

Ancient divisions, recent expansions: phylogeography and population genetics of the round goby *Apollonia melanostoma*

JOSHUA E. BROWN and CAROL A. STEPIEN

Great Lakes Genetics Laboratory, Lake Erie Center and Department of Environmental Sciences, University of Toledo, 6200 Bayshore Road, Toledo, OH 43618, USA

Abstract

During the past two decades, the round goby *Apollonia melanostoma* (= *Neogobius melanostomus*) has expanded its range via shipping transport and canals, extending north and west from the Ponto-Caspian region of Eurasia and to the North American Great Lakes. Exotic populations of the round goby have been very successful in the Baltic Sea and the Great Lakes regions, exerting significant ecological changes. Our study evaluates the population genetic and biogeographical structure of the round goby across its native and non-indigenous ranges, in light of geological history and its expansion pathways. We analyzed seven new nuclear microsatellite loci and mitochondrial DNA cytochrome *b* gene sequences from 432 individuals in 22 locations. Population structure was tested using F_{ST} -analogs, phylogenetic trees, clustering diagrams, Bayesian assignment tests and nested clade analyses. Results show that native populations in the Black vs. the Caspian Sea basins diverge by 1.4% and *c.* 350 000 years, corresponding to closure of their prior connections and supporting the taxonomic separation of the Black Sea *A. m. melanostoma* from the Caspian Sea *A. m. affinis*. Their within-basin populations diverge by ~0.4% and 100 000 years. Non-indigenous populations in the Baltic Sea and Danube and Dnieper Rivers trace to separate northern Black Sea origins, whereas the upper Volga River system houses mixed populations of *A. m. melanostoma* and *A. m. affinis*. Native populations average twice the genetic diversity of most exotic sites; however, sites in the Volga River system have high diversity due to mixing of the two taxa. Our results highlight how vicariance and anthropogenic disturbances have shaped a rapidly expanding species' genetic heritage.

Keywords: *Apollonia melanostoma*, Gobiidae, introduced species, Ponto-Caspian region, population genetics, round goby

Received 13 December 2007; revision accepted 19 March 2008

Introduction

Aquatic species introductions have increased in the past two decades due to canal construction, shipping and ballast-water transport. Newly introduced species face many challenges that may thwart their establishment – including novel habitat, food, competitors, predators and pathogens – as well as potential reproductive and recruitment difficulty. Knowledge of a nonindigenous species' phylogeographical history may aid prediction of

its relative success in novel habitats. For example, it has been hypothesized that the post-Pleistocene expansion of taxa into new aquatic habitats was ecologically similar to their present-day movements via anthropogenic transport pathways and global climate change (Hewitt 2004). Until recently, the Ponto-Caspian region of Eurasia (Black, Azov and Caspian Seas) housed an endemic fauna, such as the neogobiin flock of ~20 fish species (Stepien & Tumeo 2006), *Dreissena* mussels (Stepien *et al.* 2002, 2003; Therriault *et al.* 2005) and a variety of crustaceans (Cristescu *et al.* 2003; Audzijonyte *et al.* 2006). Due to human activities, many of these Ponto-Caspian endemics, including neogobiin fishes, have spread beyond the region (Bij de Vaate *et al.* 2002; Stepien *et al.* 2005; Ricciardi 2006).

Correspondence: Carol Stepien, Fax: +01 419-530-8399; E-mail: carol.stepien@utoledo.edu

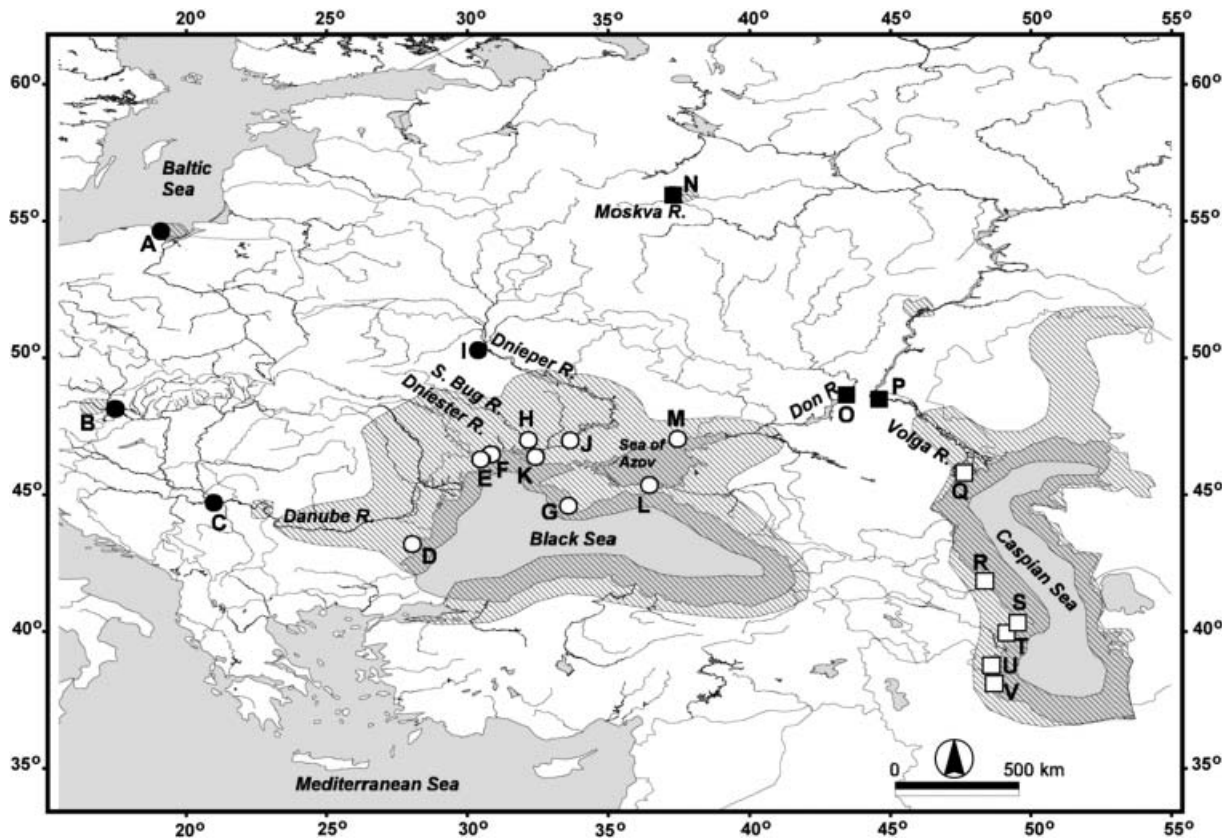


Fig. 1 Eurasian round goby distribution (hatched areas), showing sampling locations (lettered). Circles = the Black-Azov Sea basin clade, squares = the Caspian Sea basin clade, with filled symbols indicating non-native locations.

The round goby *Apollonia melanostoma* (formerly *Neogobius melanostomus*; Stepien & Tumeo 2006) is the most widespread and successful nonindigenous neogobiin fish (Neogobiinae: Gobiidae: Teleostei). It was originally distributed in the Black, Azov and Caspian Seas and the lower reaches of their tributaries. The round goby has recently spread up the great rivers of the Ponto-Caspian region and Eastern Europe (see Fig. 1), as well as to the Baltic Sea and the North American Great Lakes (both *c.* 1990; Jude *et al.* 1992; Skora & Stolarski 1993). Ballast water is a primary means of their anthropogenic transport, through the uptake of pelagic juveniles that rise nocturnally to feed in the water column (Hensler & Jude 2007). This appears to be the mechanism that led to the Baltic Sea and the North American colonizations, and likely that of the upper Volga River. Expansion populations likely migrated more slowly along canals.

The round goby inhabits a gradient of marine to freshwater habitats, reaching its largest sizes in the seas (~25 cm total length; Svetovidov 1964; C.A. Stepien, personal observation). Most individuals appear to move little in terms of geography throughout their lives, except for some seasonal migration offshore during the winter (Kostyuchenko 1969; Ray &

Corkum 2001). Males guard nests under rocks or other objects and are highly territorial (MacInnis & Corkum 2000). They eat molluscs (including dreissenids in freshwaters and mytilids in marine waters), other invertebrates and fishes (Pinchuk *et al.* 2003).

Our study analyzes population genetic divergence of the round goby across its native and nonindigenous Eurasian ranges, employing a dual approach of mitochondrial DNA (mtDNA) sequencing and seven new nuclear microsatellite (msat) loci. This investigation includes re-analysis of samples used by prior round goby genetic studies of mtDNA genes by Dougherty *et al.* (1996), Dillon & Stepien (2001), Stepien *et al.* (2005) and Stepien & Tumeo (2006), which totalled 59 samples from six Eurasian sites. We provide here the first analysis of msat loci and extend the previous work to include the entire cytochrome (*cyt*) *b* gene and add 16 additional Eurasian sites, for a total of 22 sites and 432 individuals. Our approach allows both the phylogeographical history of lineages to be assessed (through maternally inherited mtDNA), as well as fine- and broad-scale population genetic relationships to be discerned (by comparing biparentally inherited msat with mtDNA patterns). We test for historical vicariant separation

patterns of the round goby among native population sites, as well as donor–recipient population relationships for exotic locations in the Black/Azov Sea and Caspian Sea basins. Pathways for round goby expansion and possible jump transport are examined among the great rivers of the Ponto-Caspian basin, including the Danube, Dniester, Dnieper, Southern Bug, Don and Volga River watersheds, as well as the Baltic Sea.

Biogeographical history of the Ponto-Caspian region

The Ponto-Caspian region has been dominated by the intermittent union and division of its major basins in response to tectonic events and climatic shifts (Dumont 1998), which shaped its endemic fauna. It is likely that this pattern of change selected for species with flexible niche requirements and good colonization ability, as demonstrated by their historical and recent colonization successes, including postglacial and recent Eurasian population expansions and establishment in North America (c. 1990; see Ricciardi & MacIsaac 2000; Bij de Vaate *et al.* 2002). In addition to its brackish seas, the Ponto-Caspian region drains several major river systems, with the Danube, Dniester and Dnieper Rivers emptying into the Black Sea, the Don River draining into the Sea of Azov and the Volga River flowing to the Caspian Sea.

The Black and Caspian Sea basins separated c. 5 Ma, with periodic reconnections to each other through the Sea of Azov and the Manych Depression and to the Mediterranean Sea, due to tectonic activity and climate change (Briggs 1974; Reid & Orlova 2002). Pleistocene glacial cycles significantly altered the climate and hydrogeography of the Ponto-Caspian region (Hewitt 2004; Mangerud *et al.* 2004; Bahr *et al.* 2005). Water levels in the Black, Azov and Caspian Seas decreased during dry glacial periods, restricting aquatic fauna to portions of their former distributions (Zaitsev & Mamaev 1997; Dumont 1998; Major *et al.* 2006) that served as refugia for once widely distributed taxa, including fishes (Griffiths 2006; Reyjol *et al.* 2007) such as the spined loach *Cobitis taenia* (Culling *et al.* 2006), bullhead *Cottus gobio* (Englbrecht *et al.* 2000, Hanfling *et al.* 2002), brown trout *Salmo trutta* (Bernatchez 2001), chub *Leuciscus cephalus* (Durand *et al.* 1999), dace *Leuciscus leuciscus* (Costedoat *et al.* 2006) and Eurasian perch *Perca fluviatilis* (Nesbø *et al.* 1999).

