Heredity **85** (2000) 498–508 *Received 6 March 2000, accepted 22 August 2000*

Gene flow and genetic diversity: a comparison of freshwater bryozoan populations in Europe and North America

JOANNA R. FREELAND†, CHIARA ROMUALDI‡ & BETH OKAMURA*§

†*Department of Biological Sciences, Open University, Milton Keynes, Buckinghamshire MK7 6AA, U.K.,* ‡*Department of Statistics, University of Padova, Via S.Francesco 33, Padova, Italy,* §*School of Animal and Microbial Sciences, University of Reading, Whiteknights, PO Box 228, Reading, Berkshire RG6 6AJ, U.K.*

We have used microsatellite and mitochondrial sequence data to gain insight into patterns of gene flow and genetic diversity among North American and European populations of the freshwater bryozoan *Cristatella mucedo*. Mitochondrial sequence data reveal numerous, widely distributed, divergent genetic lineages in North America that can be broadly categorized into two groups, one of which is genetically homogenous and relatively similar to the European haplotypes, the other of which is more diverse. The maximum North American sequence differentiation translates into a divergence time of approximately 1.5 Myr BP. In contrast, European populations contained only three haplotypes that are all closely related. Microsatellite data reveal higher overall levels of genetic diversity in North America than Europe, although levels of within-population genetic variation are similar on the two continents. In North America, two of the three microsatellite loci show bimodal distributions of allele sizes which are significantly associated between the two loci. As a result, two microsatellite lineages are evident, and these are assortatively distributed between the mitochondrial haplotype groupings. The combined mitochondrial and microsatellite data suggest two distinct genetic lineages in North America that may represent cryptic species. Hybridization between the two presumptive species or subspecies may have contributed to the high levels of genetic diversity in North America. The overall lower levels of genetic diversity in Europe can be attributed to postglacial derivation of extant populations from a single mitochondrial lineage, and conformation to a metapopulation structure.

Keywords: freshwater bryozoans, gene flow, genetic diversity, microsatellites, mitochondrial haplotypes, phylogeography.

Freshwater invertebrate species distributed over large geographical distances often show a marked lack of morphological variation, despite the fact that many are incapable of active dispersal (reviewed in Frey, 1995). These seemingly invariant populations were regarded historically as providing evidence for frequent long-distance dispersal by a variety of vectors, particularly waterfowl (Darwin, 1859; Mayr, 1963). However, since ponds and lakes represent discrete habitats, their occupants may be expected to form subdivided populations in the absence of gene flow. This hypothesis may apply particularly to freshwater species that are incapable of active dispersal, and is supported by molecular data that reveal patterns of limited dispersal in many species. Cladoceran populations, for example, commonly show high levels of genetic differentiation despite morphological similarity across short geographical distances (e.g. Crease *et al*., 1990), although there are exceptions to this rule (e.g. Hann, 1995).

Despite frequent examples of pronounced genetic differentiation among neighbouring freshwater invertebrate populations, overall levels of mitochondrial sequence divergence are often low enough to suggest substantial historical gene flow within continents (Colbourne *et al*., 1998), although there are also examples of pronounced mitochondrial divergence at this scale (Taylor *et al*., 1998). Comparisons of populations from different continents similarly reveal variable patterns. Some freshwater zooplankton species exhibit

*Correspondence. E-mail: B.Okamura@reading.ac.uk

marked intercontinental genetic discontinuities (e.g. Taylor *et al*., 1996), and others show genetic similarities and evidence of recent intercontinental dispersal (e.g. Berg & Garton, 1994).

Dispersal is an integral part of population genetics and evolution, as the extent of gene flow influences a range of variables including population subdivision and genetic variation. Dispersal and gene flow are particularly important to populations that are prone to localized extinctions and recolonizations (Slatkin, 1977; Whitlock, 1992), and also to taxa that inhabit discrete sites such as lakes or ponds (Barrett *et al*., 1993; Berg & Garton, 1994). Nevertheless, our understanding of the effects of gene flow under different regimes remains largely theoretical because multiple factors can affect dispersal and colonization. A comparison of genetic patterns within and among conspecific populations with different histories should provide important insights into these complex processes. Here we present such a comparison by analysing the population genetics of the freshwater bryozoan *Cristatella mucedo* (Phylum Bryozoa, Class Phylactolaemata) in North America and Europe.

Colonies of *C. mucedo* commonly grow on roots, submerged branches, aquatic macrophytes, and other hard substrata in lakes and large ponds throughout the Holarctic region. Reproduction occurs largely through clonal processes, including budding, colony fission, and the production of highly resistant, dormant propagules called statoblasts, which are the only means of overwinter survival. Statoblast buoyancy allows for considerable movement within a site. In addition, statoblasts may be dispersed to more distant sites by animal vectors when they become entangled in feathers or fur, or are transported in digestive tracts. Freshwater bryozoans are hermaphroditic, and a limited phase of sexual reproduction may occur early in the growing season, at which time short-lived motile larvae are released. However, only a small proportion of colonies engage in sex (Okamura, 1997), and the consistent deviations from Hardy–Weinberg equilibrium in all *C. mucedo* populations investigated with codominant markers (Freeland *et al*., 2000a,b) are at least partially a result of this species' predominantly clonal nature. A more detailed discussion of the ecology and life history of *C. mucedo* can be found in Okamura (1997).

In this study we address the questions of whether or not there are different levels of genetic diversity in North American and European populations of *C. mucedo* and, if so, is this related to different patterns of dispersal and gene flow on the two continents? To answer these questions we analysed data from three microsatellite loci and partial 16S rDNA mitochondrial sequences. The information provided by these two molecular markers revealed significantly different patterns pertaining to the phylogeography and evolutionary history of *C. mucedo* in Europe and North America, and these patterns can be linked to biodiversity at different spatial scales.

Methods

Mitochondrial data

DNA was extracted from approximately 1 mm^3 of tissue in 500 μL of water with 5% Chelex (Bio-Rad Laboratories, Inc., California), following the manufacturer's protocol. We amplified a portion of 16S rDNA using primers 16SL (5′-CGCCTGTTTATCAAAAACAT-3′) and 16SH (5′-CCGGTCTGAACTCAGATCACGT-3′) (Palumbi *et al*., 1991). PCR reactions included 1 μL DNA, 1 × PCR buffer with 1.5 mM $MgCl_2$ (Promega), 200 μM dNTPs, 0.5 μM of each primer, and 1 U Taq DNA polymerase (Promega) in a 20-μL reaction. DNA amplification was performed in a Techne thermocycler with the following parameters: 1 cycle of 2 min at 95°C; 35 cycles of 30 s at 95°C, 30 s at 54°C, and 30 s at 72°C;1 cycle of 10 min at 72°C. The amplified products were purified with Promega Wizard kits, ligated into a Promega pGem-T Easy Vector, and transformed into XL1 MRF′ competent cells. One insert from each PCR reaction was sequenced along both strands using an ALFexpress DNA Sequencer with fluorescently labelled universal plasmid primers SP6 and M13F.

