Vol. 549: 99–112, 2016
doi: 10.3354/meps11672

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published May 10

Genetic evidence supports recolonisation by *Mya arenaria* of western Europe from North America

M. E. Cross[1,*], C. R. Bradley[2], T. F. Cross[1], S. Culloty[1], S. Lynch[1], P. McGinnity[1], R. M. O'Riordan[1], S. Vartia[3], P. A. Prodöhl[2]

[1]Aquaculture and Fisheries Development Centre, School of Biological, Earth and Environmental Sciences, University College Cork, Ireland

[2]Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, UK

[3]Carna Research Station, Ryan Institute, National University of Ireland, Galway, Carna, Connemara, Co. Galway, Ireland

ABSTRACT: The softshell clam *Mya arenaria* (L.) is currently widespread on the east and west coasts of North America. This bivalve also occurs on western European shores, where the post-Pleistocene origin of the species, whether introduced or relict, has been debated. We collected 320 *M. arenaria* from 8 locations in Europe and North America. Clams (n = 84) from 7 of the locations were examined for mitochondrial DNA variation by sequencing a section of the cytochrome oxidase 1 (*COX1*) gene. These were analysed together with 212 sequences, sourced from GenBank, from the same gene from 12 additional locations, chiefly from eastern North America but also 1 site each from western North America and from western Europe. Ten microsatellite loci were also investigated in all 320 clams. Nuclear markers showed reduced levels of variation in certain European samples. The same common *COX1* haplotypes and microsatellite alleles were present throughout the range of *M. arenaria*, although significant differences were identified in haplotypic and allelic composition between many samples, particularly those from the 2 continents (Europe and North America). These findings support the hypothesis of post-Pleistocene colonisation of European shores from eastern North America (and the recorded human transfer of clams from the east to the west coast of North America in the 19th century).

KEY WORDS: *COX1* · Europe · Microsatellite loci · Mitochondrial DNA · Softshell clam · North America

INTRODUCTION

Following the retreat of the glaciers at the end of the Pleistocene some 12 000 years ago, recolonisation of northern Europe, in particular for most terrestrial species, is presumed to have originated from 1, 2 or 3 refugia in southern Europe, i.e. Iberia, Italy and the Balkans (Hewitt 1999). In the case of shallow-water and inter-tidal marine species, the situation was different in that post-glacial colonists from the south (contemporary Iberia, western North Africa and the Mediterranean) had to follow coastlines, which experienced rapid changes in sea level (Maggs et al. 2008). The fact that many marine species have pelagic larvae is thought to increase their potential for rapid colonisation (Luttikhuizen et al. 2003, Dupont et al. 2007, Arias-Pérez et al. 2012), despite other factors (e.g. ocean currents, salinity and temperature levels) which can prevent gene flow. In more recent times, planktonic larvae of marine species could also be distributed anthropogenically in ballast or bilge water of ships (Ruiz et al. 1997, Briski et al. 2012) or as adults, particularly gastropod or bivalve mollusc species, which, when retained as fresh food items, can survive long sea journeys, if held in low-temperature sea water.

*Corresponding author: maud.cross@ucc.ie

Notwithstanding the European refugial hypothesis (see Maggs et al. 2008), certain marine mollusc species currently occupying North Atlantic coastal regions appear to have had a different route of colonisation which involved trans-oceanic movement. The edible winkle *Littorina littorea* (Linnaeus, 1758) (Prosobranchia, Gastropoda) is thought to be one such species (Chapman 2007, Chapman et al. 2007, 2008). Arguably not present on eastern North American shores after the Pleistocene glaciations (Chapman 2007, Chapman et al. 2008, Blakeslee et al. 2008, Blakeslee & Byres 2008), the suggestion is that it was introduced as a food source to Newfoundland, or through rock ballast to Nova Scotia (Chapman et al. 2007). It was first recorded in the latter area in 1840 (Brawley et al. 2009) and thereafter spread south (Beebee & Rowe 2008, Harley et al. 2013). The opposite (colonisation of European shores from eastern North America; Strasser 1999) is believed to be the case for an infaunal bivalve species, the softshell clam *Mya arenaria* (Linnaeus, 1758), considered in the present study. Now recognised as an important member of the intertidal infaunal community in the Northeast Atlantic (Strasser 1999), the species occurs over a wide geographical range from Iceland, the White Sea and northern Norway to Portugal (Strasser 1999, Conde et al. 2010), including the Barents and Black Seas, with recent reports of its presence also in the Mediterranean Sea (Zaitsev & Mamaev 1997, Weston & Buttner 2010, Crocetta & Turolla 2011). Archaeological evidence suggests that *M. arenaria* was present in this region in the late Pliocene (Strauch 1972), and that it disappeared from European shores during the Pleistocene glaciations. *M. arenaria* is then thought to have reappeared, postglacially, on European shores. The reintroduction may have been either natural or anthropogenic, with previous suggestions that Vikings were responsible (MacNeil 1965, Strauch 1972, Petersen et al. 1992).