Interglacial increases in water levels opened recolonization pathways into western and northern Eurasia via tributaries (Hewitt 2004) and restored connections between the Black/Azov and Caspian Seas at several times during the Pleistocene Epoch (Dumont 1998). Major faunal exchanges accompanied these reconnections three times during the Pleistocene (2.0–0.9 Ma, 0.7–0.4 Ma and 0.35–0.25 Ma; Reid & Orlova 2002). Changing connections also resulted in salinity fluctuations from fresh to brackish to marine

conditions, depending on connections to the Mediterranean Sea and freshwater inputs from rivers (Waters *et al.* 1994). This cycle of vicariance and reconnection resulted in many genetically divergent Black and Caspian Sea lineages, including copepod and amphipod taxa (Cristescu *et al.* 2003), mysid shrimp (Audzijonyte *et al.* 2006), dreissenid mussels (Stepien *et al.* 2003; Gelembiuk *et al.* 2006) and sturgeon (Choudhury & Dick 1998). A similar pattern may exist in round gobies, as two currently unrecognized subspecies of the round goby were once proposed: *A. m. melanostoma* in the Black Sea and *A. m. affinis* in the Caspian Sea (Navozov 1912; Pinchuk *et al.* 2003).

As a result of their distinct though related geological histories, genetic signatures of Black- and Caspian-Sea fauna may reveal divergent or congruent evolutionary patterns. Moreover, the natural separation of the Black- and Caspian-Sea drainages has been breached anthropogenically by several canals, of which the largest is the 1952 Volga–Don Canal (Zaitsev & Mamaev 1997). These canals create artificial pathways that allow taxa to move between systems, leading to gene flow.

Questions

Our objective here is to quantify the genetic diversity and divergence patterns across the native and introduced Eurasian distribution of the round goby (Fig. 1) to test the following:

- 1 Are populations in the Black and Caspian Sea basins genetically distinct?
- 2 Do populations in the Black and Caspian Seas differ in overall genetic diversity levels, and do areas of their expansion show founder effects?
- 3 Do overall biogeographical divergence patterns correspond to historical glacial refugia and show congruence with other taxa?
- 4 Do the gene pools of populations in freshwater and marine systems differ?
- 5 Did populations in the Volga–Don Canal and the northern Volga River system originate from the Sea of Azov or the Caspian Sea?
- 6 What was the likely founding population source for the nonindigenous population in the Gulf of Gdansk, Baltic Sea?

Materials and methods

Sampling, DNA extraction and genetic data collection

We sampled 22 locations from the round goby's native and introduced Eurasian ranges (Table 1, Fig. 1), including 5–30 individuals per site and a total of 432 individuals. To maximize phylogeographical information, we sampled as

Table 1 Round goby sampling locations, latitude and longitude, typical salinity (ppt), sample size (*N*), *cyt b* haplotypes, and genetic diversity values determined using mitochondrial *cyt b* sequence data and seven nuclear microsatellite loci. **Bold** haplotypes are shared among locations; *, introduced locations; H_O , observed microsatellite heterozygosity; H_E , expected heterozygosity; F_{IS} , inbreeding coefficient. Salinity values for locations A–M are from Yuriy Kvach (personal communication) and N–V are from Audzijonyte *et al.* (2006)

Water body	Location & map label	Latitude	Longitude	Salinity (ppt)	<i>N</i>	Haplotype(s)	Gene diversity	Msat		
							<i>Cyt b</i>	H_O	H_E	F_{IS}
Baltic Sea, Gulf of Gdansk	A. Gdynia, Poland*	54.54304	18.56209	7	20	1	0	0.414	0.782	0.047
Danube River	B. Bratislava, Slovakia*	48.13611	17.13509	0	39	1,23	0.051	0.377	0.609	0.038
	C. Prahovo, Serbia*	44.69633	20.72083	0	45	1,7	0.044	0.459	0.585	0.021
	D. Varna, Bulgaria	43.20138	27.88590	17–18	37	1,10,11,62,65,67,68,77,78,79	0.809	0.462	0.871	0.047
Black Sea	E. Bilgorod, Ukraine	46.29974	30.37233	0–0.5	20	1,2,3,4,5,6,9,10	0.758	0.519	0.877	0.041
	F. Odessa Ukraine	46.47082	30.73509	14–17	20	1,10,18,48,60,61,62,63,64,76	0.838	0.427	0.799	0.047
	G. Sevastopol, Ukraine	44.60404	33.54084	17–18	25	1,18,48,65,66,69,74,75,80	0.868	0.538	0.790	0.032
Southern Bug River	H. Mikolaev, Ukraine	47.00000	32.08000	0–1	27	1,18,49,50,51,58,72,73	0.459	0.520	0.889	0.041
Dnieper River	I. Kiev, Ukraine*	50.30264	30.59378	0	17	1,58,59	0.539	0.377	0.527	0.028
	J. Kakhovka, Ukraine	46.98549	33.60829	0	15	1,8	0.133	0.374	0.504	0.026
	K. Kherson, Ukraine	46.38000	32.34000	0	25	1,7,8,57	0.417	0.463	0.757	0.039
Kerch Strait	L. Kerch, Ukraine	45.35833	36.475834	12.0–14.0	10	18,41,42,43,44,45,46,47,48	0.978	0.508	0.762	0.033
Sea of Azov	M. Mariupol, Ukraine	47.03333	37.50000	6.0–8.0	13	11,12,13,14,15,16,17,18,19	0.978	0.524	0.811	0.035
Moskva River	N. Moscow, Russia*	55.93963	37.32736	0	9	24,25,26,27,28	0.857	0.426	0.539	0.021
Volga-Don Canal	O. Iliovka, Russia*	48.64327	43.61707	0	10	20	0	0.475	0.739	0.036
Volga River	P. Svetli Yar, Russia*	48.48464	44.78468	0	10	20,24,25,29,30	0.667	0.582	0.573	–0.002
	Q. Damchik, Russia	45.78835	47.88695	0–6	22	20,21,22	0.177	0.452	0.696	0.035
	R. Nabran, Azerbaijan	41.83722	48.62000	12–14	14	31,33,70,71	0.396	0.494	0.725	0.032
Caspian Sea	S. Shikh, Azerbaijan	40.30417	49.80500	12–14	14	31,32,35,81,82,83,84	0.692	0.455	0.618	0.034
	T. Alet, Azerbaijan	39.94000	49.40917	12–14	5	31	0	0.455	0.610	0.025
	U. Lenkoran, Azerbaijan	38.75194	48.86889	12–14	5	31,85	0.400	0.400	0.642	0.038
	V. Lisar, Iran	38.08083	48.97500	12–14	30	31,32,33,34,35,36,37,38,39,40,86	0.774	0.464	0.763	0.039

Table 2 Cytochrome *b* (A) and microsatellite (B) primers used for amplification of round goby DNA, with sequence, annealing temperature (T_A), and publication source. Table 2B contains the range of repeat numbers (R_N) and sizes (R_S ; bp), number of alleles (N_A), average observed heterozygosity (H_O), and average F_{IS} and F_{ST} for each microsatellite locus (A)

Primer	Sequence (5'–3')	T_A (°C)	Source					
L14724	GTGACTTGAAAAACCACCGTT	52	Meyer <i>et al.</i> 1990					
H5	GAATTYTRCGTTTGGGAG	52	Akihito <i>et al.</i> 2000					
L15066	TTGGTCGAGGCCTCTATTACG	52	This study					

Primer	Sequence (5'–3'; Feldheim <i>et al.</i> in prep.)	T_A (°C)	R_N	R_S (bp)	N_A	H_O	F_{IS}	F_{ST}
<i>Ame1F</i>	AGAACAGTCTGGAGGACTCTTTG	58	1–36	147–252	23	0.211	0.025	0.272
<i>Ame1R</i>	GCGCTTTGTGACCATGTCT	58						
<i>Ame10F</i>	ATGCGAAGCCGATTTCTG	52	2–19	196–247	15	0.288	0.065	0.292
<i>Ame10R</i>	CCATATGTCAGGCGATATTCC	52						
<i>Ame17F</i>	GGCGCAACCTCATTTAATC	58	4–55	138–291	38	0.672	0.035	0.136
<i>Ame17R</i>	GTTTAGGCGGGGTTAAGAG	58						
<i>Ame23F</i>	AAAGCATCAGCAGCAGTTGT	58	2–45	144–273	37	0.777	0.056	0.050
<i>Ame23R</i>	TATGTGAGTGTGCGGATGGT	58						
<i>Ame129F</i>	TGCTCGGTCTACTTCAAGC	56	1–56	119–339	51	0.391	0.029	0.117
<i>Ame129R</i>	GCATTCACATTCCTCCCACT	56						
<i>Ame133F</i>	GCCACCCCTTCACTCTT	56	4–27	177–269	26	0.539	0.028	0.122
<i>Ame133R</i>	GGCTATGGCATTTTCTCTCC	56						
<i>Ame194F</i>	AAACACACAGTCACAAGCACA	52	1–44	168–340	35	0.235	0.074	0.311
<i>Ame194R</i>	CACAGCTAATGGGGATCCTA	52						

many locations as possible (see Pons & Petit 1995; Culling *et al.* 2006), focusing on ports, areas of range expansion and major watersheds. Unfortunately, we were unable to sample every introduced location, especially along the Volga River in Russia. Sampling locations were defined as non-native or native based on the reports in the literature and the range presented by Pinchuk *et al.* (2003) and noted in Fig. 1 and Table 1. As both the Black and Caspian Seas are brackish, but their tributaries are freshwater, we classified locations with salinity greater than 0 ppt (Table 1) as 'marine' for purposes of comparison. These locations were along the sea coast, as opposed to being upstream in a river. Round goby samples were placed directly in 95% ethanol in labelled vials and stored at room temperature. The same individuals were analyzed for mitochondrial (mt) and msat loci, facilitating comparison. We compared variation in the round goby with its sister species, the monkey goby *Apollonia fluviatilis*; and with other neogobiin relatives, including the freshwater tubenose goby *Proterorhinus semilunaris*, the bighead goby *Neogobius kessleri* and the toad goby *Mesogobius batrachocephalus* (see Stepien & Tumeo 2006).