A total of 479–480 bp of 16S rDNA were sequenced from 34 North American colonies and 41 European colonies. The North American data included sequences of three different clones (identified by the microsatellite data, see below) from each of the populations shown in Table 1. In addition, one colony was sequenced from each of Golf Course Bay, Lake of the Woods, Ontario (49°46′ N, 94°27′ W), Mud Lake, WI (43°17′ N, 89°39′ W), Mirror Lake, WI (43°35′ N, 89°48′ W), Bolton Landing Pond, NY (43°36′ N, 73°40′ W), Morley Pond, MI (43°29′ N, 85°27′ W), Catfish Lake, WI (45°54′ N, 89°3′ W), Yellow Birch Lake, WI (45°56′ N, 89°14′ W), Adam Lake, Manitoba (49°3′ N, 100°3′ W), Whitewater Lake, WI (42°45′ N, 88°42′ W), and Washburn Lake, MO (46°51′ N, 93°59′ W). The European data included sequences of three different clones (identified by the microsatellite data, see below) from each of the populations shown in Table 1. In addition, one colony was sequenced from each of Tassjö, Sweden (56°20′ N, 13°0′ E), Hunn, Sweden (58°49′ N, 15°57′ E), Bussloo, Netherlands (52°12′ N, 6°7′ E), Hald Sø, Denmark (56°23′ N, 9°23′ E), and Etang de la Thévinère, France (47°10′ N, 1°6′ W). All 16S rDNA fragments were sequenced along both strands, and sequences have been deposited into GenBank (Accession numbers AF260067-AF260119).

Table 1 Location of populations (with numbers corresponding to population identities in Table 3), number of colonies collected, number of microsatellite genotypes identified, total number of alleles, proportion of novel clones (D^*), Shannon–Weaver diversity index (I), observed (H_o) and expected heterozygosity (H_e) with standard deviation (SD), and number of mitochondrial haplotypes (out of three colonies sequenced per location)

Location	Latitude and longitude	No. of colonies collected	No. of microsatellite genotypes	Total no.of alleles	D^*	I (SD)	H_o (SD)	H_e (SD)	No. of haplotypes
Europe									
Etang du Gros Caillou, France (1)	47°7′ N 2°5′ W	26	7	11	0.27	0.92 (0.80)	0.29 (0.51)	0.53 (0.05)	1
Le Lac de Grand Lieu, France (2)	47°3′ N 1°35′ W	30	16	7	0.53	0.78 (0.17)	0.07 (0.09)	0.48 (0.42)	2
Nationaal Park de Biesbosch, Netherlands (3)	51°40′ N 4°53′ E	28	20	15	0.71	1.22 (0.32)	0.32 (0.25)	0.65 (0.12)	2
Herpen Lake, Netherlands (4)	51°48′ N 5°47′ E	29	9	12	0.31	0.93 (0.29)	0.25 (0.44)	0.52 (0.15)	1
Ry Mølle Sø, Denmark (5)	56°5′ N 9°45′ E	29	8	9	0.28	0.63 (0.43)	0.22 (0.38)	0.39 (0.28)	2
Søndersø Nørresø, Denmark (6)	56°27′ N 9°23′ E	30	4	8	0.13	0.51 (0.68)	0.30 (0.52)	0.28 (0.37)	2
Rössjön, Sweden (7)	56°23′ N 13°5′ E	30	16	9	0.53	0.90 (0.31)	0.20 (0.34)	0.54 (0.16)	2
Sommen, Sweden (8)	58°0′ N 15°15′ E	30	16	11	0.53	1.03 (0.21)	0.09 (0.15)	0.60 (0.07)	1
Littoistenjarvi, Finland (9)	60°30′ N 22°19′ E	29	5	8	0.172	0.5 (0.47)	0.55 (0.51)	0.34 (0.30)	1
Enärjärvi, Finland (10)	60°28′ N 22°55′ E	29	24	16	0.83	1.2 (0.25)	0.14 (0.12)	0.64 (0.08)	1
Keitele, Finland (11)	62°53′ N 25°58′ E	28	17	17	0.61	1.19 (0.29)	0.13 (0.23)	0.59 (0.15)	1
Konnevesi, Finland (12)	62°40′ N 26°34′ E	30	18	13	0.60	1.17 (0.29)	0.30 (0.12)	0.64 (0.10)	2
North America									
Punderson Lake, OH (1)	41°27′ N 81°12′ W	30	23	17	0.77	1.25 (0.75)	0.08 (0.09)	0.58 (0.31)	3
Big Evans, MI (2)	43°44′ N 85°14′ W	28	15	18	0.54	1.22 (0.43)	0.39 (0.22)	0.61 (0.16)	2
Rock Lake, MO (3)	46°25′ N 94°29′ W	29	23	22	0.79	1.55 (0.43)	0.08 (0.08)	0.73 (0.13)	3
Lida Lake, MO (4)	46°32′ N 95°59′ W	29	21	17	0.72	1.23 (0.38)	0.33 (0.29)	0.63 (0.16)	3
North Star Lake, MO (5)	47°34′ N 93°39′ W	30	7	12	0.23	1.04 (0.46)	0.00 (0.00)	0.58 (0.19)	3
Thistledew Lake, MO (6)	47°48′ N 97°15′ W	30	23	8	0.77	0.75 (0.23)	0.18 (0.31)	0.48 (0.13)	2
Kewatin Bay, Lake of the Woods, Ontario (7)	49°47′ N 94°31′ W	29	22	18	0.76	1.37 (0.82)	0.07 (0.09)	0.69 (0.19)	2
Star Lake, Manitoba (8)	49°46′ N 95°15′ W	30	11	9	0.37	0.53 (0.92)	0.01 (0.02)	0.26 (0.44)	3

Sequences were aligned in CLUSTALX (Thompson *et al*., 1997). Distances (*d*) between them were calculated in MEGA v1.02 (Kumar *et al*., 1993) following the Kimura 2-parameter model of the number of nucleotide substitutions per site (Kimura, 1980). The aligned sequences were then imported into PHYLIP (Felsenstein, 1995) and a maximum likelihood tree was reconstructed using empirical base frequencies, a transition/transversion ratio of 2.0, and 1000 bootstrap replicates.