Molecular genetics, as used in the present study, provides a mechanism for testing the hypothesis of trans-Atlantic colonisation. If trans-Atlantic colonisation was exclusively or predominately the case, similar genetic composition would be expected in *M. arenaria* from the coasts of the eastern and western Atlantic. This constitutes the hypothesis explored here, i.e. that colonisation of western European shores occurred from eastern North America. In contrast, different genetic composition would be expected in softshell clams from either side of the ocean if post-Pleistocene northern European colonisation had involved southern or peri-glacial European refugia. In this case, genetic differences would have evolved as a consequence of forces such as mutation, genetic drift and selection, while the 2 assemblages were separated throughout the Pleistocene.

In addition, a single or limited number of colonisation events, limited to a small number of founders, would result in substantially lower genetic variability in northern European populations. Therefore, *M. arenaria* occurring south of the glaciated areas in eastern North America would be more genetically variable than those in Europe. In this way, another feature of the colonisation process can be elucidated.

Previous genetic studies of *M. arenaria* concentrated on eastern North American coasts and used markers such as allozymes (Levinton 1973, Morgan et al. 1978). Where European samples were included, little inter-continental variation was observed throughout the contemporary range (Lasota et al. 2004). As *M. arenaria* is known to be a successful coloniser within European waters (Bologa et al. 1995, Zaitsev & Mamaev 1997), and similar genetic composition was observed on each side of the Atlantic, a tentative agreement with the hypothesis of trans-Atlantic colonisation was formed (Lasota et al. 2004). Similarly, investigations of mtDNA using both the RFLP method and sequencing methodologies (Caporale et al. 1997, Sparagano et al. 2002) supported the trans-Atlantic colonisation concept, but discerned little or no population structure. These findings were interpreted as evidence of high levels of gene flow between areas. In their study of the cytochrome oxidase subunit I mitochondrial DNA (*COX1* mtDNA) gene of *M. arenaria*, based mainly in North America, Strasser & Barber (2009) observed a unique common haplotype in a single small sample from the Wadden Sea (Sylt, Germany). While this observation could have been linked to a possible existence of European refugia, it was dismissed on the basis of the relatively small number of individuals within samples, resulting in low discovery of rarer haplotypes (Strasser & Barber 2009).

Here, we considered sequences from a large section of the *COX1* mitochondrial gene, together with sequences previously published by Strasser & Barber (2009). Genetic data from 10 nuclear microsatellite DNA loci were also investigated. These molecular data were analysed to investigate several eastern North American samples, a single western North American and 5 European samples (the latter being the most intensive coverage to date of European samples). Both mitochondrial and nuclear data were included to test for correspondence in evolutionary inference between these 2 genomic systems. Recently published microsatellite loci were included