Genomic DNA was extracted from 25 mg of fin tissue using Qiagen DNeasy kits (Qiagen, Inc.), eluted in 200 μ L of water, stored at 4 °C until used for polymerase chain reaction (PCR) amplification, and archived at –80 °C. The entire mt cytochrome *b* (cyt *b*) gene (1138 bp), and part of

the following tRNA-Thr (66 bp), were amplified using the primers L14724, L15066 and H5 (Table 2A) in 25 μ L reactions containing 1 unit *Taq* polymerase, 200 μ M dNTPs, 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris-HCl, 0.5 μ M of primers L14724 and H5 (Table 2A), and ~30 ng of template. Amplifications were performed on a MJR DYAD thermal-cycler (Bio-Rad Laboratories), with initial denaturation at 94 °C for 120 s, followed by 35 cycles of denaturation (94 °C for 45 s), annealing (52 °C for 30 s) and extension (72 °C for 60 s), plus a final 72 °C extension for 180 s. Sequencing using the same primers was outsourced to the Cornell University Life Sciences Core Laboratories Center (<http://cores.lifesciences.cornell.edu/brcinfo>). Sequences then were aligned with CLUSTAL X version 1.8 (Thompson *et al.* 1994) and adjusted using BIOEDIT version 7.0 (Hall 1999, 2004) in our laboratory.

We analyzed variation at seven nuclear msat loci developed for the round goby by Kevin Feldheim (Field Museum of Natural History) with us (Feldheim *et al.* in preparation, Table 2B). Amplification used 10 μ L reactions of 0.6 units *Taq*, 50 μ M nucleotides, 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris-HCl and 0.5 μ M of each primer (Table 2B), and ~30 ng of template, with a sterile mineral oil overlay to maintain the reaction volume. A thermal cycle of 2 min at 94 °C for initial denaturation was followed by 35 cycles of denaturation (94 °C, 30 s), annealing (1.00 min) at

a primer-specific temperature (Table 2B) and extension (72 °C, 30 s), followed by a 5-min final extension at 72 °C. Amplification products were diluted 1 : 50 in water, of which 1 µL was added to 13 µL of formamide and ABI (Applied Biosystems Inc.) Gene Scan 500 size standard and analyzed on a 96-well plate with an ABI 3130 Genetic Analyser and GENEMAPPER version 3.7 in our laboratory. Output profiles were checked to confirm allelic size variants and representative alleles were sequenced to verify that length polymorphisms were due to variation in copy number of single repeat motifs.

ARLEQUIN (version 3.01; Excoffier *et al.* 2005) was used to assign individuals to the correct mt haplotype, which were deposited in NIH GenBank as Accession nos EU331156–EU331236 (<http://www.ncbi.nlm.nih.gov>). We identified codon positions and transitional and transversional substitutions.

Data analysis

Allelic frequencies, number of private alleles, conformance to Hardy–Weinberg equilibrium (HWE) expectations (for msat data) and linkage disequilibrium (for msats) were evaluated in GENEPOP version 3.4 (Raymond & Rousset 1995, 2004). Levels of significance for HWE and linkage disequilibrium tests were adjusted using nonsequential Bonferroni correction (Sokal & Rohlf 1995). HWE deviations were tested for heterozygosity deficiency or excess and for the presence of null alleles with MICROCHECKER version 2.23 (<http://www.microchecker.hull.ac.uk>; van Oosterhout *et al.* 2004, 2006).

Genetic composition among samples were analyzed to identify true populations (i.e. those distinguished by significantly divergent gene pools) using a pairwise *F*-statistic analogue (θ_{ST} ; Weir & Cockerham 1984) and contingency tests (Raymond & Rousset 1995; Goudet *et al.* 1996) for both mt and msat data. Relationships between recently diverged samples, such as those tested here, have been shown to be better resolved in models using contingency tests (see Balloux & Lugon-Moulin 2002), which are independent from HWE assumptions, nonparametric and little affected by small sample size (Raymond & Rousset 1995; Goudet *et al.* 1996). Our additional use of an *F*-statistic analogue facilitated direct comparisons with other studies (see Stepien *et al.* 2007). For the *cyt b* data, we also calculated mean uncorrected pairwise *p*-distances in MEGA version 3.1 (Kumar *et al.* 2004) among population groups. We then used the genetic distance values to evaluate possible divergence times among populations, based on a 2% divergence per million years (My) calibration for *cyt b* divergence of the goby *Evorthodus* across the Isthmus of Panama (Rocha *et al.* 2005).

In order to further analyze the relationships among population sites, pairwise genetic distances were calculated

using Cavalli-Sforza chord distances (Cavalli-Sforza & Edwards 1967) for the combined *cyt b* and msat data in PHYLIP (Felsenstein 1989). Neighbour-joining (NJ) trees (Saitou & Nei 1987) were constructed from the chord distances using PHYLIP. We tested correspondence between chord genetic distance (for the combined *cyt b* and msat loci) and geographical distance, measured as nearest connected waterway paths, using the Mantel (1967) approach (10 000 rearrangements) with ISOLDE in GENEPOP. The fit of the regression line was calculated using Microsoft Excel 2003. Analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) in ARLEQUIN tested for hierarchical population structure by partitioning the total θ covariance among geographical groups (Excoffier *et al.* 1992), including the Caspian Sea vs. Black Sea basins, major river systems and marine vs. freshwater locations.

NJ trees from the mt haplotype sequence data were produced using MEGA. Their phylogeographical patterns were further analyzed using nested clade analysis (NCA; Templeton *et al.* 1995) in GEODIS version 2.5 (Posada *et al.* 2000) to test the null hypothesis of no association among haplotypes and geographical location (Templeton 1998), using a statistical parsimony haplotype network generated with TCS (Clement *et al.* 2000). Ambiguous loops in the network were resolved according to Pfenninger & Posada (2002). Association between the haplotype network and geography was tested with 10 000 permutations. The GEODIS inference key (Templeton 1998; updated by Posada 2005) was used to evaluate the likely cause(s) of associations; such as isolation by distance, range and population expansion/contraction, long distance dispersal, fragmentation and demographic connectivity and shifts.

Both *F*-statistics and contingency tests use the sample location as the unit of comparison, whereas the Bayesian model-based methods of Rannala & Mountain (1997) in GENECLASS 2 version 2.0 (Piry *et al.* 2004) and STRUCTURE version 2.1 (Pritchard *et al.* 2000; Pritchard & Wen 2004) use the individual as the unit, assigning it to the most likely group (population) regardless of geographical origin. This makes these methods particularly useful for assessing likely source populations (Abdelkrim *et al.* 2007; D'Amato *et al.* 2007; Schrey *et al.* 2007). GENECLASS 2 tests were run with simulated population sizes of 10 000 individuals for each sampling site and a 0.01 rejection level (Cornuet *et al.* 1999). An exclusion test was performed to check for false positives. STRUCTURE assigned individuals to groups ranging from $K = 1$ to $K = \text{total } N$ sites, with the relative frequency of predicted group memberships totalling 1.00. Ten independent runs for each K were used, with burn-ins of 100 000 replicates and run lengths of 500 000 replicates. In order to further test the divisions within the msat data, they were subjected to a three-dimensional-factorial correspondence analysis (3D-FCA; Benzecri 1973) using GENETIX version 4.05 (Belkhir *et al.*

2004). That approach has no *a priori* assumptions about populations and shows both within- and among-population variation.

Results

Overall genetic variability and patterns

We sampled 432 individuals and recovered 81 *cyt b* haplotypes (Table 1; GenBank Accession nos EU331156–EU331236), with 17 (21%) characterizing multiple individuals and 64 (79%) singletons. There were 220 msat alleles recovered from seven loci, with 40 (18%) found exclusively in single population locations and 20 of those (9% overall) characterized single individuals. All loci were in HWE following Bonferroni correction and were independent, and no null alleles were detected by the MICROCHECKER analysis.

Most (76%) *cyt b* haplotype substitutions occurred at the third codon position, and there were 21 transversions, 83 transitions (1 : 4) and no indels (Table 1). Numbers of haplotypes per location ranged from 1 to 11 (Lisar, Iran of the Caspian Sea), and averaged 5.3 ± 0.7 . Haplotype (gene) diversity ranged from 0.0 to 0.978 (the latter in two sites in

the Sea of Azov), averaging 0.49 ± 0.07 . Number of msat alleles per locus ranged from 15 (*Ame10*) to 51 (*Ame129*; Table 2B), with a mean of 32. Observed msat heterozygosity per locus ranged from 0.211 (*Ame1*) to 0.777 (*Ame23*) and averaged 0.445. Total heterozygosity per sampling site ranged from 0.374 (the Dnieper River site at Kakhovka, Ukraine) to 0.578 (Sevastopol, Ukraine on the Black Sea) and averaged 0.46 ± 0.01 (Table 1). θ_{ST} values per msat locus ranged from 0.050 (*Ame23*) to 0.311 (*Ame194*) and averaged 0.186 (Table 2B). Pairwise θ_{ST} divergences among sampling sites were about two times greater for *cyt b* data (range = 0.000–1.000, mean = 0.277) than for the msat loci (range = 0.000–0.305, mean = 0.126, Table 4).

Black/Azov Sea vs. Caspian Sea basins

The round goby diverged from its sister species, the monkey goby *Apollonia fluviatilis*, by an average uncorrected pairwise *p*-distance of 0.12 and ~3 My divergence using a molecular clock calibration of ~2%/My for *cyt b* evolution in gobies. Round goby haplotypes were divided into two reciprocally monophyletic clades in the Black and Caspian Sea basins (Fig. 2), separated by eight fixed nucleotide differences

