Molecular Ecology (2005) **14**, 1869–1879 doi: 10.1111/j.1365-294X.2005.02565.x

Invasion genetics of the Eurasian spiny waterflea: evidence for bottlenecks and gene flow using microsatellites

ROBERT I. COLAUTTI,* MARINA MANCA,† MARKKU VILJANEN,‡ HENK A. M. KETELAARS,§ HANSRUDOLF BÜRGI,¶ HUGH J. MACISAAC* and DANIEL D. HEATH*

Great Lakes Institute for Environmental Research, University of Windsor, 401 Sunset Ave., Windsor, ON N9B 3P4 Canada, †CNR — Istituto per lo Studio degli Ecosistemi, Largo Tonolli 50, 28922 Verbania, Italy, ‡Karelian Institute, Department of Ecology, University of Joensuu, PO Box 111, FIN 80101 Joensuu, Finland, §Evides Water Company, PO BOX 59999, NL-3008 RA Rotterdam, The Netherlands, ¶Swiss Federal Institute for Water Resources and Water Pollution Control, Ch-8600 Dübendorf, Switzerland

Abstract

The Eurasian spiny waterflea (*Bythotrephes longimanus*) is a predacious zooplankter that has increased its range in Europe and is rapidly invading inland water-bodies throughout North America's Great Lakes region. To examine the genetics of these invasions, we isolated five microsatellite DNA loci with between 5 and 19 alleles per locus. We sampled three populations where *B. longimanus* has been historically present (Switzerland, Italy, and Finland) as well as an introduced European population (the Netherlands) and three North American populations (Lakes Erie, Superior, Shebandowan). Consistent with a bottleneck during colonization (i.e. founder effect), average heterozygosities of the four European populations ranged from 0.310 to 0.599, and were higher than that of three North American populations (0.151–0.220). Pairwise F_{ST} estimates among North American populations (0.002–0.063) were not significantly different from zero and were much lower than among European populations (0.208–0.474). This is consistent with a scenario of high gene flow among North American populations relative to that of European ones. Contrary to an invasion bottleneck, however, Erie and Superior populations contained similar numbers of rare alleles as European populations. Assignment tests identified several migrant genotypes in all introduced populations (the Netherlands, Erie, Superior, Shebandowan), but rarely in native ones (Switzerland, Italy and Finland). A large number of genotypes from North America were assigned to our Italian population suggesting a second, previously unidentified, invasion source somewhere in the region of northern Italy. Together, our results support an invasion bottleneck for North American populations that has been largely offset by gene flow from multiple native sources, as well as gene flow among introduced populations.

Keywords: *Bythotrephes longimanus*, genetic diversity, invasion genetics, invasive, microsatellite DNA, nonindigenous species

Received 3 December 2004; revision received 15 February 2005; accepted 1 March 2005

Introduction

Nonindigenous species are a major concern to resource managers and conservationists owing to their potential for profound ecological and economic impacts (e.g. Mack *et al*. 2000). The application of genetic techniques to investigate biological invasions (i.e. invasion genetics) is potentially useful for the management of nonindigenous species. For example, management efforts may benefit from the genetic identification of *vectors* and *routes* (*sensu* Ruiz & Carlton 2003) by which species are introduced (e.g. Cristescu *et al*. 2001; Novak & Mack 2001; Hänfling *et al*. 2002), or by the use of DNA bar coding to identify potentially invasive species (Hebert *et al*. 2003). Biological invasions by nonindigenous species also represent 'natural' colonization experiments, with effects on population genetics that are

Correspondence: Robert I. Colautti, Botany Department, University of Toronto, Toronto, ON M5S 3B2; Fax: (416) 978–5878; E-mail: colautti@botany.utoronto.ca

useful for testing and refining ecological and evolutionary theory. In particular, invasion genetics can provide insight into the early stages of population differentiation and evolution (Villablanca *et al*. 1998; Lee 2002; Hänfling & Kollmann 2003; Estoup *et al*. 2004; Kolbe *et al*. 2004). Because human-mediated invasions are often characterized by recently founded, demographically expanding populations with high dispersal, they are more likely than native populations to be out of mutation/migration–drift equilibrium (Eckert *et al*. 1996). Investigations of the effect of these nonequilibrium processes on population genetics and evolution may provide important insight into the evolution of nonindigenous species and of other populations with similar nonequilibrium dynamics (e.g. coral reef species, stocked and reintroduced species, biological control species).

To better understand the evolutionary genetics of invasive species it will be necessary (i) to identify the levels of genetic diversity and population differentiation among both introduced and native populations; (ii) to identify the most likely genetic sources for introduced populations; (iii) to compare patterns of genetic similarity with putative vectors of introduction and pathways of spread; and (iv) to examine the relationship between genetic diversity and the successful establishment, spread and proliferation of invaders. According to drift/migration models such as isolation by distance, and stepping stone, colonization events are predicted to result in reduced heterozygosity, a loss of rare alleles, and an increase in population differentiation owing primarily to founder effects and genetic drift (Wright 1943; Kimura & Weiss 1964; Nei *et al*. 1975; Slatkin 1993; Cornuet & Luikart 1996). However, human transport vectors (e.g. contaminants of imported goods, escape from cultivation, deliberate human introductions) may act to promote gene flow among introduced populations, and between introduced and native populations, thereby inflating genetic diversity in introduced populations (Kolbe *et al*. 2004). These opposing scenarios have important implications for the adaptation, and therefore the successful establishment, spread, and proliferation of nonindigenous species. Here we examine the population genetics of a rapidly spreading, abundant nonindigenous zooplankter, using populations from both its native and introduced range.

The predatory onychopod *Bythotrephes longimanus* is a widespread Palaearctic species with a facultative parthenogenetic life cycle and an invasion history in western Europe and North America (Ketelaars & Gille 1994; MacIsaac *et al*. 2000; Therriault *et al*. 2002). It was probably established in North America's Great Lakes around 1982 (Johannsson *et al*. 1991), and has since spread to over 50 inland lakes throughout Ontario, Canada (Therriault *et al*. 2002; MacIsaac *et al*. 2004; Muirhead & MacIsaac 2005). A morphologically distinct 'cederstroemi' form common to Europe is easily identified by a kink in the caudal process (Berg *et al*. 2002; Therriault *et al*. 2002). To date the cederstroemi form of *B. longimanus* is the dominant form in North America, although some animals lacking caudal process kinks have been found (R. Colautti, personal observation). North American *B. longimanus* invasions can have dramatic effects on aquatic food webs, owing primarily to profound changes in the morphology, abundance, and composition of native plankton species (Yan *et al*. 2002). The decline and later recovery of *Bythotrephes* in the native population of Lago Maggiore, Italy, had similar ecological consequences (Manca & Ruggiu 1998). Considering the ecological importance of this species, a comprehensive understanding of its population genetics in native and introduced regions has important implications for control effort.