Microsatellite data

Two earlier studies on North American and European populations of *C. mucedo* utilized different suites of microsatellite loci and therefore did not allow direct comparisons between the two continents (Freeland *et al*., 2000a,b). For this paper, we analysed a subset of these data taken from the three microsatellite loci common to both studies (loci 1.1, 2.2, and 6.7; Freeland *et al*., 1999), and representing 12 populations from north-western continental Europe and eight populations from central North America (Table 1). These populations were chosen because they are located within areas lacking major disruptions such as mountain ranges. The European populations were separated by 20–2500 km, and the North American populations were separated by 50–1500 km.

Assessment of Hardy–Weinberg equilibrium, and calculations of expected and observed heterozygosity values (H_o and H_e) were performed in POPGEN (Yeh & Boyle, 1997). Genetic diversity was measured in four ways: (i) D^* (= number of multilocus genotypes/number of colonies); (ii) Expected heterozygosity (H_e); (iii) Number of alleles; and (iv) Shannon–Weaver diversity index (*I*; Shannon & Weaver, 1949). The within-population diversity measures from the two continents were compared using a *t*-test for the expected heterozygosity values and a Mann-Witney test for the *I*-values. The cumulative measures of genetic diversity for the two continents, based on a comparison of the total number of genotypes and the total number of clonal replicates, were compared using a χ^2 contingency table analysis. Kolmogorov–Smirnov tests were used to compare the shape of the allele frequency distributions irrespective of the absolute allele sizes. Allele sizes for each continent were standardized by assigning a value of one to the smallest allele, and consecutively numbering the remaining alleles by adding one for each additional two basepairs.

F_{ST} values were calculated in GENEPOP (Raymond & Rousset, 1995) and *Rho* values, unbiased estimators of Slatkin's R_{ST} (Slatkin, 1995), were calculated in RSTCALC (Goodman, 1997). Estimates of gene flow were calculated as $Nm = 1/4[(1/F_{ST}) - 1]$, and as $Nm = 1/4[(1/Rho)-1]$. In addition, nonparametric discriminant function analyses (Kolmogorov–Smirnov test for normality: $P < 0.01$ in all populations) were conducted to classify each colony to the population whose overall genotype it most closely resembled (Manly, 1994), thereby allowing an independent inference of connectivity among populations. This entailed calculating distances based on the absolute difference in number of repeat units between each pair of colonies for each locus, and then summing these distances over all loci. This value was then squared to obtain a more realistic representation of the number of mutational steps involved. A cross-validation method in which a single colony was removed from the dataset was employed for each calculation.

The relationship between geographical and genetic distances was investigated in GENEPOP using Mantel tests with 1000 permutations to compare pairwise values of both F_{ST} and *Rho* [as $F_{ST}/(1 - F_{ST})$ or $Rho/(1 - Rho)$] against the natural log of the shortest geographical distance between sites. F_{ST} (Wright, 1951) assumes a low mutation rate and either an infinite alleles model or a *k*-alleles model, whereas R_{ST} (Slatkin, 1995) assumes a rapid mutation rate and a stepwise mutation model. We used both measurements in the test for isolation by distance, as there is no consensus as to whether F_{ST} or R_{ST} (*Rho*) is more appropriate for analysing microsatellite data.

Results

Mitochondrial data

The distances among the 19 16S rDNA sequences are shown in Table 2. One indel meant that the total number of sequenced base pairs ranged from 479 bp (in 16 haplotypes) to 480 bp (in 3 haplotypes). There were 39 variable sites, of which 18 were phylogenetically informative. The ratio of transitions/transversions ranged from 0.0 to 7.0 in pairwise comparisons, with a mean ratio of 1.8. Sixteen haplotypes were found in the 34 North American colonies that were sequenced, with a maximum sequence divergence of 0.036. In contrast, only three haplotypes were found in the 41 European colonies that were sequenced, with a maximum sequence divergence of 0.006. The distances between North American and European sequences ranged from 0.008 to 0.026. Some North American sequences were more similar to the European sequences than to other North American haplotypes. Many North American populations contained multiple haplotypes, while European populations more commonly revealed a single haplotype (Table 1). Figure 1 shows the scaled relationships among haplotypes as depicted by a maximum likelihood tree.