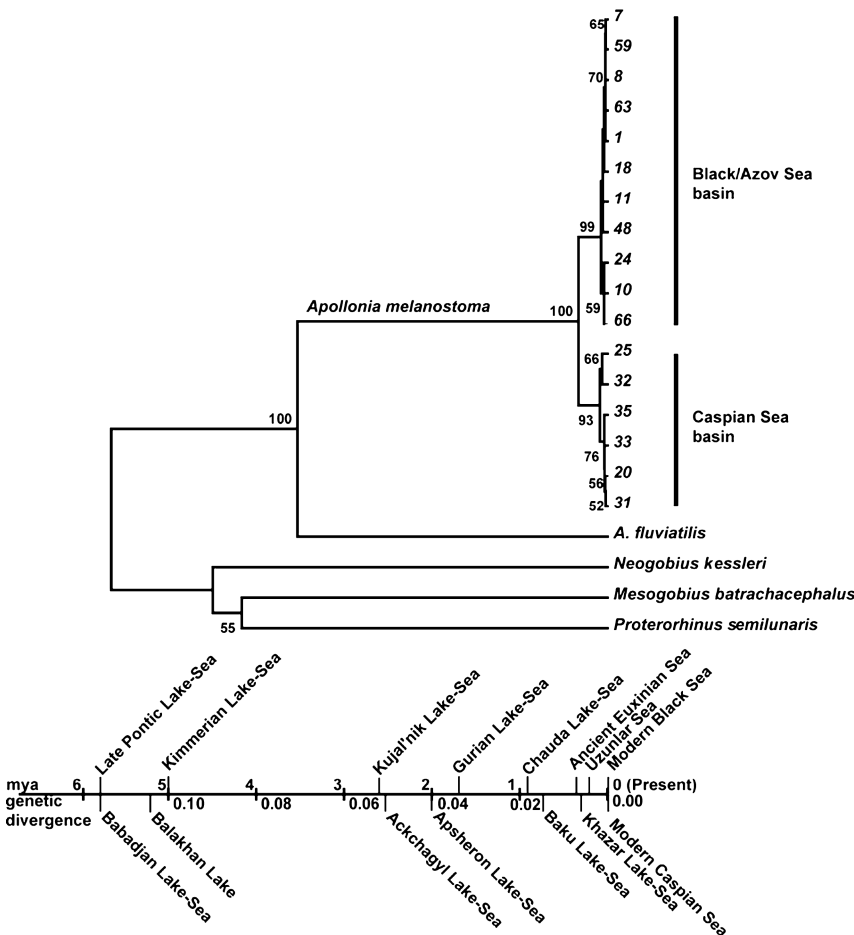


Fig. 2 Neighbour-joining tree of phylogenetic relationships among common Eurasian round goby *Apollonia melanostoma* haplotypes (found in multiple individuals), constructed in MEGA version 3.1 (Kumar *et al.* 2004) and rooted to its monkey goby sister species (*Apollonia fluviatilis*) and three other neogobiin species. Bootstrap percentage $\geq 50\%$ pseudo-replications. Branch lengths are proportional to genetic divergence. Timeline shows major geological Ponto-Caspian events and is proportional to tree length, using a molecular clock calibration of 2%/million years for cytochrome *b* variation in gobies (Rocha *et al.* 2005). The Black and Caspian Sea clades split following the Chauda-Baku faunal exchange ~350 000 years ago.

Table 3 (A) Partitioning of genetic variation between round goby population divisions and among sampling sites within them using AMOVA (Excoffier *et al.* 1992). Measures include Φ_{ST} and % variation between the divisions and among sampling sites within them (for which all $P < 0.0001$), and between group mean uncorrected p -distances for *cyt b*. The remainder percent of the variation is found within the sampling sites. Within group mean uncorrected p -distance was 0.003 for all Black, Azov and Caspian Sea groupings. (B) Summary statistics for regional comparisons.

(A)

	AMOVA (Φ_{ST})		Genetic variation (<i>cyt b</i> /Msat)		Uncorrected p -distance
	<i>Cyt b</i>	Msat	Between divisions	Among sites	<i>Cyt b</i>
Black/Azov Seas vs. Caspian Sea basins	0.839	0.193	81%/8%	3%/12%	0.014
Black Sea freshwater vs. marine vs. Sea of Azov	0.282	0.149	18%/6%	10%/9%	0.004
Caspian Sea vs. Volga River system	0.467	0.122	9%/2%	37%/10%	0.004

(B)

Comparison	Mean N <i>cyt b</i> haplotypes	N of private haplotypes	Gene diversity (<i>Cyt b</i>)	H_o (Msat)
Black/Azov S. Basin	5.9 ± 1.0	3.4 ± 0.7	0.573 ± 0.098	0.459 ± 0.017
Caspian S. Basin	4.3 ± 0.9	2.3 ± 0.6	0.440 ± 0.109	0.467 ± 0.017
Black S. Freshwater	2.6 ± 0.4	0.6 ± 0.2	0.237 ± 0.101	0.410 ± 0.021
Black S. Marine	9.0 ± 0.4	4.8 ± 0.2	0.746 ± 0.074	0.493 ± 0.021
Azov S.	9.0 ± 0.0	7.0 ± 0.0	0.978 ± 0.000	0.516 ± 0.008
Iranian Caspian S.	11.0	7.0	0.774 ± 0.000	0.469 ± 0.000
Azerbaijani Caspian S.	3.5 ± 1.3	1.8 ± 0.6	0.372 ± 0.284	0.421 ± 0.029
Volga R.	3.5 ± 0.9	1.8 ± 0.9	0.425 ± 0.201	0.452 ± 0.036
Native	6.5 ± 0.9	3.7 ± 0.6	0.578 ± 0.082	0.459 ± 0.017
Non-Native	2.7 ± 0.6	1.0 ± 0.4	0.352 ± 0.144	0.423 ± 0.025

and an uncorrected p -distance of 0.014, equivalent to ~350 000 y divergence. This division was additionally supported by $\Phi_{ST} = 0.839$ (*cyt b*) and $\Phi_{ST} = 0.193$ (msat) (AMOVA; Table 3A), assignment-test distinctions (Table 6) and 3D-FCA clustering results. NCA inferred that this division resulted from allopatric fragmentation (clades 4-1 and 4-2, Table 6, Supplementary material, Fig. S1). In contrast, the majority of lower-level clades were defined by restricted gene flow with isolation by distance or by contiguous range expansion (Table 6).

No round goby haplotypes were shared between the Black and Caspian Sea basins, which additionally housed 83 (38%) and 19 (9%) private msat alleles, respectively. Both basins housed similar numbers of haplotypes and private alleles and had similar levels of *cyt b* and msat gene diversity (Table 3B). Within each system, marine locations had more private alleles, including the Kerch Strait (7 *cyt b*, 2 msat), Sea of Azov (7 *cyt b*, 1 msat), and the southern basin of the Caspian Sea (in Iran; 7 *cyt b*, 8 msat).

Mantel-test results based on combined data from *cyt b* and msat loci revealed significant genetic isolation with geographical distance within each basin ($P < 0.0001$ for each; Black Sea $y = 0.1005x - 0.4155$, $R^2 = 0.200$; Caspian Sea $y = 0.1011x - 0.4211$; $R^2 = 0.232$; figures not shown). Assignment tests of combined msat and *cyt b* data using STRUCTURE Bayesian analysis identified 15 round goby population groups (posterior probability = 0.999, Fig. 5).

Most sampling locations assigned to one or two primary groups. Individuals from nonindigenous locations had the greatest assignment levels to single population groups, including the Danube River samples at Slovakia and Serbia, the upper Dnieper River at Kiev and the Volga-Don Canal, along with native Sea of Azov samples. GENECLASS 2 results were congruent with those from STRUCTURE. The 3D-FCA revealed four clusters that were separated by three axes (explaining a total of 43% of the variation; figure not shown). Axis 1 explained 17.95% of the data and separated the Dnieper River, Danube River and Odessa populations from all others. Axis 2 explained 9.43% of the data and split the Dnieper River samples from a cluster containing the Danube River and Odessa samples. Axis 3 explained 8.79% of the data and separated the Baltic Sea and marine Black Sea samples from those in the Sea of Azov, Volga River and Caspian Sea.

Patterns within the Black/Azov Sea basins

Pairwise θ_{ST} estimates between sampling locations in the Black/Azov Sea basin ranged from 0.001 to 0.751 (*cyt b*) and from 0.028 to 0.305 (msat) (Table 4A), indicating additional population genetic structure. A three-way division separating the Black Sea freshwater, Black Sea marine and Sea of Azov population groups was supported by AMOVA (Table 3A) and pairwise comparisons (Tables 4A

Table 4 Pairwise divergences between round goby population samples using θ_{ST} for the Black Sea basin (A) and the Caspian Sea basin (B): *cyt b* below diagonal, microsatellites above diagonal. *, significant ($P < 0.05$); **, remained significant after Bonferroni correction.

(A)

	Danube River		NW Black Sea			Dnieper River			NC Black Sea	Sea of Azov			
	Baltic S.	Slovakia	Serbia	Varna	Bilgorod	Odessa	S. Bug R.	Kiev	Khakhovka	Kherson	Crimea	Kerch Strait	Azov S.
A. Baltic Sea	—	0.260**	0.142**	0.049**	0.077**	0.075**	0.083**	0.270**	0.212**	0.164**	0.062**	0.090**	0.065**
B. Slovakia	0	—	0.116**	0.196**	0.193**	0.115**	0.136**	0.207**	0.229**	0.121**	0.264**	0.290**	0.210**
C. Serbia	0	0.0002	—	0.108**	0.146**	0.028*	0.093**	0.159**	0.109**	0.146**	0.128**	0.186**	0.121**
D. Varna	0.115*	0.198**	0.217**	—	0.028**	0.043**	0.022**	0.237**	0.189**	0.107**	0.033**	0.081**	0.045**
E. Bilgorod	0.085*	0.159**	0.177**	0.073*	—	0.065**	0.040**	0.231**	0.199**	0.099**	0.078**	0.108**	0.038**
F. Odessa	0.165*	0.271**	0.295**	0	0.093*	—	0.046**	0.156**	0.125**	0.094**	0.076**	0.095**	0.055**
H. S. Bug R.	0	0.020*	0.026*	0.086*	0.067*	0.123**	—	0.211**	0.168**	0.071**	0.078**	0.119**	0.055**
I. Kiev	0.656**	0.731**	0.751**	0.282**	0.323**	0.310**	0.392**	—	0.108**	0.184**	0.247**	0.305**	0.281**
J. Khakhovka	0.009	0.023	0.029	0.101*	0.072*	0.145*	0	0.586**	—	0.136**	0.195**	0.243**	0.214**
K. Kherson	0.020	0.061*	0.057*	0.127**	0.093**	0.169**	0.023*	0.451**	0	—	0.173**	0.187**	0.120**
G. Crimea	0.261**	0.375**	0.401**	0.049*	0.180**	0.008	0.220**	0.353**	0.239**	0.264**	—	0.067**	0.041**
L. Kerch Strait	0.549**	0.685**	0.711**	0.145**	0.322**	0.098*	0.437**	0.534**	0.505**	0.493**	0.050	—	0.032*
M. Sea of Azov	0.523**	0.666**	0.692**	0.161**	0.335**	0.126**	0.440**	0.485**	0.482**	0.481**	0.091*	0.019	—

(B)

	Volga River		Azerbaijani Caspian Sea				Iranian Caspian Sea		
	Moskva R.	Volga-Don Canal	Svetli Yar	Damchik	Nabran	Shikh	Lenkoran	Alet	Lisar
N. Moskva River	—	0.110**	0.215**	0.181**	0.225**	0.258**	0.295**	0.290**	0.128**
O. Volga-Don Canal	0.725**	—	0.052*	0.032*	0.075**	0.097**	0.127**	0.107**	0.095**
P. Svetli Yar	0.349*	0.157	—	0.040*	0.079**	0.085**	0.111**	0.113**	0.173**
Q. Damchik	0.811**	0	0.267**	—	0.070**	0.078**	0.092**	0.085**	0.115**
R. Nabran	0.659**	0.503**	0.133*	0.563**	—	0.022*	0.019	0.039*	0.105**
S. Shikh	0.649**	0.468**	0.131*	0.537**	0	—	0	0.011	0.152**
U. Lenkoran	0.601*	0.892**	0.088	0.775**	0	0	—	0	0.143**
T. Alet	0.577*	1.00**	0.049	0.803**	0	0	0	—	0.149**
V. Lisar	0.716**	0.437**	0.191*	0.491**	0.020	0.030	0.051	0.005	—