Earlier allozyme and mtDNA studies on *B. longimanus* identified low levels of genetic variation (Weider 1991; Berg & Garton 1994; Berg *et al*. 2002; Therriault *et al*. 2002). Therefore, higher resolution markers are required to assess patterns of gene flow and the identification of potential reproductive barriers among *B. longimanus* populations. Here, we develop high-resolution, microsatellite DNA markers and use them to examine patterns of genetic diversity and differentiation among European and North American populations of *B. longimanus*. Specifically we ask: (i) Is there evidence for founder effects or bottlenecks in the introduced populations? (ii) What are the patterns of genetic diversity within vs. among populations in both Europe and North America? (iii) Is there evidence for multiple invasion sources or multiple introductions through human-mediated processes?

Materials and methods

Sample collection

Live samples were collected from three introduced populations in North America, as well as three native and one introduced population in Europe (Fig. 1). North American samples were collected from the inland Lake Shebandowan (SHE: 48°N, 90°W), and the Great Lakes Erie (ERI: 42°N, 83°N) and Superior (SUP: 46°N, 92°W) in June (Lakes Shebandowan and Superior) or August (Lake Erie) of 2001 by vertical hauls of a 250 μm mesh plankton net. *Bythotrephes longimanus* was found in all of the Great Lakes by 1989, but was not identified in Lake Shebandowan until 1995. In Europe, samples were collected from Petrusplaat reservoir, the Netherlands (NET 51°N, 5°E), Lake Puruvesi, Finland (FIN: 61°N, 28°E), Lago Maggiore, Italy (ITA: 46°N, 9°E), and Lake Lucerne, Switzerland (SWI: 47°N, 8°E) between August 2000 and November 2003 using plankton nets with either 126 μm, 250 μm or 450 μm mesh. *Bythotrephes longimanus* has been historically observed in all European lakes except for the Petrusplaat reservoir

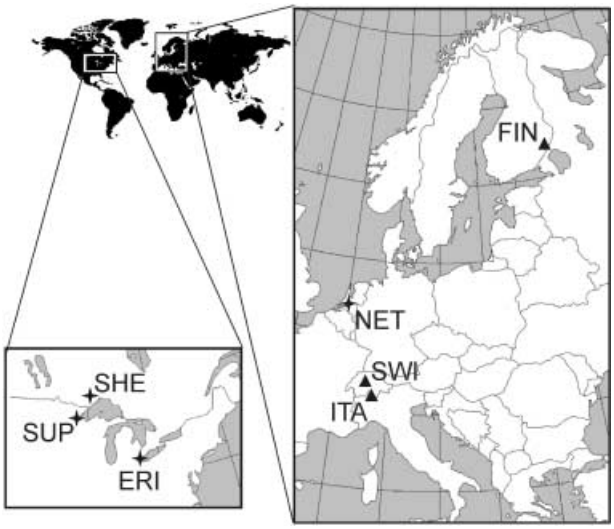

Fig. 1 Map of *Bythotrephes longimanus* sampling sites of native (triangles) and introduced (stars) populations in Europe and North America. European sites are Petrusplaat reservoir, the Netherlands (NET); Lake Puruvesi, Finland (FIN); Lago Maggiore, Italy (ITA); and Lake Lucerne, Switzerland (SWI). Note that the Netherlands population is a recently colonized population in Europe. North American sites are composed of the inland Lake Shebandowan (SHE), and the Great Lakes Erie (ERI) and Superior (SUP).

(NET), which was first invaded some time around 1987. Although our Swiss and Italian populations are in close geographical proximity, they are separated by the Alps, a major geographical barrier. All samples were stored in 95% ethanol and sorted in the laboratory under dissecting microscope to separate individual *B. longimanus* from other planktonic species. Over a period of days, pigment leached out of the samples resulting in a yellowish preservative, so ethanol was refreshed periodically (typically 30 days) until no colouration was visible.

Microsatellite isolation and characterization

Total genomic DNA was extracted from whole *B. longimanus* organisms using Promega's Wizard Kit®. Microsatellite DNA-enriched genomic libraries were developed following Hamilton *et al*. (1999), with modifications from O'Reilly *et al*. (2000). Briefly, high-quality DNA was extracted from eight individuals from Shebandowan Lake, Ontario, Canada, combined, and precipitated using a standard salt precipitation procedure (Sambrook *et al*. 1989). We conducted separate enrichment procedures using $(GACA)_4$, $(AGAT)_4$, $(AC)_8$, and $(AGC)_4$ probes. The enriched genomic DNA library was transformed into One Shot® *Escherichia coli* cells using a TOPO TA Cloning® Kit (Invitrogen) or a Promega® pGEM-T Vector System I with Promega's competent cells. A total of 204 transformed colonies were subcloned, and the insert polymerase chain reaction (PCR) was amplified using universal M13 primers. Twenty-three clones amplified more than one band and were discarded, as were 37 additional clones with inserts < 150 bp. Sequencing of the 144 remaining clones was performed with a Beckman Coulter CEQ 8000 Genetic Analysis System, according to protocols provided by Beckman Coulter. We selected 24 clones that contained more than three copies of the repeat sequence, and sufficient flanking regions to design primer sets (PRIMER3: Whitehead Institute for Biomedical Research and NetPrimer–Premier Biosoft®).

Primer sets were tested using standard PCR conditions: 94 °C for 2 min, followed by 35 cycles of 94 °C for 1 min, optimal melting temperature (as determined by a previous temperature-gradient PCR; see Table 1) for 1 min, and 72 °C for 90 s. These 35 cycles were followed by a final 5-min step at 72 °C. Extended primers were designed for polymorphic primers by adding the complimentary M13 F sequence to the 5′ end of the forward primer in each set. This enabled us to use dye-labelled M13 F primers (25 cycles × 58 °C) for fragment analysis on the CEQ 8000. In some cases, secondary amplification products were identified from a single sample (e.g. more than two fragments per locus); however, when this occurred the primary amplified fragments were identified by their higher intensity and characteristic chromatogram pattern.