Table 2 Kimura 2-parameter distances between mitochondrial haplotypes. Numbers in bold represent comparisons involving the three European haplotypes. Haplotype identity numbers correspond to colonies from the following populations (number after location refers to number of colonies with that haplotype):**1:** Kewatin Bay, Lake of the Woods, Ontario (2); Golf Course Bay, Lake of the Woods, Ontario (1); Washburn Lake, MO (1); Rock Lake, MO (1); Lida Lake, MO (1); Punderson Lake, OH (1); Big Evans Lake, MI (2); Morley Pond, MI (1); **2:** Mirror Lake, Wisconsin (1); **3:** Mud Lake, WI (1) **4:** Bolton Landing Pond, NY (1); **5:** Punderson Lake, OH (1); **6:** Rock Lake, MO (1); North Star Lake, MO (1); Kewatin Bay, Lake of the Woods, Ontario (1); Lida Lake, MO (1); Catfish Lake, WI (1); **7:** Star Lake, Manitoba (1); Whitewater Lake, WI (1); **8:** Thistledew Lake, MO (2); **9:** Yellow Birch Lake, WI (1); **10:** Adam Lake, Manitoba (1); **11:** Star Lake, Manitoba (1); Lida Lake, MO (1); **12:** Star Lake, Manitoba (1); **13:** North Star Lake, MO (1); **14:** Big Evans Lake, MI (1); Rock Lake, MO (1); **15:** Punderson Lake, OH (1); **16:** North Star Lake, MO (1); Thistledew Lake, MO (1); **17:** Etang de la Thévinère, France (1); Etang du Gros Caillou, France (3); Le Lac de Grand Lieu, France (2); Bussloo, Netherlands (1); Nationaal Park de Biesbosch, Netherlands (2); Herpen, Netherlands (3); Hald Sø, Denmark (1); Hunn, Sweden (1); Sommen, Sweden (3); Keitele, Finland (3); Enärjärvi, Finland (3); Littoistenjarvi, Finland (3); Søndersø Nørresø, Denmark (2); Rössjön, Sweden (2); Konnevesi, Finland (2); Ry Mølle Sø, Denmark (2); **18:** Le Lac de Grand Lieu, France (1); Ry Mølle Sø, Denmark (1); Søndersø Nørresø, Denmark (1); Tassjö, Sweden (1); Nationaal Park de Biesbosch, Netherlands (1); **19:** Rössjön, Sweden (1); Konnevesi, Finland (1)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
2	0.021																	
3	0.002	0.023																
4	0.013	0.026	0.015															
5	0.021	0.004	0.023	0.026														
6	0.023	0.032	0.026	0.011	0.032													
7	0.021	0.030	0.023	0.013	0.030	0.011												
8	0.015	0.015	0.017	0.028	0.015	0.034	0.032											
9	0.004	0.026	0.006	0.017	0.026	0.028	0.026	0.019										
10	0.002	0.023	0.004	0.011	0.023	0.021	0.019	0.017	0.006									
11	0.026	0.034	0.028	0.013	0.034	0.002	0.013	0.036	0.030	0.023								
12	0.002	0.023	0.004	0.015	0.023	0.026	0.023	0.017	0.006	0.004	0.028							
13	0.004	0.025	0.006	0.017	0.026	0.028	0.026	0.019	0.008	0.006	0.030	0.006						
14	0.002	0.023	0.004	0.015	0.023	0.026	0.023	0.017	0.006	0.004	0.028	0.004	0.006					
15	0.004	0.026	0.006	0.017	0.026	0.028	0.026	0.019	0.008	0.006	0.030	0.006	0.008	0.006				
16	0.004	0.026	0.006	0.017	0.026	0.028	0.026	0.019	0.008	0.006	0.030	0.006	0.008	0.006	0.008			
17	**0.008**	**0.017**	**0.011**	**0.013**	**0.017**	**0.019**	**0.017**	**0.019**	**0.013**	**0.011**	**0.021**	**0.011**	**0.013**	**0.011**	**0.013**	**0.013**		
18	**0.010**	**0.019**	**0.013**	**0.015**	**0.019**	**0.021**	**0.019**	**0.021**	**0.015**	**0.013**	**0.023**	**0.013**	**0.015**	**0.013**	**0.015**	**0.015**	**0.002**	
19	**0.013**	**0.021**	**0.015**	**0.017**	**0.021**	**0.023**	**0.021**	**0.023**	**0.017**	**0.015**	**0.026**	**0.015**	**0.017**	**0.015**	**0.017**	**0.017**	**0.004**	**0.006**

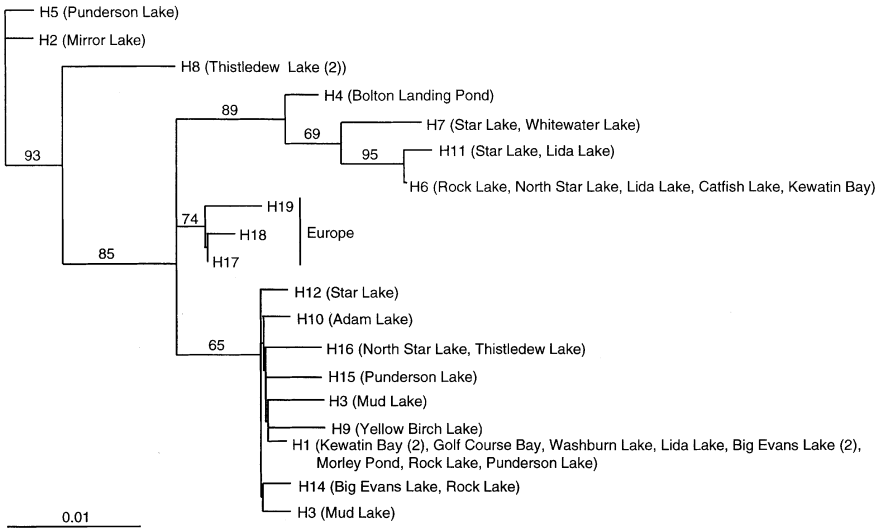

Fig. 1 Scaled maximum likelihood tree of the 19 mitochondrial haplotypes. Number in brackets after the population name refers to the number of individuals with that sequence (only specified when >1). Bootstrap values >50 are written above the branches.

Microsatellite data

The 348 European colonies yielded 160 multilocus genotypes and 188 replicates, which gives a total European D^* value of 0.46. The 235 North American colonies yielded 145 multilocus genotypes and 90 replicates, which gives a total North American D^* value of 0.62. The χ^2 test indicated that North America has a relatively higher number of unique genotypes than Europe ($\chi^2 = 13.281$, d.f. $= 1$, $P < 0.001$). This is reflected by the overall higher number of alleles in North America (61 alleles in 235 colonies) compared to Europe (43 alleles in 348 colonies; Fig. 2). While two of the three microsatellite loci had some alleles common to both North America and Europe, the distributions of allele frequencies were significantly different at all three loci (Kolmogorov–Smirnov: Locus 1.1: $P = 0.001$; Locus 2.2: $P = 0.000$; Locus 6.7: $P = 0.000$; Fig. 2). Loci 2.2 and 6.7 revealed bimodal patterns of distribution of allele sizes in North American populations (Fig. 2). At