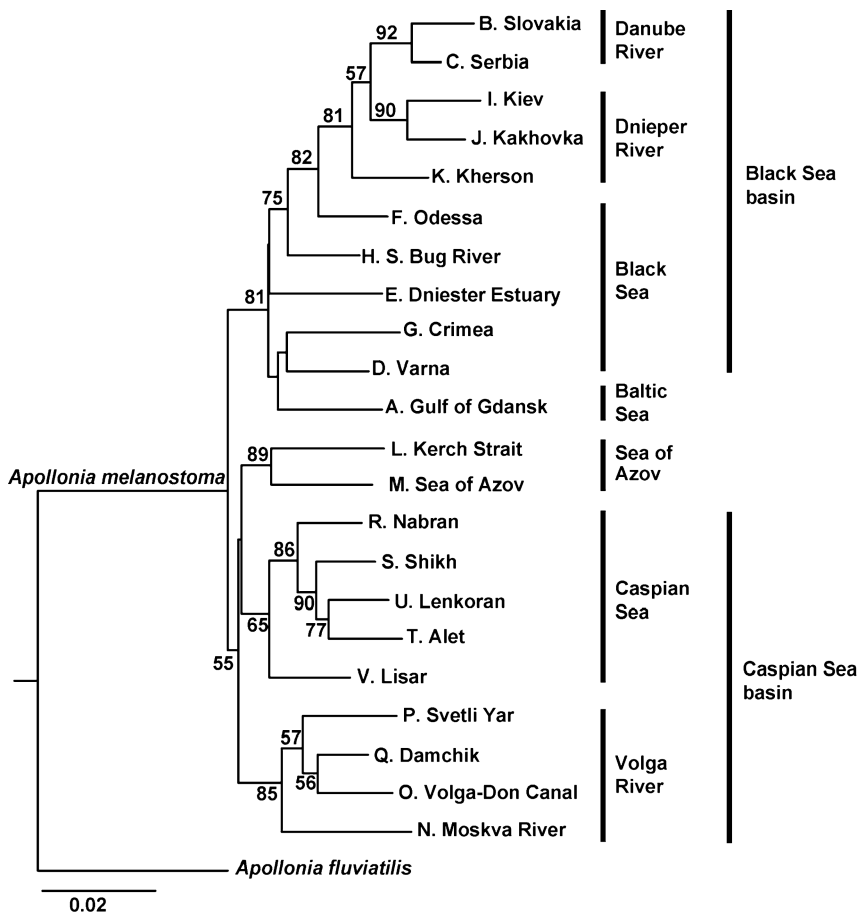


Fig. 3 Genetic distance tree among round goby population sites constructed in PHYLIP v3.6 (Felsenstein 1989) with Cavalli-Sforza chord distances (Cavalli-Sforza & Edwards 1967) and combined mitochondrial cytochrome *b* and microsatellite data. Bootstrap percentage values from 1000 pseudo-replicates. Branch lengths are proportional to genetic divergence.

and 5A). These groups diverged by average *p*-distances of 0.004, corresponding to ~100 000 years (Table 3A). Freshwater locations in the Black Sea had fewer haplotypes and private haplotypes, as well as lower gene diversities than did the two marine groups (Table 3B). The 3D-FCA and the chord distance NJ tree (Fig. 3) grouped the Sea of Azov samples as evolutionarily closer to those from the Volga River and the Caspian Sea.

Five population divisions within the Black/Azov Sea basin were designated by STRUCTURE assignments (Fig. 5). Samples from the western Black Sea, the Crimean Peninsula, the Dnieper River and the Sea of Azov formed distinctive groups. Nonindigenous sites in the Danube River assigned with some samples from the Black Sea at Odessa. Within the Dnieper River, the nonindigenous round goby population from Kiev appeared distinct, although related to those further downstream.

Patterns within the Caspian Sea basin

Within the Caspian Sea basin, samples from Iran had the most *cyt b* haplotypes, more private haplotypes and higher haplotype diversity (Table 3B). Pairwise θ_{ST} comparisons

of Caspian Sea basin samples ranged from 0.000 to 1.000 (*cyt b*) and from 0.000 to 0.295 (msat). Division into Volga River and Caspian Sea population groups was supported by AMOVA (Table 3A), pairwise comparisons (Tables 4B and 5B), the Cavalli-Sforza chord distance tree (Fig. 3) and STRUCTURE assignment tests (Fig. 5). Further division of the Caspian Sea basin round goby samples into three population groups in the Volga River, Azerbaijani Caspian Sea and the Iranian Caspian Sea was indicated by STRUCTURE assignment tests (Fig. 5) and AMOVA (Table 3A), corresponding to $p = 0.004$ (Table 3A) and ~100 000 years.

Non-native locations

Mean gene diversity was greater in native round goby populations than in nonindigenous locations (Tables 1 and 3B; *cyt b*: 0.578 ± 0.082 native vs. 0.352 ± 0.144 nonindigenous; msat: 0.459 ± 0.017 native vs. 0.423 ± 0.025 nonindigenous). However, some exotic population sites were unusually diverse for *cyt b*, including the upper Dnieper River (0.539) and the Moskva River (of the Volga system; 0.857) (see Table 1). Genetic diversity among native populations was highest in the marine sites at the

Table 5 Pairwise contingency test results for round goby samples from the Black/Azov Sea (A) and Caspian Sea (B) basins: Cyt *b* below diagonal, Msats above diagonal. ns, not significant; *, significant ($P < 0.05$); **, remained significant after Bonferroni correction

(A)

	Danube River			NW Black Sea			S. Bug R.	Dnieper River			NC Black Sea	Sea of Azov	
	Baltic S.	Slovakia	Serbia	Varna	Bilgorod	Odessa		Kiev	Khakhovka	Kherson	Crimea	Kerch Strait	Azov S.
A. Baltic Sea	—	*	**	ns	ns	**	*	**	**	ns	ns	ns	**
B. Slovakia	ns	—	**	**	*	*	*	**	**	*	**	*	**
C. Serbia	ns	ns	—	**	**	*	*	**	**	**	**	**	**
D. Varna	*	**	**	—	ns	*	ns	**	**	ns	ns	ns	ns
E. Bilgorod	*	**	**	*	—	**	ns	**	*	ns	ns	ns	*
F. Odessa	**	**	**	ns	ns	—	ns	**	**	*	**	ns	*
H. S. Bug R.	ns	*	*	*	*	*	—	**	*	ns	ns	ns	ns
I. Kiev	**	**	**	**	**	*	**	—	**	**	**	**	**
J. Khakhovka	ns	ns	ns	ns	*	*	ns	**	—	ns	*	**	*
K. Kherson	ns	*	*	*	**	ns	ns	**	ns	—	ns	ns	ns
G. Crimea	**	**	**	*	**	ns	**	**	**	**	—	ns	*
L. Kerch Strait	**	**	**	**	*	*	**	**	**	**	ns	—	ns
M. Sea of Azov	**	**	**	*	**	*	**	**	**	**	*	ns	—

(B)

	Moskva R.		Volga River		Azerbaijani Caspian Sea				Iranian Caspian Sea	
	Moskva R.	Volga-Don Canal	Svetli Yar	Damchik	Nabran	Shikh	Lenkoran	Alet	Lisar	
N. Moskva River	—	**	**	**	**	**	**	**	**	
O. Volga-Don Canal	**	—	*	*	**	**	**	**	**	
P. Svetli Yar	*	ns	—	*	**	**	**	**	**	
Q. Damchik	**	ns	ns	—	**	**	**	**	**	
R. Nabran	**	**	**	**	—	*	ns	*	**	
S. Shikh	**	**	**	**	ns	—	ns	ns	**	
U. Lenkoran	*	**	*	**	ns	ns	—	ns	**	
T. Alet	*	**	*	**	ns	ns	ns	—	**	
V. Lisar	**	**	**	**	ns	ns	ns	ns	—	

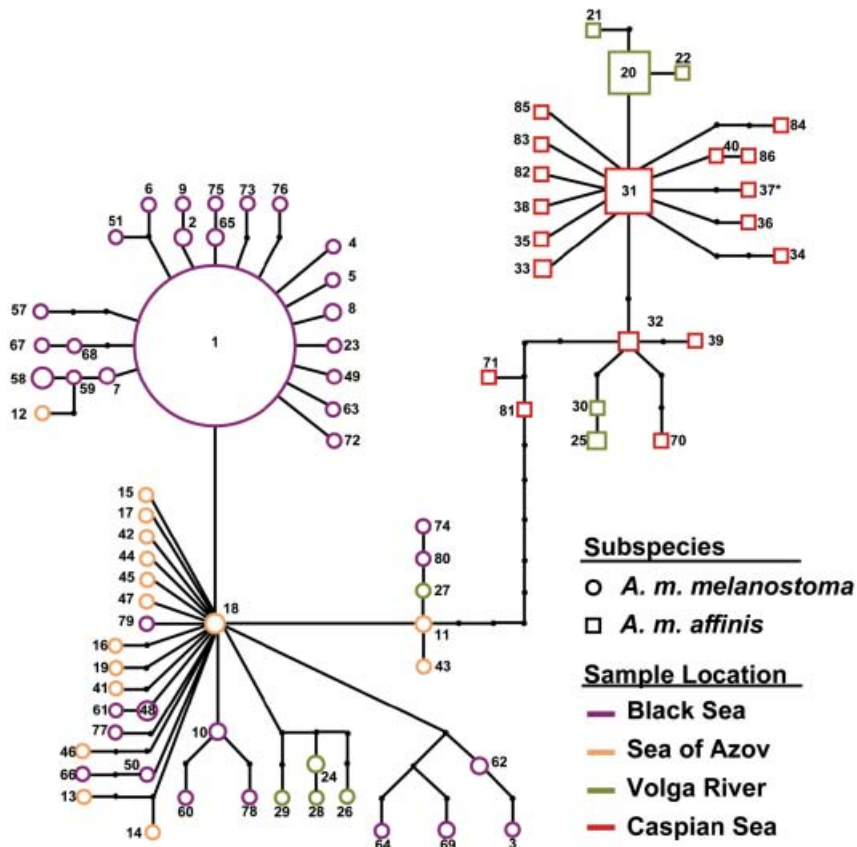


Fig. 4 Parsimony network among round goby *Apollonia melanostoma* cytochrome *b* haplotypes. Symbol size is proportional to haplotype frequency and are coloured to indicate location. The sister species, the monkey goby *A. fluviatilis*, connects to the network at haplotype 37 (*), after < 138 mutational steps. The nesting diagram is available in Fig. S1.

Kerch Strait and the Sea of Azov (cyt *b*, 0.978; msat: 0.509 and cyt *b*, 0.978; msat, 0.586).