Statistical analyses

To test for population structure among all populations, we used the TFPGA software package (Tools for Population Genetic Analysis version 1.3) and GENEPOP version 3.4 (Raymond & Rousset 1995) to calculate observed and expected heterozygosities (H_O and H_E), to test for Hardy–Weinberg equilibrium (HWE), to test for linkage disequilibrium among all pairwise sets of loci, and to calculate pairwise F_{ST}, and Nei's standard (1972) genetic distances (D_s). Linkage disequilibrium and HWE were calculated using a Markov chain Monte Carlo (MCMC) approach, with a dememorization number of 1000 and 100 batches of 1000 iterations per batch. A genetic distance matrix (Cavalli-Sforza & Edwards' chord distance, D_c, 1967) between all pairs of populations was calculated using the program POPULATIONS, version 1.2.28 (O. Langella, Centre National de la Recherche Scientifique, Laboratoire Populations, Génétique et Evolution, Gif sur Yvette; http://www.cnrs-gif.fr/pge/bioinfo/populations) using the neighbour-joining (NJ) cluster method. Bootstrap values were calculated over all loci and NJ data were exported to TREEVIEW (release 1.6.6; page 2001) for graphing. To test for recent gene flow among populations, we used GENECLASS (version 1.0; Cornuet *et al*. 1999) to test for migrant genotypes. We identified individuals that were cross-assigned to populations from which they were not sampled, using a conservative probability of 0.5 for

Table 1 Primer sequences for five novel polymorphic microsatellite DNA marker loci in *Bythotrephes longimanus*. The locus name, number of individuals genotyped (*n*), repeat array sequence, annealing temperatures (T_a), fragment size range, and number of alleles (*A*) are given for each locus

Locus	*n*	Repeat array	Primer sequences (5′–3′)	T_a (°C)	Size (bp)	*A*
Blo4	216	$(GTCT)_3(N)_2(CTGT)_6(N)_{56}(TG)_7$	F: CGC ACG TCT GTC TGT CTT TC R: GGC GGA AAT ATG CAA ATC AT	54.6	176–290	19
Blo14	209	$(CA)_3(N)_{12}(CA)_3(N)_{26}(CAGA)_4CA(CAGA)_3$	F: GAA GGC CCG ACG GTT GA R: AGG ATG TGC TGT GAT TGT GC	54.6	163–221	9
Blo20	222	$GTCTGTTT(GTCT)_3$	F: TTC CAA ATT TCT TTC CGA GGT R: CCT TAG CTC AAT CAA TTG CT	54.6	180–254	13
Blo35	192	$(CAAA)_3(CAGA)_3CAAACAGA$	F: CTT CCC TTT AAA TAG CCC GAA T R: CGG TAA TTG GAC CGT GAA TG	59	109–153	5
Blo158	166	$(GTT)_4(GGT)_7(GCT)_4(GGT)_2(GCT)_4$	F: GCC AGG GCG TAG GAA TTA TG R: AGC ACC TTC ATC ACC AGC A	57.8	98–191	17

the assignment test. These individuals likely represent recent migrant genotypes, or individuals containing recently introgressed genomes.

We used four methods to test for bottlenecks. First, we used the program BOTTLENECK version 1.2 (Cornuet & Luikart 1996; Cornuet *et al*. 1999) utilizing the Wilcoxon test for heterozygote excess, under the stepwise-mutation model (SMM) as well as the two-phase model (TPM). We chose the Wilcoxon test because it is preferable when few loci are used (Cornuet & Luikart 1996; Cornuet *et al*. 1999). Second, the FSTAT program version 2.9 (Goudet 1995) was used to measure the total number of individual genotypes (N) and unique alleles (N_a) and to calculate allele diversity (k) after Weir & Cockerham (1984). While both N_a and k may be informative, the latter controls for variation in sample size and is therefore more appropriate for comparisons of allelic richness when sample size is uneven. Third, we calculated M ratios after Garza & Williamson (2001). The M ratio compares the number of alleles (N_a) to their size range, measured in repeat number (r). At equilibrium, N_a/r should approach 1 under the SMM, while persistent bottlenecks are expected to decrease N_a faster than r, resulting in a lower M ratio (Garza & Williamson 2001). In extreme cases, sustained bottlenecks combined with drift can result in fixation of alleles. This would give an M ratio of 1, leading to an artificially inflated M when averaged over loci. Similarly, a fixed locus that recently mutated would result in two alleles with a small r, and therefore an M ratio close to 1. To reduce bias from extreme bottlenecks like this, in each population we excluded loci with one allele, as well as loci with two alleles if the frequency of the second allele was < 0.05. Finally, we graphed the allele frequencies for each population based on methods in Luikart *et al*. (1998). Excluding alleles with a frequency of zero (i.e. empty size classes), bottlenecks will cause a decrease in rare alleles and a resultant shift in the mode of the allele frequency histogram.

Results

Primer evaluation

We developed five polymorphic microsatellite loci for *Bythotrephes longimanus*, each with imperfect repeat motifs (Table 1). Two loci (*Blo4* and *Blo14*) contained both a di- and tetra-repeat motif, two others (*Blo20* and *Blo35*) contained mixed tetra-repeats, and the fifth (*Blo158*) was composed of a mixed tri-repeat motif (GenBank Accession nos AY819750–AY819754). When data from all populations were pooled, the number of alleles at each locus ranged from 5 to 19 (Table 1). Twelve of 35 population × locus combinations were significantly out of HWE after Bonferroni correction ($1 - [1 - \alpha]$), all of which were due to homozygote excess (Table 2). Perhaps surprisingly, the difference between the expected (H_E) and observed heterozygosities (H_O) was less than 3% of the H_O for all 12 loci not in HWE. In other words, the total numbers of expected and observed heterozygotes were similar, but the heterozygote genotypes often contained alleles that were not found in homozygotes and vice versa. A total of 10 pairwise linkage tests were possible within each population, resulting in 70 linkage comparisons. Of these, only three locus pairs showed evidence of linkage disequilibrium after Bonferroni correction, and all were combinations of the *Blo4*, *Blo35* and *Blo158* loci in the Italian (ITA) population. Twenty-five locus pairs could not be tested because of fixed alleles in one of the populations.

Genetic diversity and population divergence

Average H_O ranged from 0.3098 to 0.5993 in European populations and from 0.1507 to 0.2199 in North American populations (Table 2). The heterozygosity of the introduced Netherlands population was intermediate to native European and introduced North American populations.