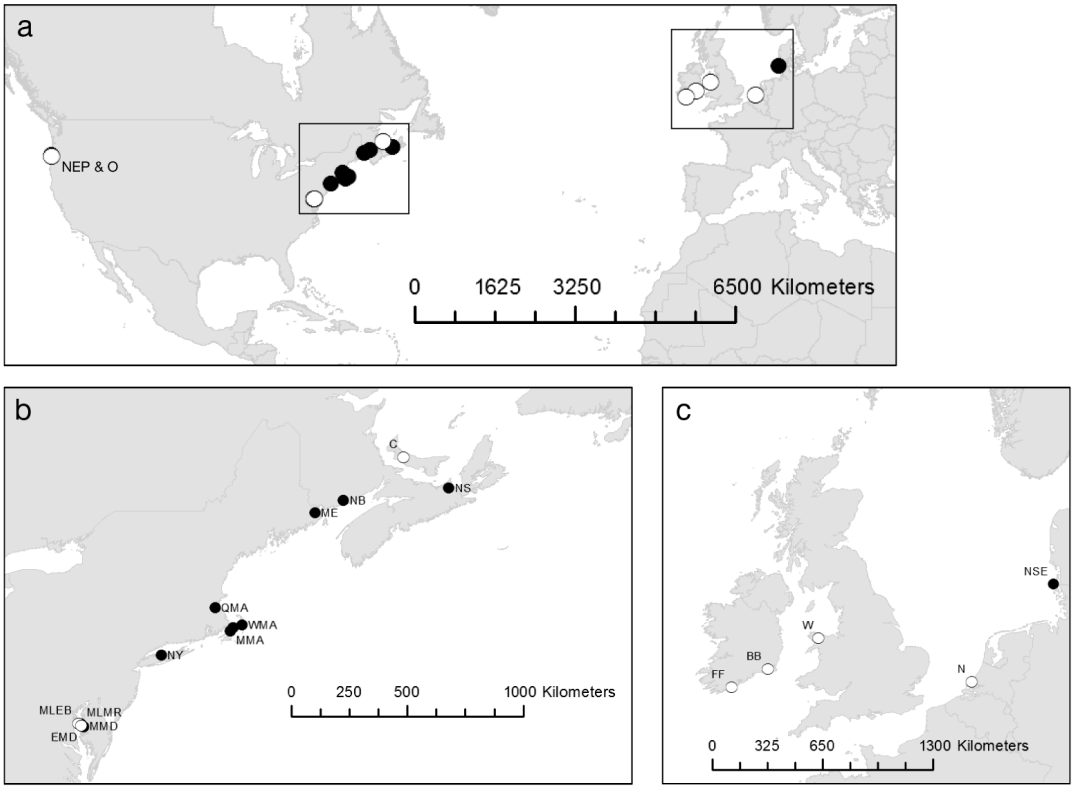

Fig. 1. Sampling locations of *Mya arenaria* (a) throughout its range, (b) in the eastern North American region (38.816° N, 76.237° W to 46.446° N, 63.767° W) and (c) on the western European coast (51.648° N, 08.698° W to 53.238° N, 04.091° W). Black circles represent samples analysed in the previous study of Strasser & Barber (2009), and white circles are samples collected for the present study. Sampling codes for the current samples are as provided in Table 2. The NEP site in Oregon is obscured in panel (a), and some other overlap of neighbouring sites occurs in panels (a) and (b) (see Tables 1 and 2 for additional site information)

(St-Onge et al. 2011, Krapal et al. 2012), as their intensive polymorphism and high evolutionary rate make this marker type ideal for determining routes of colonisation, quantifying relative numbers of founders and determining contemporary population structure in donor and colonised areas.

In summary, the aims of the current study were to investigate genetic composition at the *COX1* gene and 10 microsatellite loci in samples of *M. arenaria* from eastern and western North America and from north-western Europe, and to test the hypothesis about the origin of European colonisation.

MATERIALS AND METHODS

Sampling

Specimens of *Mya arenaria* (n = 320) were sampled from 8 intertidal locations throughout the contemporary species range, in Ireland, North Wales, the Netherlands, eastern Canada and 3 locations in the USA (Fig. 1). Specimens were collected by digging in the lower intertidal, or with a benthic hydraulic escalator dredge, at each site. Gill tissue samples from each individual were preserved in 90% molecular grade ethanol. Details of sampling locations are shown in Fig. 1 and Table 1.