The exotic Baltic Sea population site

The nonindigenous Baltic Sea population site had relatively low gene diversity (cyt *b*: 1 haplotype and $h = 0.000$, msat: $h = 0.379$; Table 1) and clustered with samples from the Black Sea basin (Figs 3 and 4). Pairwise tests based on cyt *b* showed that the Baltic Sea sample significantly differed from those in the Crimean Peninsula (Black Sea) and the Sea of Azov (Tables 4A and 5A). However, the Baltic Sea sample was statistically indistinguishable from other locations in the remainder of the Black Sea and its rivers using cyt *b*. Pairwise θ_{ST} (Table 4A) and contingency tests (Table 5A) using msat data showed significant divergence of the Baltic Sea sample from all other locations. Assignment tests also discerned the Baltic Sea sample as different from all others, which precluded identification of its source (Fig. 5).

Black Sea non-native locations

The two non-native round goby samples from the Danube River in Slovakia and Serbia had low mt gene diversity

(0.051 and 0.044, respectively). However, their msat gene diversities (0.377 and 0.459, respectively) were comparable to mean values for native locations. Nonindigenous sites in the Danube River assigned with each other and with samples from the Black Sea at Odessa in the STRUCTURE (Fig. 5) analysis and in the 3D-FCA. The non-native sample in the Dnieper River at Kiev had very high gene diversity (cyt *b*: 0.539, msat: 0.369) and a private cyt *b* allele. Both STRUCTURE (Fig. 5) and the 3D-FCA grouped it with native locations further downstream, while also highlighting its distinctiveness.

Caspian Sea non-native locations

Of the three exotic sites in the Caspian Sea basin, the Volga–Don Canal sample had very low gene diversity and the introduced location from the Moskva River had very high gene diversity (cyt *b*, 0.857; msat, 0.425) – higher than almost all of the native locations sampled. The sample from the central Volga River near Svetli Yar also housed high genetic variability (Table 1). Introduced round gobies in the Volga–Don Canal, the Volga River and the Moskva River appeared related to native population locations nearer the mouth of the Volga River. However, NCA additionally revealed a lineage of cyt *b* haplotypes that

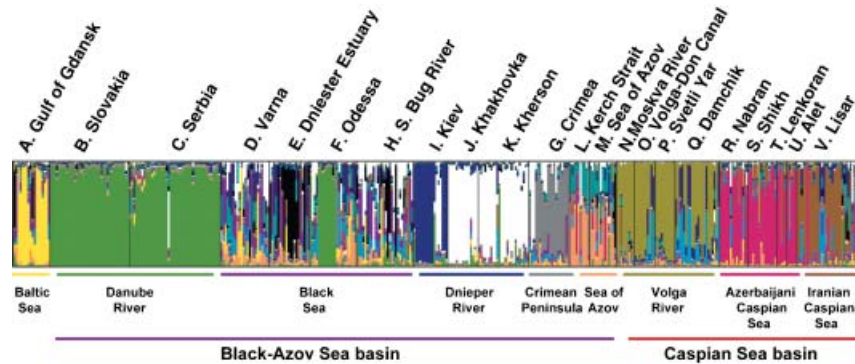


Fig. 5 Bayesian STRUCTURE version 2.1 (Pritchard *et al.* 2000; Pritchard & Wen 2004) analysis of round goby populations using combined data from seven microsatellite loci and cytochrome *b* sequences. $K = 15$ (posterior probability, $pp = 0.999$). Each individual is represented by a thin vertical line, which is partitioned into K coloured segments that represent the individual's estimated membership fractions. Black lines separate individuals from different sampling sites, which are labelled above the figure. Ten STRUCTURE runs at each K produced nearly identical individual membership coefficients, having pairwise similarity coefficients above 0.95, and the figure illustrating a given K is based on the highest probability run at that K . Results show five main population genetic divisions in the Black/Azov Sea basin and three within the Caspian Sea basin.

was shared between samples from the Moskva River and the Volga River at Svetli Yar, and appeared related to those from the Black/Azov Sea population group (clade 3-6, Fig. S1). There thus are two disparate Volga River lineages – one descendent from the Caspian Sea and one similar to the Black/Azov Sea group. NCA was unable to discern whether they originated from contiguous range expansion, long distance colonization or past fragmentation (Table 6). They likely reflect anthropogenic activities, such as long-distance transport and the breaching of watershed boundaries with canals.

Discussion

Ancient divisions

Phylogeography of the Eurasian round goby revealed two major lineages within its native range, corresponding to the two primary Ponto-Caspian basins – the Black/Azov Sea and the Caspian Sea – which likely diverged following a faunal exchange between the Euxinian (Black) and the Khazar (Caspian) Sea basins ~350 000 years ago (Zubakov 1988). A subspecies divergence between *Apollonia m. melanostoma* in the Black Sea and *A. m. affinis* in the Caspian Sea was once recognized (Navozov 1912; Berg 1949), with the Caspian Sea subspecies having lower scale counts (Iljin 1938) and a lower craniological index (see Pinchuk *et al.* 2003). Our results show that the two lineages are markedly divergent based on both mt and nuclear DNA data, distinguished by eight fixed *cyt b* substitutions, and are reciprocally monophyletic (Fig. 2). Our study thus supports their taxonomic distinctiveness (Pinchuk *et al.* 2003), and they should be recognized as subspecies and further evaluated for possible elevation to species.

Marked genetic divergences between round goby taxa from the Black/Azov and Caspian Sea basins are congruent with those of other fishes and invertebrates. They diverged in separate glacial refugia in the Black and Caspian Sea basins and then followed separate glacial meltwater pathways along the Danube, Dnieper and Volga Rivers. Examples of invertebrates include planktonic cladocerans (~1 Ma) and the amphipod *Pontogammarus crassus* (~1–1.6 Ma; Cristescu *et al.* 2003). Fishes include populations of ruffe *Gymnocephalus cernuus* across Eurasia that showed pronounced mtDNA genetic divergence between lineages in the Black and Caspian Sea basins dated as ~500 000 ± 180 000 years (Stepien *et al.* 1998), comparable to the separation estimated here and attributed to the most extensive cold period during the glaciations (see Rohling *et al.* 1998). Similarly, Culling *et al.* (2006) discerned Ponto-Caspian clades of the spined loach *Cobitis taenia*, with separation between the Black Sea rivers and the Volga River system dated to ~500 000 years ago. Black Sea and Caspian/Aral Sea clades of the brown trout *Salmo trutta* were estimated to have diverged more recently (~150 000–310 000 years ago) based on nested clade analysis (Bernatchez 2001).

Other studies of divergence between the Black and Caspian Sea basins discerned older patterns, dating to their formation ~5 Ma (Dumont 1998). Work in our laboratory on the monkey goby *Apollonia fluviatilis* (the sister species of the round goby; Stepien & Tumeo 2006) shows a much older division between Black/Azov and Caspian Sea locations, dating to ~3.75 Ma (M. Neilson and C. A. Stepien, unpublished). The chub *Leuciscus cephalus* is divided in two Pliocene-dated Ponto-Caspian clades; one in the Danube River system and the other in the Caspian Sea basin (Durand *et al.* 1999). Shallower divergence was found for zebra mussel populations dating to ~100 000–200 000 years

Table 6 Nested clade analysis of round goby cytochrome *b* haplotypes showing significant nesting clades and significant subclades, clade dispersion (D_C) and displacement values (D_N), the inference chain, inference results and map locations (lettered according to Fig. 1, with Caspian Sea locations italicized). For clade dispersion and displacement values, (S) = a significantly small value, (L) = a significantly large value. Figure is in Fig. S1.

Nesting clade	Significant subclades	D_C	D_N	Inference chain	Result	Map locations
Total	4-1	(S) $P < 0.0001$	(S) $P < 0.0001$	1,19 – N	Allopatric Fragmentation	A*,B*,C*,D,E,F,G,H,I*,J,K,L,M,N*,P*,N*,O*,P*,Q,R,S,T,U,V
	4-2	(S) $P < 0.0001$	(L) $P < 0.0001$			
	1-T	(S) $P = 0.3190$	(L) $P < 0.0001$			
4-1	3-1	(S) $P = 0.0180$	(S) $P = 0.0130$	1,2,3,4 – N	Restricted Gene Flow with Isolation by Distance	A*,B*,C*,D,E,F,G,H,I*,J,K,M D,E,F,G,H,L,M,N*,P* E,F,G
	3-2		(L) $P = 0.0750$			
	3-3		(S) $P = 0.0270$			
	1-T		(L) $P = 0.0180$			
4-2	3-5	(L) $P = 0.0010$	(L) $P = 0.0080$	1,2,3,4 – N	Restricted Gene Flow with Isolation by Distance	N*,P*,R,S,U O*,P*,Q,R,S,T,U,V
	3-6	(S) $P = 0.0100$	(S) $P = 0.0110$			
	1-T	(L) $P = 0.0010$	(L) $P = 0.0080$			
3-1	2-1	(L) $P = 0.0010$	(L) $P = 0.0020$	1,2,3,4 – N	Restricted Gene Flow with Isolation by Distance	A*,B*,C*,D,E,F,G,H,I*,J,K C*,H,I*,K,M
	2-2	(S) $P = 0.0030$	(S) $P = 0.0160$			
	1-T	(L) $P = 0.0030$	(L) $P = 0.0120$			
3-2	2-3	(S) $P < 0.0001$	(S) $P < 0.0001$	1,2,11,12 – N	Contiguous Range Expansion	E,F,H,G,L,M M D,E,F N*,P*
	2-4	(S) $P = 0.0130$	(S) $P = 0.0210$			
	2-5					
	2-6		(L) $P < 0.0001$			
	1-T		(S) $P = 0.0010$			
	3-6		(S) $P < 0.0001$			
2-13	(S) $P < 0.0001$		(L) $P < 0.0001$			
2-1	1-2	(S) $P < 0.0001$	(S) $P < 0.0001$	1, 2, 3, 4 – N	Restricted Gene Flow with Isolation by Distance	E E,G A*,B*,C*,D,E,F,G,H,I*,J,K
	1-3	(S) $P = 0.0260$				
	1-8	(L) $P < 0.0001$	(L) $P < 0.0001$			
	1-T	(L) $P < 0.0001$	(L) $P < 0.0001$			
2-3	1-11	(S) $P = 0.0370$		1, 2, 11, 12 – N	Contiguous Range Expansion	E,F,G,H,L,M E,G,L
	1-15	(S) $P = 0.0370$				
2-11	1-36	(S) $P = 0.0060$		1, 2, 11, 12 – N	Contiguous Range Expansion	S,V
	1-T	(S) $P = 0.0280$				
1-8	1	(L) $P = 0.0020$	(L) $P = 0.0030$	1, 2, 3, 4 – N	Restricted Gene Flow with Isolation by Distance	A*,B*,C*,D,E,F,G,H,I*,J,K J,K
	8	(S) $P = 0.0160$				
	1-T	(L) $P = 0.0040$	(L) $P = 0.0020$			
1-9	7		(L) $P = 0.0030$	1, 2, 3, 4 – N	Restricted Gene Flow with Isolation by Distance	C,K H,I*
	58	(S) $P = 0.0110$	(S) $P = 0.0110$			
	1-T	(L) $P = 0.0210$	(L) $P = 0.0050$			
1-38	1-T	(S) $P = 0.0260$		1, 2, 11, 12 – N	Contiguous Range Expansion	R,S,T,U,V

according to Stepien *et al.* (2003, 2005) and Gelembiuk *et al.* (2006), corresponding to the closure of the most recent connection between the Black and Caspian Sea basins (Reid & Orlova 2002).