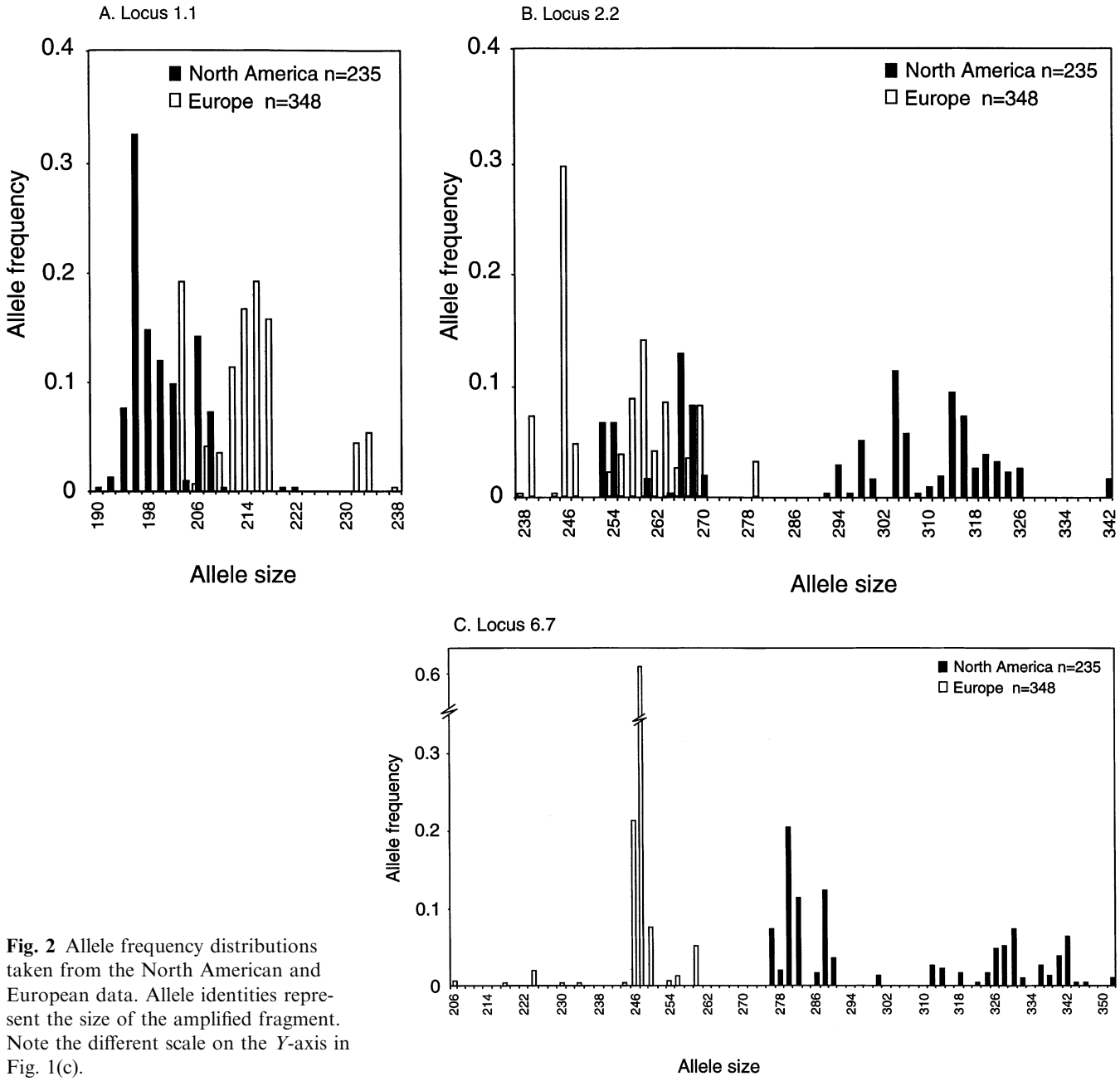

Fig. 2 Allele frequency distributions taken from the North American and European data. Allele identities represent the size of the amplified fragment. Note the different scale on the *Y*-axis in Fig. 1(c).

locus 2.2, group 1 alleles ranged in size from 252 to 270 bp, and group 2 alleles ranged from 292 to 342 bp. At locus 6.7, group 1 alleles ranged from 276 to 290 bp, and group 2 alleles ranged from 312 to 352 bp. Note that the 300 bp alleles at locus 6.7 cannot be assigned clearly to either group and, therefore, the four genotypes that contained this allele were excluded from further analysis. Of the remaining 141 North American microsatellite genotypes, 30 (21.3%) contained alleles from group 1 at locus 2.2 and group 2 at locus 6.7; 72 (51.1%) contained alleles from group 2 at locus 2.2 and group 1 at locus 6.7; 14 (9.9%) contained alleles from group 1 of each locus; 19 (13.5%) contained alleles from group 2 of each locus; and 6 (4.3%) were heterozygous for alleles from each group in at least one of the two loci. A contingency coefficient tested with a χ^2 analysis showed that there is an overall association between the allele groups of the two loci ($\chi^2 = 39.65$, d.f. $= 1$, $P < 0.001$).

In order to compare the allele groups at loci 2.2 and 6.7 with the mitochondrial haplotypes, the latter were divided into two groups based on the pairwise distances between sequences. The first group included 9 haplotypes with pairwise divergences of 0.002–0.008 (haplotypes 1,3,9,10,12–16; Table 2). Haplotypes within this group are separated from the European haplotypes by pairwise sequence divergences of 0.008–0.013, and are contained within the cluster that is closest to the European haplotypes in Fig. 1. The second mitochondrial group contains the 7 remaining North American haplotypes, and these have pairwise sequence divergences of 0.011–0.028 within the group, and 0.013–0.021 in comparison with the European haplotypes. While the division of these two groups is not based on stringent criteria, they can be described as a cohesive group with relatively high similarity to the European haplotypes, and a more diverse group that shows greater differentiation from the European haplotypes.

Seventeen individuals from haplotype group one were microsatellite genotyped. Of these, 13 had alleles from group 2 and group 1 of loci 2.2 and 6.7, respectively; 2 had alleles from group 1 and group 2 of loci 2.2 and 6.7, respectively; and 1 had alleles from group 2 of both loci. Ten individuals from haplotype group 2 were microsatellite genotyped, and of these 7 had alleles from group 1 and group 2 of loci 2.2 and 6.7, respectively, and 3 had alleles from group 2 and group 1 of loci 2.2 and 6.7, respectively. The proportion of individuals with group 1, locus 2.2, and group 2, locus 6.7, compared to the proportion of individuals with group 2, locus 2.2, and group 1, locus 6.7, is significantly different between the two haplotype groups (Fisher's exact test: $P = 0.009$). Therefore, two genetic lineages can be identified in North America based on the mitochondrial haplotypes and the alleles at loci 2.2 and 6.7. A small proportion of individuals does not agree with these lineage classifications, and these may be explained by hybridization between the two lineages (see Discussion).

Within-population values of genetic variation, measured as the number of alleles, D^*, I and H_e are given in Table 1. Comparisons of the within-population values of H_e (t-test: $t = -1.51$, d.f. $= 18$, $P = 0.148$) and I (Mann–Witney: $T = 95.5$, $P = 0.396$) show that levels of within-population diversity did not differ between continents (see also Fig. 3 for a graphic comparison of population H_e values). In addition, levels of observed heterozygosity in the European and North American populations (Table 1) were not significantly different from one another (t-test: $t = 1.54$, d.f. $= 18$, $P = 0.142$). None of the populations on either continent were found to be in Hardy–Weinberg equilibrium.