Mitochondrial DNA

Ten to 16 *M. arenaria* specimens representing the sampling locations covered in the current study (Table 1) were randomly chosen for mtDNA sequencing analysis. Genomic DNA was extracted from gill tissue using a QIAamp DNA mini Kit (Qiagen). DNA quality and concentration were assessed by gel electrophoresis (0.8% 0.5× TBE agarose gel). For each of these samples, a 533 base pair (bp) region of the mtDNA *COX1* gene was amplified using the universal PCR primers HCO-2198 and LCO-1490 (Folmer et al. 1994). PCR amplifications were carried out in 50 µl volumes, containing 5 µl of 10× Invitrogen buffer,

Table 1. Sampling locations and numbers of *Mya arenaria* specimens analysed in this study, including summary statistics for 10 microsatellite loci (*a*: average number of alleles, R_s: average allelic richness, H_e: expected heterozygosity, H_o: observed heterozygosity)

Code	(n)	Sampled area	Lat (°N), Long (°W)	*a*	R_s	H_e	H_o
MLEB	43	Eastern Bay, Maryland, eastern USA, North America	38.886, 76.340	22.2	11.14	0.901	0.798
MLMR	53	Miles River, Maryland, eastern USA, North America	38.816, 76.237	24.8	11.15	0.901	0.831
C	30	Prince Edward Island, Gulf of St. Lawrence, eastern Canada	46.446, 63.767	14.6	9.01	0.844	0.747
N	30	Hinderplaat, The Netherlands, Europe	51.843, –04.009	12	8.04	0.837	0.737
BB	48	Bannow Bay, Wexford, Ireland, Europe	52.251, 06.763	8.9	5.99	0.752	0.678
FF	31	Flaxfort, Cork, Ireland, Europe	51.648, 08.698	8.9	6.21	0.733	0.686
W	25	Menai Bridge, Wales, Europe	53.238, 04.091	7.2	5.65	0.706	0.759
O	60	Alsea Bay, Oregon, western USA, North America	44.584, 123.964	21.9	10.10	0.894	0.914

3.5 µl $MgCl_2$ (50 mM), 2 µl dNTPs (5 mM), 1 µl of each primer (10 pM $µl^{-1}$), 0.5 µl BSA (10 mg ml^{-1}), 0.8 µl (4 U) Ampli-*Taq* polymerase (Invitrogen), 34.2 µl dH_2O and 2 µl (~50 ng $µl^{-1}$) template DNA. PCR amplifications were carried out on a Techne TC-Plus thermocycler as follows: 95°C for 15 min, 5× (95°C for 60 s, 50°C for 60 s, 72°C for 60 s), 30× (95°C for 60 s, 52°C for 60 s, 72°C for 60 s), and 72°C for 5 min.

Amplified products were checked for quality by gel electrophoresis (1.5% 0.5× TBE agarose gel) and purified using EXOSAP-IT (Affymetrix, USB Products) purification kits. Purified products were bi-directionally sequenced using HCO-2198 and LCO-1490 primers and the Big Dye Terminator (V3.1) sequencing chemistry (Applied Biosystems) following the manufacturer's recommendation. Cycle sequencing reactions were carried out on a Techne TC-Plus thermocycler as follows: 94°C for 3 min, 25× (94°C for 10 s, 50°C for 5 s, 60°C for 4 min) and 8°C for 10 min. Resulting sequencing products were purified by ethanol precipitation (125 mM EDTA and 100% molecular grade ethanol). Sequencing was carried out using an ABI 3730XL 96 capillary system DNA analyser (Applied Biosystems).

The sequencing analysis of 212 *M. arenaria* specimens from 12 additional sites (Fig. 1, Table 2) were included in the current analysis. This allowed the integration of novel mtDNA sequencing data generated within this investigation to existing lineages identified in a previous study by Strasser & Barber (2009).

Table 2. Summary sample statistics derived from the mitochondrial *COX1* sequencing data including number of individuals (N), number of haplotypes (nh), haplotype diversity ($h \pm SD$), nucleotide diversity ($\pi \times 10^{-2}$), and average number of nucleotide differences (*k*). **Bold** font identifies samples collected in the current investigation