Patterns within the Black/Azov Sea basin. Within the Black/Azov Sea region, there are significant divergences among round goby population groups that correspond to geographical features, notably basins and tributaries. Round goby populations in the Black Sea rivers, the Black Sea proper and the Sea of Azov significantly differ, dating to ~100 000 years ago. Their resolution is shallow, and may have resulted from expansions into river systems. The Crimean Peninsula in the Black Sea houses round goby lineages from both the Black Sea and the Sea of Azov, reflecting its intermediate geographical location.

Patterns within the Caspian Sea basin. We identified three round goby population groups in the Caspian Sea basin – one in the Volga River and two within the Caspian Sea proper, dating to 100 000 years. Despite expansions within the Volga River watershed, no Caspian derived lineages were found in the locations we sampled in the Black and Azov Seas. The two Caspian Sea groups may have differentiated in the central and southern deep basins, which likely served as refugia during low water periods.

Round goby populations within the Caspian Sea and Black/Azov Sea basins thus differentiated about the same time. About 100 000 years ago, the Riss Glaciation lowered Caspian water levels 50 m below sea level and the Black Sea became salinized via contact with the Mediterranean Sea (Reid & Orlova 2002). These events may have isolated the respective taxa into groups within these basins. Additional research is needed to further explore these fine-scale patterns.

Recent expansions

We sampled four sets of non-native populations in this study: the Baltic Sea, the Slovak and Serbian portions of the Danube River, the upper Dnieper River near Kiev and portions of the Volga River system. Nonindigenous round goby populations in central Europe appear derived from the Black Sea primary clade, whereas those in the Volga River comprise a genetic admixture from native Volga River locations and the Black/Azov Sea region.

Msat data similarities suggest that the Dniester and Dnieper Rivers in the northwest Black Sea were likely donors to the Baltic Sea nonindigenous population site and should be investigated further. This assignment gains additional support from the study of the round goby invasion by Dillon & Stepien (2001) using the mtDNA control region, which eliminated the western Black Sea region of Varna, Bulgaria, as a likely donor source for the Baltic Sea

population site but did not have samples available from the Dnieper and Dniester Rivers. Round gobies from the Dniester and Dnieper Rivers thus should be sequenced for the control region.

Nonindigenous round goby sites in the Danube River at Bratislava and Prahovo had relatively low genetic diversity in our study. Similarly low genetic diversity was found in an expansion population of the racer goby *Neogobius gymnotrachelus* from the Bratislava site (Ohayon & Stepien 2007). In contrast, a racer goby sample from another nonindigenous Danube River location located southward at Tekija, Serbia, had much higher diversity. Exotic populations of the racer goby thus appeared to lose genetic variation as they progressed northward along the Danube River (Ohayon & Stepien 2007).

The nonindigenous round goby population in the upper Dnieper River contrasts with those from the Baltic Sea and the Danube River in having higher genetic diversity (cyt *b*, 0.539; msat, 0.369) and a private cyt *b* allele. The differences observed between this location and others downstream are likely due to founder effects and colonization admixture resulting from extensive shipping activity, as the Dnieper River is a commercially important waterway.

Round goby samples from the Volga–Don Canal had a single cyt *b* haplotype that also occurred in the central Volga River and the Volga River delta. This was surprising, since the water filling the canal reservoirs originates from the Don River (V. Boldyrev, State Institute of Lake and River Fisheries, Volgograd, Russia, personal communication). However, Audzijonyte *et al.* (2006) found that mysid shrimp populations in the Canal were of Volga River origin, despite stockings from the Don River.

Round goby samples from the Moskva River and the central Volga River contained both *A. m. melanostoma* (the Black Sea basin subspecies that we are resurrecting here) and *A. m. affinis* (the Caspian Sea basin subspecies) and consequently had high gene diversity. The Black Sea taxon likely reached the Volga River via human activities, whereas the Caspian Sea taxon likely spread upriver from the Volga River delta. The Volga River and its tributaries need to be further sampled to determine whether the taxa hybridize.

Summary and conclusions

Our findings support long-term primary genetic division between round goby taxa in the Black/Azov and Caspian Sea basins leading to their differentiation as *Apollonia melanostoma melanostoma* and *A. m. affinis*, respectively. These taxa should be recognized as valid subspecies and further evaluated for possible elevation to species. Both have expanded beyond their historical distributions due to human activities, especially during the past two decades (Bij de Vaate *et al.* 2002; Stepien *et al.* 2005; Ricciardi 2006).

Expansion populations in central and eastern Europe are composed of *A. m. melanostoma*. Both taxa co-occur in the Volga River system, making this region potentially valuable for future evolutionary studies and worthy of more intensive research. Our results highlight how glacial and anthropogenic disturbances have shaped the genetic heritage of a fish that has recently escaped the Ponto-Caspian and rapidly expanded its range on two continents.

Acknowledgements

The project was funded by NSF DEB-0456972, USEPA CR-83281401-0 and Lake Erie Protection Fund 02-14 to CAS. Samples were provided by S. Abdolmalaki, N. Bogutskaya, V. Boldyrev, I. Grigorovich, S. Ibrahimov, H. Jenner, D. Jude, V. Kovac, Y. Kvach, M. Neilson, J. Ram, S. Rudnicka, M. Sapota, P. Simonovic, A. Smirnov and C. Wiesner. Data collection was aided by Great Lakes Genetics Laboratory technicians D. Murphy and C. Taylor. We especially thank the late J. Coss of the Lake Erie Center for G.I.S. support and the maps, and P. Uzmans for administrative assistance. We thank B. Bodamer, J. Bossenbroek, L. Corkum, M. Diaz, A. Haponski, R. Lohner, C. Mayer, J. Miner, M. Neilson and J. Sepulveda Villet for valuable comments on the manuscript. This is publication 2008-004 from the Lake Erie Research Center.

References

- Abdelkrim J, Pascal M, Samadi S (2007) Establishing causes of eradication failure based on genetics: case study of ship rat eradication in Ste. Anne Archipelago. *Conservation Biology*, **21**, 719–730.
- Akihito, Iwata A, Kobayashi T *et al.* (2000) Evolutionary aspects of gobioid fishes based upon a phylogenetic analysis of mitochondrial cytochrome *b* genes. *Gene*, **259**, 5–15.
- Audzijonyte A, Daneliya ME, Vainola R (2006) Comparative phylogeography of Ponto-Caspian mysid crustaceans: isolation and exchange among dynamic inland sea basins. *Molecular Ecology*, **15**, 2969–2984.
- Bahr A, Lamy F, Arz H *et al.* (2005) Late glacial to Holocene climate and sedimentation history in the NW Black Sea. *Marine Geology*, **214**, 309–322.
- Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**, 155–165.
- Belkhir K, Borsari P, Chikhi L *et al.* (2004) GENETIX 4.05, logiciel sous Windows TM pour la genétique des populations. Laboratoire Genome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier. <http://www.genetix.univ-montp2.fr/genetix/genetix.htm>.
- Benzecri JP (1973) *L'analyse Des Donneés, Tome II: L'analyse Des Correspondances*. Dunod Press, Paris.
- Berg LS (1949) *Freshwater Fishes of the U.S.S.R. & Adjacent Countries*, 4th edn. **3**, 927–1382. Izvestia Akademiia Nauk SSSR, Moscow (in Russian).
- Bernatchez L (2001) The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analysis of mitochondrial DNA variation. *Evolution*, **55**, 351–379.
- Bij de Vaate A, Jazdzewski K, Ketelaars HAM *et al.* (2002) Geographical patterns in range extension of Ponto-Caspian macroinvertebrate species in Europe. *Canadian Journal of Fisheries and Aquatic Sciences*, **59**, 1159–1174.
- Briggs JC (1974) *Marine Zoogeography*. McGraw-Hill, Inc., New York.
- Cavalli-Sforza LL, Edwards AW (1967) Phylogenetic analysis. Models and estimation procedures. *American Journal of Human Genetics*, **19**, 233–257.
- Choudhury A, Dick TA (1998) The historical biogeography of sturgeons (*Osteichthyes: Acipenseridae*): a synthesis of phylogenetic, palaeontology, and palaeogeography. *Journal of Biogeography*, **25**, 623–640.
- Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.
- Cornuet JM, Piry S, Luikart G *et al.* (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989–2000.
- Costedoat C, Chappaz R, Barascud B *et al.* (2006) Heterogeneous colonization pattern of European Cyprinids, as highlighted by the dace complex (Teleostei: Cyprinidae). *Molecular Phylogenetics and Evolution*, **41**, 127–148.
- Cristescu MEA, Hebert PDN, Onciu TM (2003) Phylogeography of Ponto-Caspian crustaceans: a benthic-planktonic comparison. *Molecular Ecology*, **12**, 985–996.
- Culling MA, Janko K, Boron A *et al.* (2006) European colonization by the spined loach (*Cobitis taenia*) from Ponto-Caspian refugia based on mitochondrial DNA variation. *Molecular Ecology*, **15**, 173–190.
- D'Amato ME, Esterhuyse MM, van der Waal BCW *et al.* (2007) Hybridization and phylogeography of the Mozambique tilapia *Oreochromis mossambicus* in southern Africa evidenced by mitochondrial and microsatellite DNA genotyping. *Conservation Genetics*, **8**, 475–488.
- Dillon AK, Stepien CA (2001) Genetic and biogeographic relationships of the invasive round (*Neogobius melanostomus*) and tubenose (*Proterorhinus marmoratus*) gobies in the Great Lakes versus Eurasian populations. *Journal of Great Lakes Research*, **27**, 267–280.
- Dougherty JD, Moore WS, Ram JL (1996) Mitochondrial DNA analysis of round goby (*Neogobius melanostomus*) and tubenose goby (*Proterorhinus marmoratus*) in the Great Lakes basin. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 474–480.
- Dumont HJ (1998) The Caspian Lake: history, biota, structure, and function. *Limnology and Oceanography*, **43**, 44–52.
- Durand JD, Persat H, Bouvet Y (1999) Phylogeography and postglacial dispersal of the chub (*Leuciscus cephalus*) in Europe. *Molecular Ecology*, **8**, 989–997.
- Engbrecht CC, Freyhof J, Nolte A *et al.* (2000) Phylogeography of the bullhead *Cottus gobio* (Pices: Teleostei: Cottidae) suggests a pre-Pleistocene origin of the major central European populations. *Molecular Ecology*, **9**, 709–722.
- Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Excoffier L, Laval G, Schneider S (2005) ARLEQUIN ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50. <http://cmpg.unibe.ch/software/arlequin3/>.
- Felsenstein J (1989) PHYLIP – phylogeny inference package Version 3.2. *Cladistics*, **5**, 164–166.
- Gelembiuk GW, May GE, Lee CE (2006) Phylogeography and systematics of zebra mussels and related species. *Molecular Ecology*, **15**, 1033–1050.

