Bell, M.A., Reimchen, T.E., Meyer, A., 1994. Global survey of mitochondrial-DNA sequences in the threespine stickleback—evidence for recent migrations. Evolution 48, 608–622.

Posada, D., Crandall, K.A., 2001. Intraspecific gene genealogies: trees grafting into networks. Trends in Ecology and Evolution 16, 37–45.

Raeymaekers, J.A.M., Maes, G.E., Audenaert, E., Volckaert, F.A.M., 2005. Detecting Holocene divergence in the anadromous-freshwater threespined stickleback (*Gasterosteus aculeatus*) system. Molecular Ecology 14, 1001–1014.

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

Reusch, T.B.H., Wegner, K.M., 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.

Saillard, J., Forster, P., Lynnerup, N., Bandelt, H.J., Norby, S., 2000. mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. American Journal of Human Genetics 67, 718–726.

Schneider, S., Roessli, D., Excoffier, D., 1997. Arlequin: A Software for Population Genetic Data Analysis. University of Geneva, Geneva.

Sissons, J.B., 1979. The Loch Lomond Stadial in the British Isles. Nature 280, 199–203.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Taylor, E.B., McPhail, J.D., 1999. Evolutionary history of an adaptive radiation in species pairs of threespine sticklebacks (*Gasterosteus*): insights from mitochondrial DNA. Biological Journal of the Linnean Society 66, 271–291.

Thompson, C.E., Taylor, E.B., McPhail, J., 1997. Parallel evolution of lake-stream pairs of threespine sticklebacsk (*Gasterosteus*) inferred from mitochondrial DNA variation. Evolution 51, 1955–1965.

Vrijenhoek, R.C., Angus, R.A., Schultz, R.J., 1977. Variation and heterozygosity in sexually versus clonally reproducing populations of *Poeciliopsis*. Evolution 31, 767–781.

Watanabe, K., Mori, S., Nishida, M., 2003. Genetic relationships and origin of two geographic groups of the freshwater threespine stickleback, 'hariyo'. Zoological Science 20, 265–274.