When *Nm* is derived from *Rho*, values are higher among European populations [range $= 0.11$–14.86, $n = 66$, mean $= 1.21 \pm 0.31$(SE)] compared to the North American populations (range $= 0.03$–0.87, $n = 28$, mean $= 0.20 \pm 0.04$). When *Nm* is derived from F_{ST} this difference is considerably reduced, although *Nm* remains higher for the European (range $= 0.11$–4.2, $n = 66$, mean $= 0.70 \pm 0.07$) than the North American populations (range $= 0.19$–0.73, $n = 28$, mean $= 0.44 \pm 0.03$). Higher levels of connectivity among European popu-

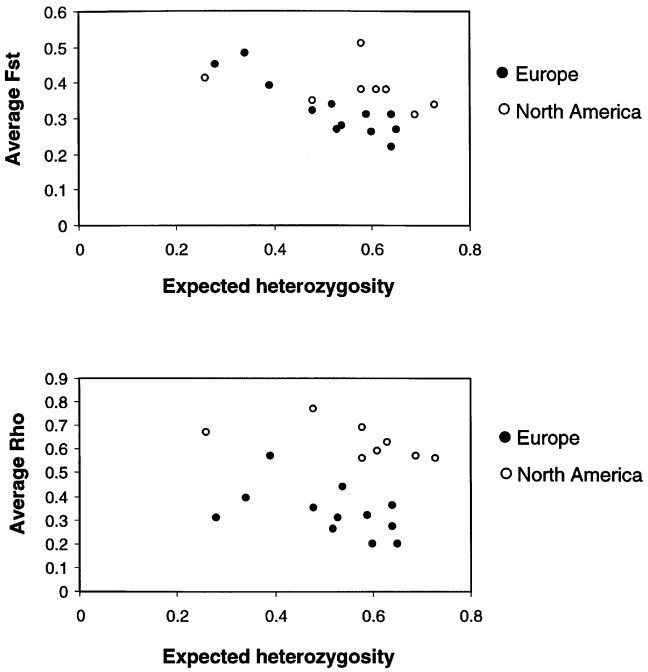

Fig. 3 Summary of within-population diversity (measured as H_E) and among population divergence (measured as average pairwise F_{ST} or *Rho* for each population) for all North American and European data. Note the different scales on the *Y*-axes.

Table 3 Classification of colonies according to the discriminant function analyses. Numbers in bold represent individuals assigned to a population other than the one in which they were found. When clonal replicates are included in this number, the number in brackets written immediately afterward refers to the number of unique multilocus genotypes. Rows correspond to the site from which the colony was collected, columns correspond to the sites to which the colony was classified. Column and row numbers correspond to the population numbers in Table 1

	1	2	3	4	5	6	7	8	9	10	11	12
(a) European populations												
1	24	0	0	0	0	0	**1**	0	0	0	0	**1**
2	0	11	0	**2**	0	**17(9)**	0	0	0	0	0	0
3	0	0	12	**3**	**1**	**3**	**5**	**2(1)**	0	**1**	**1**	0
4	**1**	**1**	0	25	0	**1**	**1**	0	0	0	0	0
5	0	**1**	0	0	25	**3(2)**	0	0	0	0	0	0
6	0	0	0	0	0	28	0	**2(1)**	0	0	0	0
7	**2**	0	0	0	0	**1**	24	**1**	**1**	0	0	**1**
8	**1**	**2**	**3**	**3**	**2(1)**	**5(2)**	**2**	15	0	0	0	0
9	0	0	0	0	0	0	0	**1**	28	0	0	0
10	**3**	**1**	**4**	**3**	0	0	0	**1**	0	18	**2(1)**	**1**
11	0	0	**2**	**3**	0	**2(1)**	**1**	**1**	0	0	20	**1**
12	**1**	**0(1)**	0	0	0	**4(3)**	0	**3(2)**	0	**4**	**1**	14
(b) North American populations												
1	28	0	**2**	0	0	0	0	0				
2	0	21	**2**	**5(3)**	0	0	0	0				
3	**2**	**1**	16	**10(9)**	0	0	0	0				
4	**1**	0	**4**	23	0	0	0	**1**				
5	0	0	0	0	25	0	0	**5(2)**				
6	0	0	0	0	0	30	0	0				
7	0	0	0	0	**4**	0	23	**2**				
8	0	0	0	0	**4(2)**	0	0	26				

lations are also suggested by the discriminant function analyses which showed that 32.5% of European colonies (representing 51.3% of European multilocus genotypes) were classified to populations other than the one from which they were collected, compared to 18.3% of North American colonies (representing 24% of North American multilocus genotypes; Table 3).

Figure 3 summarizes within-population diversity and among population differentiation in the two continents. This depiction reiterates the similar levels of within-population genetic variation (here shown as H_e) from two continents, and also represents the variable levels of population subdivision. According to pairwise F_{ST} values, the European and North American populations are separated by similar genetic distances, but according to *Rho* the North American populations show considerably higher levels of subdivision than the European populations. The Mantel tests showed no pattern of isolation by distance in either North America ($P = 0.62$ when F_{ST} values are used, $P = 0.94$ when *Rho* values are used) or Europe ($P = 0.80$ when F_{ST} values are used, $P = 0.77$ when *Rho* values are used).

Discussion

Gene flow in North America and Europe

In North America, multiple mitochondrial haplotypes have been maintained both within populations and within the continent as a whole. The maximum mitochondrial sequence divergence of 0.036 in North America translates into a divergence time of approximately 1.5 Myr BP, assuming a molecular clock similar to that of 2.3% per million years for arthropod mitochondrial DNA (Brower, 1994). Because much of North America was glaciated up to 10–20 000 yr BP, numerous haplotypes must have been maintained throughout glacial–interglacial cycles, possibly in multiple refugia. The presence of multiple haplotypes within the same populations indicates repeated colonization events.

Estimates of gene flow in North America based on microsatellite data are consistently low, but the reason for this is not readily apparent. The pronounced levels of genetic differentiation revealed by the microsatellite data must be reconciled with the lack of population differentiation shown by the mitochondrial data. This discrepancy may be attributed to the relatively rapid mutation rate of microsatellite loci that, in conjunction with little ongoing gene flow, could explain the relatively rapid acquisition of population-specific microsatellite alleles. However, it should be noted that the high mutation rate of microsatellite loci could generate a large number of different alleles over a relatively short time period. Our limited sample sizes mean that a number of these alleles would remain undetected, potentially leading us to underestimate the numbers of shared alleles among populations and, in turn, to

underestimate gene flow. Furthermore, we sampled from only a small proportion of the large number of sites that may support *C. mucedo* populations, and sampling from a greater number of lakes may have revealed patterns of connectivity among some locations.