Table 2 Summary information of the five loci for the introduced and native populations of *Bythotrephes longimanus* sampled. Number of individuals sampled (N), number of alleles (N_a), allelic diversity (k), and M ratio were calculated by FSTAT. Expected (H_E) and Nei's (1978) unbiased observed heterozygosities (H_O), were calculated by TFPGA. Introduced populations are in bold type. Locus–population combinations that were out of Hardy–Weinberg equilibrium after Bonferroni correction are denoted by bold type H_O with asterisks. Population abbreviations are as described in Fig. 1

	ITA	SWI	FIN	**NET**	**ERI**	**SUP**	**SHE**	Mean
Blo4								
N	48	24	22	39	36	24	23	30.9
N_a	11	9	3	2	4	1	2	4.6
k	7.21	8.60	2.77	1.97	3.06	1.00	1.94	3.79
M	0.216	0.375	0.273	0.400	0.190			0.291
H_E	0.7174	0.8559	0.5093	0.1420	0.1570	0.0000	0.0832	0.3521
H_O	**0.7250***	0.8741	0.5211	0.1439	0.1592	0.0000	0.0850	0.3583
Blo14								
N	48	23	20	38	37	24	19	29.9
N_a	4	3	2	4	3	1	1	2.6
k	2.29	2.48	1.85	2.34	2.42	1.00	1.00	1.91
M	0.333	0.130		0.500	0.136			0.275
H_E	0.0809	0.0841	0.0487	0.0769	0.1037	0.0000	0.0000	0.0563
H_O	**0.0818***	0.0860	0.0500	0.0779	**0.1051***	0.0000	0.0000	0.0573
Blo20								
N	47	24	22	45	37	23	24	31.7
N_a	5	1	2	6	5	7	1	3.9
k	4.94	1.00	2.00	4.32	4.34	6.52	1.00	3.44
M	0.250		0.667	0.162	0.143	0.212		0.287
H_E	0.7245	0.0000	0.4907	0.4328	0.3696	0.4707	0.0000	0.3555
H_O	**0.7323***	0.0000	0.5021	**0.4377***	**0.3747***	**0.4812***	0.0000	0.3611
Blo35								
N	46	24	21	39	24	18	20	27.4
N_a	4	1	3	4	3	1	2	2.6
k	3.75	1.00	2.81	3.12	2.63	1.00	2.00	2.33
M	0.174		0.600	0.190	0.600		0.095	0.332
H_E	0.6146	0.0000	0.1757	0.3849	0.1189	0.0000	0.2188	0.2161
H_O	0.6214	0.0000	0.1800	**0.3899***	0.1215	0.0000	0.2244	0.2196
Blo158								
N	39	24	18	21	23	24	17	23.7
N_a	10	5	4	5	3	7	3	5.3
k	8.37	4.34	3.94	4.62	2.99	6.31	3.00	4.79
M	0.323	0.105	0.108	0.169	0.207	0.318	0.120	0.193
H_E	0.8254	0.5764	0.6651	0.6224	0.3318	0.5096	0.4308	0.5659
H_O	**0.8362***	0.5887	**0.6841***	**0.6376***	0.3391	**0.5177***	0.4439	0.5782
Mean								
N	45.6	23.8	20.6	36.4	31.4	22.6	20.6	28.7
N_a	6.8	3.8	2.8	4.2	3.6	3.4	1.8	3.8
k	5.31	3.48	2.67	3.27	3.09	3.16	1.79	3.26
M	0.259	0.203	0.412	0.284	0.255	0.265	0.108	0.255
H_E	0.5926	0.3033	0.3779	0.3318	0.2162	0.1955	0.1465	0.3091
H_O	0.5993	0.3098	0.3875	0.3374	0.2199	0.1988	0.1506	0.3148

There was also considerable divergence evident among some populations. Pairwise F_{ST} estimates were highest between Swiss (SWI) and Netherlands (NET) populations, and generally were high among all European populations (0.208–0.598). In contrast, pairwise F_{ST} values between North American populations were more than two orders of magnitude smaller (0.002–0.063) and not significantly different from zero (Table 3). Pairwise F_{ST} between each North American and European population (0.287–0.571) were similar to, but generally higher than, measurements between pairs of native European populations (0.208–0.598; Table 3). The NET population was genetically more similar

Table 3 Matrix of pairwise F_{ST} (above diagonal), and Nei's (1972) standard genetic distance (D_s) (below diagonal) for four introduced and three native populations of *Bythotrephes longimanus*. Estimates between introduced populations are in bold type. All pairwise comparisons except those indicated by '†' were significantly different from zero based on Fisher exact test with Bonferroni correction. Population abbreviations are as described in Fig. 1

	ITA	SWI	FIN	**NET**	**ERI**	**SUP**	**SHE**
ITA	0	0.216	0.208	0.527	0.364	0.351	0.380
SWI	0.275	0	0.235	0.598	0.517	0.523	0.571
FIN	0.325	0.190	0	0.424	0.297	0.284	0.326
NET	0.313	0.474	0.373	0	**0.406**	**0.413**	**0.429**
ERI	0.542	0.482	0.196	**0.324**	0	**0.002**†	**0.039**†
SUP	0.518	0.478	0.183	**0.336**	**0.008**	0	**0.063**†
SHE	0.552	0.522	0.191	**0.320**	**0.015**	**0.021**	0

to North American than to European populations but relatively different from either, given pairwise F_{ST} values between 0.406 and 0.598. Pairwise estimates of Nei's (1972) standard genetic distances (D_s) showed a similar pattern, with values ranging from 0.190 to 0.598 for European pairs and 0.008 to 0.021 for North American pairs (Table 3). Neighbour-joining (NJ) analysis using chord distance (D_c) grouped the three introduced North American populations together, and away from all European populations including the NET. All nodes were supported ≥ 53% when bootstrapped over loci. All three of our measurements of genetic structure were similar; however, D_c showed relatively greater divergence among introduced populations compared to pairwise F_{ST} and D_s estimates (Fig. 2). In addition, the North American populations grouped slightly closer to the Netherlands (NET) population than to the Finland (FIN) population based on NJ analysis (Fig. 2).