- Goudet J, Raymond M, de Meeus T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933–1940.
- Griffiths D (2006) Pattern and process in the ecological biogeography of European freshwater fish. *Journal of Animal Ecology*, **75**, 734–751.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95.
- Hall T (2004) BioEdit v. 7.0.0. (Available online at www.mbio.ncsu.edu/BioEdit/).
- Hanfling B, Hellemans B, Volckaert FAM, Carvalho GR (2002) Late glacial history of the cold-adapted freshwater fish *Cottus gobio*, revealed by microsatellites. *Molecular Ecology*, **11**, 1717–1729.
- Hensler SR, Jude DJ (2007) Diel vertical migration of round goby larvae in the Great Lakes. *Journal of Great Lakes Research*, **33**, 295–302.
- Hewitt GM (2004) Genetic consequences of climatic oscillations in the. *Quaternary. Philosophical Transactions of the Royal Society of London, Series B: Biological Science*, **359**, 183–195.
- Ilijin BS (1938) Gobies (Gobiidae) based upon materials collected by the USSR Academy of Sciences Expedition to Mertvyi Kultuk and Kaidak. *Trudy Komm. Kompl. Izuch. Kasp. Moria (Proceedings of the Commission for Complex Investigation of the Caspian Sea)*, **2**, 111–131 (in Russian).
- Jude DJ, Reider RH, Smith GR (1992) Establishment of Gobiidae in the Great Lakes basin. *Canadian Journal of Fisheries and Aquatic Sciences*, **49**, 416–421.
- Kostyuchenko AA (1969) Regularity of distribution and migrations of the round goby in the Sea of Azov. *Trudy Azov.-Chernom. Nauchn.-Issled. Rybn. Khoz. Okeanogr. (Proceedings of the Commission for Complex Investigation of the Caspian Sea)*, **26**, 14–19 (in Russian).
- Kumar S, Tamura K, Jakobsen IB, Nei M (2004) MEGA vers. 3. Molecular Evolutionary Genetics Analysis software, Arizona State University, Tempe, AZ.
- MacInnis AJ, Corkum LD (2000) Fecundity and reproductive season of the round goby *Neogobius melanostomus* in the upper Detroit River. *Transactions of the American Fisheries Society*, **129**, 135–144.
- Major CO, Goldstein SL, Ryan WBF *et al.* (2006) The co-evolution of Black Sea level and composition through the last deglaciation and its paleoclimatic significance. *Quaternary Science Reviews*, **25**, 2031–2047.
- Mangerud J, Jakobsson M, Alexanderson H *et al.* (2004) Ice-dammed lakes and rerouting of the drainage of northern Eurasia during the Last Glaciation. *Quaternary Science Reviews*, **23**, 1313–1332.
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.
- Meyer A, Kocher TD, Bsbwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, **347**, 550–553.
- Navozov NP (1912) Materials for an ichthyofauna of the Ural River basin. *Vestnik Ryby*, **1912**, 252–283 (in Russian).
- Nesbø CL, Fosshem T, Vøllestad LA, Jakobsen KS (1999) Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology*, **8**, 1387–1404.
- Ohayon JL, Stepien CA (2007) Genetic and biogeographic relationships of the racer goby *Neogobius gymnotrachelus* (Gobiidae: Teleostei) from introduced and native Eurasian locations. *Journal of Fish Biology*, **71** (Suppl. C), 360–370.
- Pfenniger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution*, **56**, 1776–1788.
- Pinchuk VI, Vasil'eva ED, Vasil'ev VP, Miller PJ (2003) *Neogobius melanostomus* (Pallas, 1814). In: *The Freshwater Fishes of Europe, Vol. 8/I: Mugilidae, Atherinidae, Atherinopsidae, Blenniidae, Odontobutidae, Gobiidae I* (ed. PJ Miller), pp. 293–345. AULA-Verlag, Wiesbaden.
- Piry S, Alapetite A, Cornuet J-M *et al.* (2004) GENECLASS 2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, **95**, 536–539. <http://www.montpellier.inra.fr/CBGP/software/>.
- Pons O, Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity I. Haploid locus. *Theoretical and Applied Genetics*, **90**, 462–470.
- Posada D (2005) GeoDis Inference Key. <http://darwin.uvigo.es/software/geodis.html>.
- Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.
- Pritchard JK, Wen W (2004) Documentation for STRUCTURE software: Version 2. Department of Human Genetics, University of Chicago, Chicago, IL. <http://pritch.basd.uchicago.edu>.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences, USA*, **94**, 9197–9201.
- Ray WJ, Corkum LD (2001) Habitat and site affinity of the round goby. *Journal of Great Lakes Research*, **27**, 329–334.
- Raymond M, Rousset F (1995) GENEPOP (version 3.4): Population genetics software for exact test and ecumenicism. *Journal of Heredity*, **86**, 248–249.
- Raymond M, Rousset F (2004) GENEPOP (version 3.4): Population genetics software for exact tests and ecumenicism. Laboratoire de Genetique et Environment, Montpellier, France. <http://wbiomed.curtin.edu.au/genepop>.
- Reid DF, Orlova MI (2002) Geological and evolutionary underpinnings for the success of Ponto-Caspian species invasions in the Baltic Sea and North American Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 2513–2525.
- Reyjol Y, Huguency B, Pont D *et al.* (2007) Patterns in species richness and endemism of European freshwater fish. *Global Ecology and Biogeography*, **16**, 65–75.
- Ricciardi A (2006) Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. *Diversity and Distributions*, **12**, 425–433.
- Ricciardi A, MacIsaac HJ (2000) Recent mass invasion of the North American Great Lakes by Ponto-Caspian species. *Trends in Ecology and Evolution*, **15**, 62–65.
- Rocha LA, Robertson DR, Rocha CR, Van Tassell JL, Craig MT, Bowen BW (2005) Recent invasion of the tropical Atlantic by an Indo-Pacific coral reef fish. *Molecular Ecology*, **14**, 3921–3928.
- Rohling EJ, Fenton M, Jorissen FJ *et al.* (1998) Magnitudes of sea-level lowstands of the past 500 000 years. *Nature*, **394**, 162–165.
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.
- Schrey AW, Sloss BL, Sheehan RJ, Heidinger RC, Heist EJ (2007) Genetic discrimination of middle Mississippi River *Scaphirhynchus* sturgeon into pallid, shovelnose, and putative hybrids with multiple microsatellite loci. *Conservation Genetics*, **8**, 683–693.

- Skora KE, Stolarski J (1993) New fish species in the Gulf of Gdansk, *Neogobius* sp [cf. *Neogobius melanostomus* (Pallas 1811)]. *Bulletin of the Sea Fisheries Institute*, **1**, 83–84.
- Sokal RR, Rohlf FJ (1995) *Biometry*, 3rd edn. W.H. Freeman, Inc., New York.
- Stepien CA, Tumeo MA (2006) Invasion genetics of Ponto-Caspian gobies in the Great Lakes: a 'cryptic' species, absence of founder effects, and comparative risk analysis. *Biological Invasions*, **8**, 61–78.
- Stepien CA, Dillon AK, Chandler MD (1998) Genetic identity, phylogeography, and systematics of ruffe *Gymnocephalus* in the North American Great Lakes and Eurasia. *Journal of Great Lakes Research*, **24**, 361–378.
- Stepien CA, Taylor CD, Dabrowska KA (2002) Genetic variability and phylogeographic patterns of a nonindigenous species invasion: a comparison of exotic versus native zebra and quagga mussel populations. *Journal of Evolutionary Biology*, **15**, 314–328.
- Stepien CA, Taylor CD, Grigorovich IA *et al.* (2003) DNA and systematic analysis of invasive and native dreissenid mussels: is *Dreissena bugensis* really *D. rostriformis*? *Aquatic Invaders*, **14**, 1–10.
- Stepien CA, Brown JE, Neilson ME, Tumeo MA (2005) Genetic diversity of invasive species in the Great Lakes versus their Eurasian source populations: insights for risk analysis. *Risk Analysis*, **25**, 1043–1060.
- Stepien CA, Murphy DJ, Strange RM (2007) Broad- to fine-scale population genetic patterning in the smallmouth bass *Micropterus dolomieu* across the Laurentian Great Lakes and beyond: an interplay of behaviour and geography. *Molecular Ecology*, **16**, 1605–1624.
- Svetovidov AN (1964) [*Fishes of the Black Sea*]. Nauka, Moscow. (in Russian).
- Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.
- Templeton AR, Routman E, Phillips C (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767–782.
- Therriault TW, Orlova MI, Docker MF *et al.* (2005) Invasion genetics of a freshwater mussel (*Dreissena rostriformis bugensis*) in eastern Europe: high gene flow and multiple introductions. *Heredity*, **95**, 16–23.
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680. <http://bips.u-strasbg.fr/fr/Documentation/ClustalX/>.
- Van Oosterhout C, Hutchison WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.
- Van Oosterhout C, Hutchison WF, Wills DPM, Shipley P (2006) micro-checker. URL www.microchecker.hull.ac.uk.
- Waters JM, Lintermans M, White RWG (1994) Mitochondrial DNA variation suggests river capture as source of vicariance in *Gadopsis bispinosus* (Pisces: Gadopsidae). *Journal of Fish Biology*, **44**, 549–551.
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Zaitsev Y, Mamaev V (1997) *Marine Biological Diversity in the Black Sea: A Study of Change and Decline*. United Nations Publications, New York.
- Zubakov VA (1988) Climatostratigraphic scheme of the Black Sea Pleistocene and its correlation with the oxygen-isotope scale and glacial events. *Quaternary Research*, **29**, 1–24.

This study forms part of J.E. Brown's Ph.D. dissertation project under the sponsorship of Professor C.A. Stepien. Stepien's research programme focuses on the genetic patterns of invasions by Ponto-Caspian species in the North American Great Lakes. Her Great Lakes Genetics Laboratory is also studying the invasions of dreissenid mussels and Eurasian ruffe, as well as the systematics of the Neogobiinae in the Ponto-Caspian region. Other laboratory research focuses on the conservation genetics of native fishes. Further information and updates are available at <http://lakeerie.utoledo.edu>.

Supplementary material

The following supplementary material is available for this article:

Figure S1 Statistical parsimony network with nesting groups from nested clade analysis. Results of this analysis are given in Table 6. Significant clades in the analysis are shaded.

This material is available as part of the online article from: <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-294X.2008.03777.x>
(This link will take you to the article abstract).

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