In Europe, rather different patterns of population differentiation were evident. Only three mitochondrial haplotypes were detected, all of which were closely related. The relatively low number of haplotypes in Europe as compared to North America is particularly notable given the substantially larger sampling area in Europe. Because these European sequences are similar to some of the North American sequences, and fall within the range of divergences among all North American and European haplotypes, it appears possible that *C. mucedo* colonized Europe from North America. Intercontinental dispersal events between the two continents have been identified in other freshwater invertebrates (e.g. Berg & Garton, 1994; Taylor *et al.*, 1996). Alternatively, the current distribution of mitochondrial haplotypes may be explained if the European and North American populations represent two ends of an essentially continuous distribution of *C. mucedo* populations across Asia and the Bering Straits. The paucity of mitochondrial haplotypes in north-western Europe today may have resulted from a loss of diversity during glacial cycles, as opposed to a founder effect following colonization from North America.

Measures of population differentiation and subsequent inferences of gene flow are similar in the North American and European populations when F_{ST} values are calculated, but differ considerably according to calculations of R_{ST} (*Rho*). While R_{ST} may be more suitable for microsatellite data owing to its inherent SMM and allowance for a high mutation rate (Slatkin, 1995), measure of F_{ST} may be more appropriate at smaller geographical distances (Estoup *et al.*, 1995). Calculations of R_{ST} may be more appropriate for populations that have low migration rates, as these will show significant effects of mutation (Rayboud *et al.*, 1998). Estimates of population subdivision based on *Rho* and F_{ST} differ most strikingly with respect to the relatively high levels of differentiation among North American populations according to *Rho* values (Fig. 3). This may be attributed to the high variance in allele sizes in the North American populations, as R_{ST} is equivalent to the fraction of the total variance in allele size that is between populations (Slatkin, 1995; Goodman, 1997). A downward bias in estimates of population differentiation calculated from F_{ST} has been predicted for microsatellite data, as the IAM of F_{ST} does not allow for forward and backward mutation in allele size and will therefore underestimate the total divergence between populations (Slatkin, 1995). However, the complex patterns of microsatellite mutation and evolution (Primmer & Ellegren, 1998 and references therein) mean that the most appropriate method for analysing microsatellite data has not yet been agreed upon.

The relatively low levels of population subdivision in Europe revealed by F_{ST} and *Rho* suggest correspondingly high levels of connectivity among populations on that continent, a conclusion that is reinforced by the discriminant function analyses. This genetic similarity of populations may be partially attributed to the postglacial (re)colonization of lakes by a single genetic lineage. However, it may also be influenced by ongoing events, as the process of population differentiation will be slowed or even arrested in the presence of ongoing gene flow among sites. The European samples were from water bodies located along a transect that roughly corresponds to a migratory route followed by many waterfowl (Scott & Rose, 1996). Occasional dispersal of statoblasts by waterfowl may explain why European populations exhibited higher levels of connectivity than the North American populations, which were collected from sites traversed by a network of divergent migratory routes (Bellrose, 1968). Waterfowl are obvious agents for long-distance dispersal of bryozoan statoblasts. *C. mucedo* statoblasts ensnared in moulted feathers are frequently encountered along the shores of lakes (Okamura, pers. obs.), and have recently been observed to hatch following passage through the digestive tracts of ducks (I. Charambilidou, pers. comm.). In addition, statoblasts are produced most prolifically in autumn, a time that coincides with one of the two main periods of waterfowl migration.

Genetic variation

As discussed above, mtDNA data revealed a substantially greater number of genetic lineages both within and among populations in North America. Similarly, the microsatellite data revealed overall higher levels of diversity in North America. In North America, the two distinct genetic lineages suggested by the combinations of mitochondrial and microsatellite data may reflect cryptic species. While *C. mucedo* is assumed to be a monospecific genus, this conclusion is based on a fairly superficial knowledge of this species throughout much of its range. Furthermore, two different genome sizes have been discovered in North American *C. mucedo* (Potter, 1979). The relationship among the microsatellite alleles and the mitochondrial sequences in the two lineages identified in this study is imperfect, but this may be explained by occasional differential introgression of DNA following hybridization between the two lineages.

Although overall genetic diversity was higher in North America, within-population variation was the

same in the two continents. This suggests that similar local processes may affect genetic variation, such as the contribution of sexual reproduction. While the ecology of North American *C. mucedo* has not been well studied, similar measures of observed heterozygosity and a lack of Hardy–Weinberg proportions on the two continents suggest comparable levels of sex. We suggest that the disparate levels of within-continent genetic variation in North America and Europe are related to a historical bottleneck in north-western Europe, compounded by a current metapopulation structure. Ongoing gene flow provides evidence that dispersal links subpopulations within a metapopulation in Europe (Freeland *et al.*, 2000a), although there is little evidence for such a structure in North America (Freeland *et al.*, 2000b). Notably, the effective size of a metapopulation should be smaller than the sum of its component populations (Hedrick & Gilpin, 1997). Our European data agree with this prediction, as they show maintenance of genetic variation within local populations, but depletion of variation within the metapopulation as a whole.

Conclusions

In *C. mucedo* populations, both historical events and current population dynamics influence patterns of genetic diversity. In North America, a relatively large number of divergent mitochondrial genetic lineages have been maintained for the past 1.5 Myr, presumably within multiple glacial refugia throughout much of the Pleistocene. In marked contrast, a small number of mitochondrial lineages are found in north-western Europe, and these may be traced to the (re)colonization of this area either from North America or from European glacial refugia. However, neither a residual founder effect nor an apparent metapopulation structure has diminished levels of within-population genetic variability in Europe, perhaps in part because ongoing dispersal results in the continued introduction of novel genotypes into local populations. In contrast, North American *C. mucedo* have an overall higher level of genetic diversity which may result in part from hybridization between cryptic species, a process that has previously been inferred to increase genetic variation (Freeland & Boag, 1999). Our results suggest that caution should be employed when extrapolating population genetic patterns throughout a species range, and reveal that both historical and ongoing events affect current patterns of biodiversity.

Acknowledgements

We thank Viv Rimmer and the AMS sequencing lab for technical assistance. David Balding, Sarah Cluck, and Kelvin Conrad provided advice on analyses. Our work was supported by the Natural Environment Research Council (grant GR3/11068).

References

BARRETT, S. C. H., ECKERT, C. G. AND HUSBAND, B. C. 1993. Evolutionary processes in aquatic plant populations. *Aquat. Bot.*, **44**, 105–145.