Allele diversity and assignment tests

Average allelic diversity ranged from 1.79 in the introduced population of Shebandowan (SHE) to 5.31 in the native Italian (ITA). The two other introduced populations, Superior (SUP) and Erie (ERI), had values similar to native populations (Table 2). Comparison of mean *M* ratios reveals a similar pattern, ranging from 0.108 to 0.412. Shebandowan (SHE) had the lowest mean *M* ratio while ERI and SUP had values similar to native populations. The ITA population had the highest mean allelic diversity, while the FIN population had the highest mean *M* ratio. The distributions of the frequency of alleles for each population revealed that most alleles were rare (< 5%), but some populations had a high proportion of common alleles (> 45%); however, this distribution of allele frequencies was not consistent among populations (Fig. 3). None of the populations showed significant bottlenecks based on the Wilcoxon test under either the SMM or TPM.

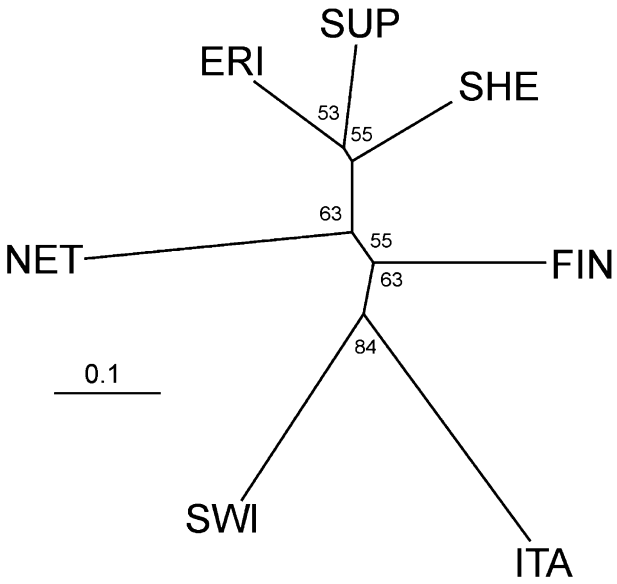

Fig. 2 Unrooted neighbour-joining cluster diagram based on Cavalli-Sforza & Edwards (1967) chord distance depicting the relationships among *Bythotrephes longimanus* populations from Europe (NET, FIN, ITA, SWI) and North America (SHE, ERI, SUP). Bootstrap values shown are percentage support based on 1000 bootstrap iterations. See Fig. 1 for population locations.

Results from the assignment test were more conclusive. Individuals from native populations tended to be assigned to the population from which they were sampled. Only four individuals from all of the native European populations were not assigned to the populations from which they were sampled, while in the NET population, 19 individuals were cross-assigned to other populations (Table 4). Conversely, many individuals from North American populations were cross-assigned to populations from which they were not obtained (Table 4). For example, 10 SUP individuals were classified as Lake Superior genotypes, while 11 and 10 individuals were assigned to ERI and SHE, respectively. In addition to cross-classification among introduced populations, between 6 and 25 individuals from ERI were cross-assigned to each of the native populations (Table 4).

Discussion

Evidence for bottlenecks

At first glance, our results suggest contradictory evidence for bottlenecks in the North American introduced populations. Consistent with an invasion bottleneck, heterozygosity (H_O) was much lower in North American than in European populations (Table 2). The introduced population of Shebandowan Lake (SHE), had the lowest values of H_O, *k* and *M* of all populations examined in this study, strongly

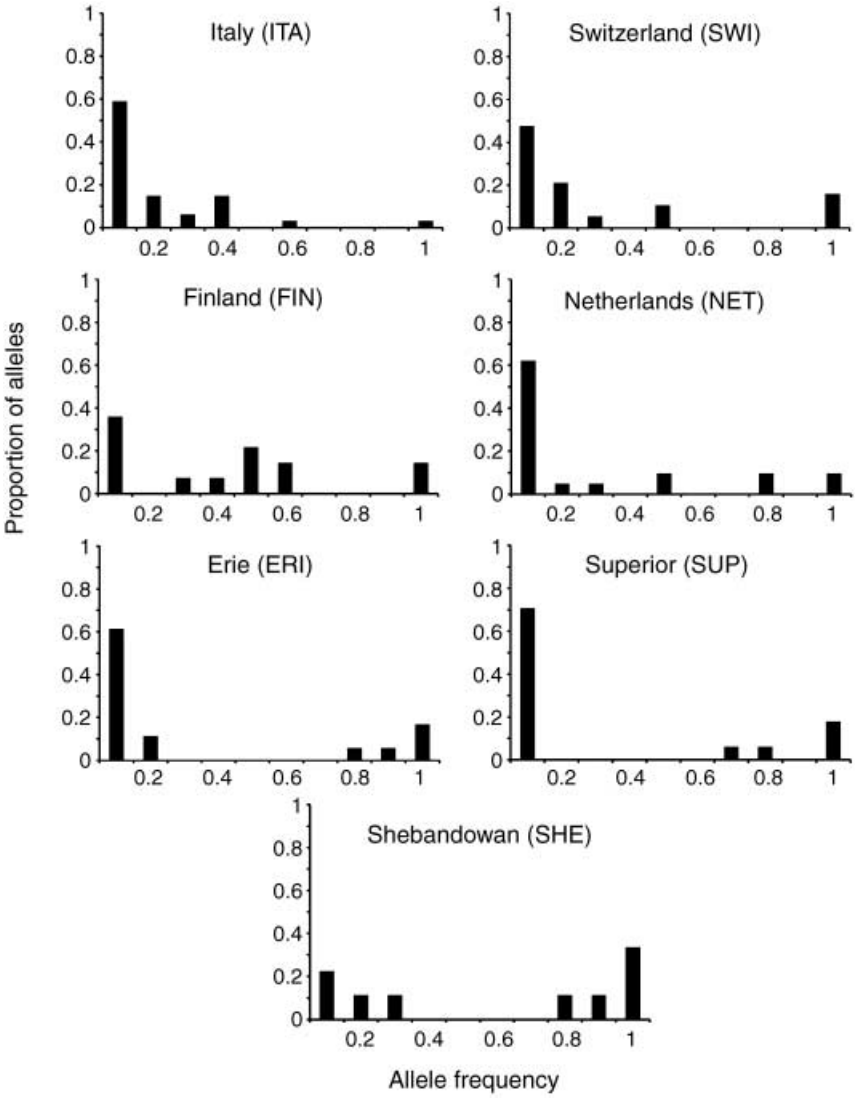

Fig. 3 Distribution histograms of allele frequencies over all loci, for each of four European (NET, FIN, ITA, SWI) and three North American (SHE, ERI, SUP) populations. Bars on the right of the histogram represent fixed alleles (i.e. frequency of 1), while rare alleles are on the left. Bottlenecks are expected to shift the mode from low to high frequencies (i.e. from left to right). Sampling locations are shown in Fig. 1.

supporting founder effects resulting from a bottleneck for this population – the only inland lake in the introduced region that we examined. However, allelic diversity (k), M ratios (Table 2), and allelic frequency diagrams (Fig. 3) showed weak evidence for bottlenecks in the other introduced populations when compared to native ones. These measurements, along with the bottleneck procedure, appear to have been affected by the presence of a large number of rare alleles in all populations except SHE (Fig. 3).

The genetic similarity (based on F_{ST}, D_s, D_c and assignment test results) between SUP and ERI suggests high levels of gene flow among these populations. This is consistent with the putative invasion vectors for *Bythotrephes longimanus* in North America. The two Great Lakes populations (SUP and ERI) are linked by transoceanic vessels and by regional fleets that move large amounts of ballast water, and associated biota, throughout the Great Lakes. Inland lakes like Shebandowan likely represent rare (or a single) secondary invasions mediated by recreational boaters and involve much smaller inocula (MacIsaac *et al*. 2004). Our results are most consistent with a scenario of initial colonization (i.e. an invasion bottleneck) followed by multiple subsequent invasions, along with high migration among Lakes Erie and Superior. The result is an invasion bottleneck that has been largely offset by gene flow from at least two, and probably more, native populations. An alternative scenario of a single large introduction event from a highly diverse pool cannot be completely ruled out. However, such a scenario is hard to harmonize with our results and those of previous studies. This would require that the source population contained a large number of rare alleles that are shared with our NET, FIN and ITA populations. Given the level of divergence of these populations, and the phylogeographical structure evident from our four European samples, it seems unlikely to us that such a population exists, unless it too is a recently colonized population.

Source	Assigned to							Unclass.
	ITA	SWI	FIN	**NET**	**ERI**	**SUP**	**SHE**	
ITA	30/30	1/1	0/0	0/0	0/0	0/0	0/0	17
SWI	0/0	10/10	0/0	0/0	0/0	0/0	0/0	14
FIN	3/0	0/0	15/12	0/0	0/0	0/0	0/0	7
NET	3/1	0/0	6/0	**35/29**	**6/0**	**7/0**	**6/0**	10
ERI	25/2	6/0	24/0	**9/0**	**39/2**	**35/0**	**28/0**	
SUP	1/0	0/0	2/0	**0/0**	**11/1**	**10/0**	**10/0**	12
SHE	3/0	0/0	12/	**1/0**	**16/0**	**16/0**	**17/1**	6

Table 4 Results of an assignment test on *Bythotrephes longimanus* genotypes using GENECLASS software. The number of times an individual was assigned to one of the populations is shown on the left side of the slash, and the numbers of unique assignments are shown on the right. Bold numbers in shaded area show assignment results within and among introduced populations. Individuals that failed to meet the simulation minimum criterion (threshold = 0.50) are listed as unclassified. Population abbreviations are as described in Fig. 1

A potentially confounding factor in our analysis is the facultative parthenogenetic life cycle of *B. longimanus*. Indeed, the *M* ratios of even native populations in this study are much lower than those of bottlenecked mammals in Garza & Williamson (2001), while the allele frequency histograms in Fig. 3 are quite different than those expected under SMM or TPM (see Luikart *et al.* 1998). Frequent cycles of asexual reproduction, and seasonal 'crashes' in population size could account for the number of fixed loci and the slight homozygote excess in loci out of HWE. Parthenogenesis reduces the rate of recombination, a result consistent with our uncommon HWE results. That is, seasonal bouts of sexual reproduction allow for recombination, leading to similar H_O and H_E results, but several generations of asexual reproduction increases the frequency of particular genotypes, resulting in a deviation from HWE at the level of genotype. The parthenogenetic life history of *B. longimanus* may therefore affect our ability to clearly identify bottlenecks in the introduced range. However, the lack of any clear patterns of linkage disequilibrium in any of our populations suggests that any confounding effects of parthenogenesis are small. Moreover, it is difficult to explain the existence of a large number of rare alleles in a scenario of low recombination.

Genetic variation and invasion sources

Based on NJ analysis and on pairwise F_{ST} and D_s estimates, our results support strong isolation among European populations of *B. longimanus*. Inferred gene flow was substantially higher among the introduced populations, although they also displayed some level of genetic differentiation based on the presence of unique alleles. Interestingly, the two Great Lakes populations studied—ERI and SUP—were genetically more similar to each other than either is to SHE, an inland lake that was likely invaded from Lake Superior (MacIsaac *et al.* 2004). Neighbour-joining analysis grouped all North American populations somewhere between the NET sample and the FIN sample, while Therriault *et al.* (2002) concluded based on sequence data of the mtDNA cytochrome oxidase I (*COI*) gene that Lake Puruvesi (our FIN population) was a more likely source than a population from the Netherlands.

With more extensive sampling and with more loci, our results might tend towards a more accurate relationship among these populations. However, if our results accurately portray the genetic structure of *B. longimanus* populations, then migration inferences based on averaged estimates like D_c or F_{ST} may not be equivalent to gene flow under many scenarios, such as when migrants are subject to selection at loci linked to neutral markers. Conversely, both strong invasion bottlenecks and significant gene flow from multiple sources could account for the patterns observed in our analyses. For example if SUP and ERI received genes from divergent areas of Europe (such as Finland and Italy), the high gene flow suggested by our analyses would have homogenized most of this variation so that each introduced population would group together and separate from either source.

Previous studies have identified only a single source region somewhere in the northeast Baltic region. Berg *et al.* (2002) suggested Lake Ladoga, Russia, was the source based on allozymes, and Therriault *et al.* (2002) identified Lake Puruvesi, Finland, as the likely source of the North American invasion based on the mtDNA *COI* gene. Lake Ladoga was not sampled in the latter study, but both support a single invasion source somewhere in the northeast Baltic. Our pairwise F_{ST} calculations, which implicate our FIN population over the others we sampled, also support a source in this region. By contrast, our NJ analysis suggests sources near Finland and the Netherlands, the latter of which is itself an introduced population.