BELLROSE, F. C. 1968. *Waterfowl migration corridors East of the Rocky Mountains in the United States*. Biological Notes no. 61. Illinois Natural History Survey, Urbana, IL.

BERG, D. J. AND GARTON, D. W. 1994. Genetic differentiation in North American and European populations of the cladoceran Bythotrephes. *Limnol. Oceanogr.*, **39**, 1503–1516.

BROWER, A. V. Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Helioconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 6491–6495.

COLBOURNE, J. K., CREASE, T. J., WEIDER, L. J., HEBERT, P. D. N., DUFRESNE, F. AND HOBÆK, A. 1998. Phylogenetics and evolution of a circumarctic species complex (Cladocera: *Daphnia pulex*). *Biol. J. Linn. Soc.*, **65**, 347–365.

CREASE, T. J., LYNCH, M. AND SPITZE, K. 1990. Hierarchical analysis of population genetic variation in mitochondrial and nuclear genes of *Daphnia pulex*. *Mol. Biol. Evol.*, **7**, 444–458.

DARWIN, C. 1859. *On the Origin of Species by Means of Natural Selection*. John Murray, London.

ESTOUP, A., GARNERY, L., SOLIGNAC, M. AND CORNUET, J.-M. 1995. Microsatellite variation in honey bee (*Apis mellifera* L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics*, **140**, 679–695.

FELSENSTEIN, J. 1995. PHYLIP *(Phylogeny Inference Package), v. 3.5c*. University of Washington (distributed by the author).

FREELAND, J. R. AND BOAG, P. T. 1999. The mitochondrial and nuclear genetic homogeneity of the phenotypically diverse Darwin's Ground Finches. *Evolution*, **53**, 1553–1563.

FREELAND, J. R., JONES, C. S., NOBLE, L. R. AND OKAMURA, B. 1999. Polymorphic microsatellite loci identified in the highly clonal freshwater bryozoan *Cristatella mucedo*. *Mol. Ecol.*, **8**, 341–342.

FREELAND, J. R., NOBLE, L. R. AND OKAMURA, B. 2000a. Genetic consequences of the metapopulation biology of a facultatively sexual freshwater invertebrate. *J. Evol. Biol.*, **13**, 383–395.

FREELAND, J. R., NOBLE, L. R. AND OKAMURA, B. 2000b. Genetic Diversity of North American populations of *Cristatella Mucedo*, inferred from microsatellite and mitochondrial DNA. *Mol. Ecol.*, **9**, 1375–1389.

FREY, D. G. 1995. Changing attitudes toward chydorid anomopods since 1769. *Hydrobiology*, **307**, 43–55.

GOODMAN, S. J. 1997. R-st calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Mol. Ecol.*, **6**, 881–885.

HANN, B. J. 1995. Genetic variation in *Simocephalus* (Anomopoda, Daphniidae) in North America: patterns and consequences. *Hydrobiology*, **307**, 9–14.

HEDRICK, P. W. AND GILPIN, M. E. 1997. Genetic effective size of a metapopulation. In: Hanski, I. and Gilpin, M. E., (eds) *Metapopulation Biology*, pp. 166–182. Academic Press, San Diego, CA.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, **16**, 111–120.

KUMAR, S., TAMURA, K. AND NEI, M. 1993. *MEGA: Molecular Evolutionary Genetics Analysis, v. 1.02*. Pennsylvania State University (distributed by the authors).

MANLY, B. F. J. 1994. *Multivariate Statistical Methods: a Primer*. Chapman & Hall, London.

MAYR, E. 1963. *Animal Species and Evolution*. Harvard University Press, Cambridge.

OKAMURA, B. 1997. The ecology of subdivided populations of a clonal freshwater bryozoan in southern England. *Arch. Hydrobiol.*, **141**, 13–34.

PALUMBI, S. R., MARTIN, A., ROMANO, S., MCMILLAN, W. S., STICE, S. AND GRABOWSKI, G. 1991. *The Simple Fool's Guide to PCR*. University of Hawaii Press, Honolulu, HI.

POTTER, R. 1979. Bryozoan karyotypes and genome sizes. In: Larwood, G. P. and Abbott, M. B. (eds) *Advances in Bryozoology*, pp. 11–32. Academic Press, London.

PRIMMER, C. R. AND ELLEGREN, H. 1998. Patterns of molecular evolution in avian microsatellites. *Mol. Biol. Evol*, **15**, 997–1008.

RAYBOUD, A. F., MOGG, R. J., ALDAM, C., GLIDDON, C. J., THORPE, R. S. AND CLARKE, R. T. 1998. The genetic structure of sea beet (*Beta vulgaris* ssp. *maritima*) populations. III. Detection of isolation by distance at microsatellite loci. *Heredity*, **80**, 127–132.

RAYMOND, M. AND ROUSSET, F. 1995. GENEPOP (Version 1.2) – Population genetics software for exact test and ecumenicism. *J. Hered.*, **86**, 248–249.

SCOTT, D. A. AND ROSE, P. M. 1996. *Atlas of Anatidae Populations in Africa and Western Eurasia*. Wetlands International, The Netherlands.

SHANNON, C. E. AND WEAVER, W. 1949. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, IL.

SLATKIN, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinctions. *Theor. Pop. Biol.*, **12**, 253–262.

SLATKIN, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

TAYLOR, D. J., HEBERT, P. D. N. AND COLBOURNE, J. K. 1996. Phylogenetics and evolution of the *Daphnia longispina* group (Crustacea) based on 12S rDNA sequence and allozyme variation. *Mol. Phyl. Evol.*, **5**, 495–510.

TAYLOR, D. J., FINSTON, T. L. AND HEBERT, P. D. N. 1998. Biogeography of a widespread freshwater crustacean: pseudocongruence and cryptic endemism in the North American *Daphnia laevis* complex. *Evolution*, **52**, 1648–1670.

THOMPSON, J. D., GIBSON, T. J., PLEWNIAK, F., JEANMOUGIN, F. AND HIGGINS, D. G. 1997. *The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools*. *Nucleic Acids Research.*, **24**, 4876–4882.

WHITLOCK, M. C. 1992. Temporal fluctuations in demographic parameters and the genetic variance among populations. *Evolution*, **46**, 608–615.

WRIGHT, S. 1951. The genetical structure of populations. *Eugenics*, **15**, 323–354.

YEH, F. C. AND BOYLE, T. J. B. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belg. J. Botany*, **129**, 157.