Results from our assignment test are more consistent with gene flow from multiple sources in the introduced populations and implicate an additional source in the northern Mediterranean region. The cross-assignment of North American individuals to European populations suggests gene flow primarily from somewhere near FIN and ITA, with less input from areas near our NET and SWI populations. It is therefore somewhat surprising that the NET population groups closest to North America in our NJ analysis, while assignment results suggest FIN and ITA as

more likely sources of many recent migrants. It is also interesting to contrast specific alleles from the five loci. Of the 63 total alleles, 18 are shared between North American and European populations, while 10 and 35 alleles are unique to each region, respectively. Given that only four populations were sampled in the native range, the presence of unique alleles strongly implicates gene flow from other sources that were not sampled in the present study. In sum, our results combined with those of previous studies (Berg & Garton 1994; Berg *et al*. 2002; Therriault *et al*. 2002) support both a founder source from somewhere near the northeast Baltic region and subsequent gene flow from one or more sources, most likely in the region of northern Italy/Switzerland and perhaps the Netherlands. Alternatively, the similarity between the Netherlands and North American population could owe to their both being recently established populations from one or more similar sources. This pattern of gene flow is consistent with major areas of shipping linking the Great Lakes with North Sea and Baltic Sea countries (Colautti *et al*. 2003). We should stress that the European sources implicated in this study do not represent putative source *populations*, but rather the general, geographical regions from whence North American colonists were derived. The assignment of migrant individuals to both our most northern (Finland) and southern (Italy) populations strongly suggests either (i) multiple introductions from quite distant locations of Europe or (ii) the existence of one or more unsampled European sources that is linked to these regions, perhaps by commercial shipping or other human vectors.

Genetic diversity and invasion success

Genetic studies of Great Lakes invaders have reported two contrasting patterns of genetic diversity. Our microsatellite data, as well as previous data from allozymes and mtDNA strongly suggest bottlenecks for introduced relative to native populations of *B. longimanus* (Berg & Garton 1994; Berg *et al*. 2002; Therriault *et al*. 2002). The case is similar for populations of *Cercopagis pengoi*, the closest extant relative of *B. longimanus*, based on mtDNA (Cristescu *et al*. 2001). By contrast, mtDNA and allozyme data show similar levels of genetic diversity for both introduced and native populations of two other Great Lakes invaders, *Dreissena* mussels and *Neogobius melanostomus* gobies (Marsden *et al*. 1995; Dillon & Stepien 2001; Stepien *et al*. 2002). In the case of *Dreissena* sp., introduced populations were even more diverse at mtDNA loci than were several native populations (Stepien *et al*. 2002). Both *Dreissena* sp. and *N. melanostomus* were probably established in the Great Lakes several years after *B. longimanus*. All species represent rapidly spreading, prolific invaders in the Great Lakes, likely share similar introduction vectors and pathways, and likely receive gene flow from multiple populations, yet the genetic diversity available for selection in these species may be quite different.

Genetic bottlenecks are likely the result of circumstances underlying invasions (e.g. founder effects, strong selection), but a recent reduction in genetic variation revealed by neutral loci can also have important consequences for evolution. Extreme bottlenecks during invasion can reduce fitness by exposing recessive detrimental alleles and/or can reduce the amount of quantitative variation available for selection (Fisher 1930; Crnokrak & Roff 1999; Reed & Frankham 2001). However, parthenogenesis can effectively increase inbreeding rates by increasing the probability that clones or close relatives recombine to produce offspring. Thus, in general, species like *B. longimanus*, which can potentially establish new populations from a single individual, are likely more resistant to inbreeding depression than are obligately sexual organisms, since a long history of inbreeding will have purged recessive detrimental alleles by natural selection (Hedrick 1994; Lande *et al*. 1994). We found evidence for both bottlenecks and gene flow from multiple sources; thus it is not clear that the loss of genetic diversity at neutral loci is related to invasion success in this species. However, the combination of founder effects, selection, and ongoing gene flow from multiple sources may represent the 'best of both worlds': repeated bottlenecks break up epistatic interactions and expose rare, recessive genes to purging by natural selection while gene flow from multiple sources introduces new additive variation for selection to act upon.

Conclusion

Previous invasion genetics studies of *Bythotrephes longimanus* have been hampered by a dearth of variable markers; our five novel microsatellite loci provide a valuable tool for further research into the genetic structure of introduced populations in this species. We found that North American populations are significantly differentiated from European ones, and that the former contain only a subset of the genetic variability of the latter (i.e. a bottleneck). However, the genetic consequences of bottlenecks have been offset by high gene flow between the two Great Lakes populations and by gene flow from multiple European sources. Our novel finding of gene flow from southern Europe as well as the northeast Baltic suggests caution should be used when inferring genetic relationships among recently introduced populations. Further studies on the invasion genetics of introduced and native populations will lead to a better understanding of the effects of bottlenecks and gene flow from multiple sources on the evolution of populations significantly out of mutation/migration–drift equilibrium, with important implications for the evolution of species that are commensal with human activity. Ours is the first invasion genetics study of *B. longimanus* to use

rapidly evolving markers and one of a growing number on biological invaders in general (e.g. Berg & Garton 1994; Villablanca *et al*. 1998; Cristescu *et al*. 2001; Dillon & Stepien 2001; Novak & Mack 2001; Berg *et al*. 2002; Hänfling *et al*. 2002; Lee 2002; Stepien *et al*. 2002; Therriault *et al*. 2002; Estoup *et al*. 2004; Kolbe *et al*. 2004). We have successfully (i) identified the severity of founder effects, (ii) revealed source regions (northeast Baltic and southern Europe) of invasion and subsequent gene flow, and (iii) confirmed, at a qualitative level, putative scenarios of gene flow among introduced populations. Our findings demonstrate the utility of invasion genetics to inform management decisions (e.g. identifying invasion sources) and to improve our understanding of the genetic variation underlying the evolution of invasive species.

Acknowledgements

The authors are most grateful to V. Rixon, C. Busch, M. Docker, R. Hepburn, and T. Therriault for assistance in the laboratory. D. Gray, M. Gessner, G. Cimbura, M. Rawson, C. Willox and many others provided sampling assistance. S.C.H. Barrett, J. Friedman, S. Walker, and M. Burbidge, D. McCauley and two anonymous reviewers provided valuable feedback on the manuscript. Financial support to RIC from the Natural Science and Engineering Research Council of Canada (NSERC) and the University of Windsor, to HJM from NSERC and from a DFO Research Chair, and to DDH from NSERC Discovery and Canada Research Chair programs is gratefully acknowledged.

References

Berg DJ, Garton DW (1994) Genetic differentiation in North American and European populations of the cladoceran *Bythotrephes*. *Limnology and Oceanography*, **39**, 1503–1516.

Berg DJ, Garton DW, MacIsaac HJ, Panov VE, Telesh IV (2002) Changes in genetic structure of North American *Bythotrephes* populations following invasion from Lake Ladoga, Russia. *Freshwater Biology*, **47**, 275–282.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, **32**, 550–570.

Colautti RI, Niimi AJ, van Overdijk CDA, Mills EL, Holeck K, MacIsaac HJ (2003) Spatial and temporal analysis of transoceanic shipping vectors to the Great Lakes. In: *Invasive Species: Vectors and Management Strategies* (eds Carlton GM, Ruiz JT), pp. 459–504. Island Press, Washington, D.C.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989–2000.

Cristescu MEA, Hebert PDN, Witt JDS, MacIsaac HJ, Grigorovich IA (2001) An invasion history for *Cercopagis pengoi* based on mitochondrial gene sequences. *Limnology and Oceanography*, **46**, 224–229.

Crnokrak P, Roff DA (1999) Inbreeding depression in the wild. *Heredity*, **83**, 260–270.

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.

Eckert CG, Manicacci D, Barrett SCH (1996) Genetic drift and founder effect in native versus introduced populations of an invading plant, *Lythrum salicaria* (Lythraceae). *Evolution*, **50**, 1512–1519.

Estoup A, Beaumont M, Sennedot F, Moritz C, Cornuet JM (2004) Genetic analysis of complex demographic scenarios: spatially expanding populations of the cane toad, *Bufo marinus*. *Evolution*, **58**, 2021–2036.

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

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.

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

Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques*, **2**, 500–507.

Hänfling B, Carvalho GR, Brandl R (2002) mtDNA sequences and possible invasion pathways of the Chinese mitten crab. *Marine Ecology Progress Series*, **238**, 307–310.

Hänfling B, Kollmann J (2003) An evolutionary perspective of biological invasions. *Trends in Ecology & Evolution*, **17**, 545–546.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **270**, 313–321.

Hedrick PW (1994) Purging inbreeding depression and the probability of extinction—full-sib mating. *Heredity*, **73**, 363–372.

Johannsson OE, Mills EL, O'Gorman R (1991) Changes in the nearshore and offshore zooplankton communities in Lake Ontario: 1981–88. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**, 1546–1557.

Ketelaars HAM, Gille L (1994) Range extension of the predatory cladoceran *Bythotrephes longimanus* Leydig, 1960 (Crustacea, Onychopoda) in western Europe. *Netherlands Journal of Aquatic Ecology*, **28**, 175–180.

Kimura M, Weiss GH (1964) The stepping-stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49**, 561–576.

Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, Losos JB (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.

Lande R, Schemske DW, Schultz ST (1994) High inbreeding depression, selective interference among loci, and the threshold selfing rate for purging recessive lethal mutations. *Evolution*, **48**, 965–978.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, **17**, 386–391.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238–247.

MacIsaac HJ, Borbely JVM, Muirhead JR, Graniero PA (2004) Backcasting and forecasting biological invasions of inland lakes. *Ecological Applications*, **14**, 773–783.

MacIsaac HJ, Ketelaars HAM, Grigorovich IA, Ramcharan R, Yan

ND (2000) Modeling *Bythotrephes longimanus* invasions in the Great Lakes basin based on its European distribution. *Archiv für Hydrobiologie*, **149**, 1–21.

Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, Bazzaz FA (2000) Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications*, **10**, 689–710.

Manca M, Ruggiu D (1998) Consequences of pelagic food-web changes during a long-term lake oligotrophication process. *Limnology and Oceanography*, **43**, 1368–1373.

Marsden JE, Spidle A, May B (1995) Genetic similarity among zebra mussel populations within North America and Europe. *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 836–847.

Muirhead JR, MacIsaac HJ (2005) Evolution of lakes as hubs in an invasion network. *Journal of Applied Ecology*, **42**, 80–90.

Nei M (1972) Genetic distance between populations. *American Naturalist*, **106**, 283–292.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.

Novak SJ, Mack RN (2001) Tracing plant introduction and spread: genetic evidence from *Bromus tectorum* (Cheatgrass). *Bioscience*, **51**, 114–122.

O'Reilly PT, Canino MF, Bailey KM, Bentzen P (2000) Isolation of twenty low stutter di- and tetra nucleotide microsatellites for population analyses of walleye pollock and other gadoids. *Journal of Fish Biology*, **56**, 1074–1086.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software fore exact tests and ecumenicism. Journal of Heredity, 86, 248–249. Online: http://wbiomed.curtin.edu.au/genepop/index.html.

Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution*, **55**, 1095–1103.

Ruiz JT, Carlton GM (2003) Invasion vectors: a conceptual framework for management. In: *Invasive Species: Vectors and Management Strategies* (eds Carlton GM, Ruiz JT), pp. 459–504. Island Press, Washington, D.C.

Sambrook J, Fritsh EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

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

Stepien CA, Taylor CD, Debrowska KA (2002) Genetic variability and phylogeographical patterns of a nonindigenous species invasion: a comparison of exotic vs. native zebra and quagga mussel populations. *Journal of Evolutionary Biology*, **15**, 314–328.

Therriault TW, Grigorovich IA, Cristescu ME *et al.* (2002) Taxonomic resolution of the genus *Bythotrephes* Leydig using molecular markers and a re-evaluation of its global distribution. *Diversity and Distributions*, **8**, 67–84.

Villablanca FX, Roderick GK, Palumbi SR (1998) Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. *Molecular Ecology*, **7**, 547–560.

Weider LJ (1991) Allozymic variation in *Bythotrephes cederstroemi*: a recent invader of the Great Lakes. *Journal of Great Lakes Research*, **17**, 141–143.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

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

Yan ND, Girard R, Boudreau S (2002) An introduced invertebrate predator (*Bythotrephes*) reduces zooplankton species richness. *Ecology Letters*, **5**, 481–485.

This research was part of Rob Colautti's M.Sc. thesis work with Hugh MacIsaac and Dan Heath at the Great Lakes Institute, University of Windsor, Ontario, Canada. Rob is currently working on his Ph.D. thesis in the Botany Department of the University of Toronto, exploring patterns of evolution and adaptation of the invasive weed purple loosestrife (*Lythrum salicaria*) in eastern North America. Dan Heath holds a Canada Research Chair in Conservation Genetics, and Hugh MacIsaac holds an Invasive Species Research Chair from the Department of Fisheries and Oceans Canada.