Sampled area	Sample code	N	nh	*h*	π	*k*
Eastern Bay, Maryland A	MLEB	12	4	0.455 ± 0.17	0.094	0.500
Miles River, Maryland B	MLMR	12	8	0.894 ± 0.08	0.244	1.303
Eastern Bay, MD	EMD	9	4	0.583 ± 0.18	0.167	0.889
Miles River, MD	MMD	15	4	0.371 ± 0.15	0.096	0.514
Stony Brook, NY	NY	15	5	0.476 ± 0.16	0.100	0.533
Wareham, MA	WMA	25	7	0.430 ± 0.12	0.119	0.633
Mashpee, MA	MMA	20	6	0.447 ± 0.14	0.110	0.600
Barnstable, MA	BMA	19	3	0.001 ± 0.12	0.039	0.211
Quincy, MA	QMA	22	3	0.177 ± 0.11	0.034	0.182
Pembroke, ME	ME	21	4	0.271 ± 0.12	0.071	0.381
St. John, NB	NB	11	4	0.491 ± 0.03	0.102	0.545
Antigonish, NS	NS	20	1	–	–	–
Prince Edward Island, Canada	C	16	2	0.125 ± 0.11	0.023	0.125
Sylt, Germany	NSE	15	3	0.530 ± 0.14	0.108	0.576
Netherlands	N	12	3	0.530 ± 0.14	0.108	0.576
Bannow Bay, Ireland	BB	10	4	0.800 ± 0.08	0.226	1.200
Flaxfort, Ireland	FF	10	3	0.711 ± 0.09	0.167	0.889
Wales	W	12	1	–	–	–
Newport, OR	NEP	20	5	0.574 ± 0.74	0.0013	0.668
Overall		296	34	0.422 ± 0.04	0.1071	0.521

Microsatellite DNA

For microsatellite analysis, genomic DNA was extracted from all *M. arenaria* specimens (n = 320) using the CHELEX-100 resin (Bio-Rad) extraction method. All samples were screened for 10 microsatellite marker loci described by St-Onge et al. (2011) (*Mar*01, *Mar*04, *Mar*06, *Mar*07, *Mar*08) and Krapal et al. (2012) (Ma02, Ma06, Ma11, Ma14, Ma15), which were produced using *M. arenaria* microsatellite clone sequences from GenBank (accession numbers: JN850609.1, JN850610.1, JN850611.1, JN850612.1, JN850613.1, JN850614.1, JN850615.1,

JN850616.1, JN850617.1; Krapal et al. 2012). Primers were designed using Primer3Plus (Rozen & Skaletsky 2000, Untergasser et al. 2007) with optimal primer length as 20 bp and optimal maximum temperature (T_m) at 60°C. The forward primers were tailed with 1 of 4 universal dye-labelled tails at their 5′ end: T3: PET: 5′-AAT TAA CCC TCA CTA AAG GG-3′, M13 Reverse: NED: 5′-GGA TAA CAA TTT CAC AC AGG-3′ (Diniz et al. 2007), Hill: 6FAM: 5′-TGA CCG GCA GCA AAA TTG-3′ (Tozaki et al. 2001) and Neomycin rev: VIC: 5′-AGG TGA GAT GAC AGG AGA TC-3′. PIG-tails were added to the 5′ end of all the reverse primers (Brownstein et al. 1996). Primers were combined into 2 multiplex PCR reactions with Multiplex Manager 1.0 (Holleley & Geerts 2009).

Microsatellite screening was performed using multiplex PCR amplification carried out in reaction volumes of 5 µl, including 1 µl of DNA (approximately 50 ng), 2.5 µl of Qiagen 2× Multiplex Mix™, 0.5 µl of 10× primer mix (outlined below) and 1 µl dH_2O. The forward primer in each case incorporated a tail, to which a labelled tail would anneal during PCR. Final reaction primer concentrations were 0.2 µM of forward (tailed) primer, 0.2 µM of reverse primer (which was PIG-tailed) and 0.05 µM of labelled tail primer. Four different labelled tails were used (1 for each channel, FAM, NED, VIC and PET). Tailing methods for each locus are outlined in Table S1 in Supplement 1 at www.int-res.com/articles/suppl/m549p099_supp.pdf. Amplifications were performed on a Techne TC-Plus thermocycler as follows: 15 min at 95°C; 30 cycles of 30 s at 94°C, 90 s at 56°C and 60 s at 72°C; 8 cycles of 30 s at 94°C, 90 s at 53°C and 60 s at 72°C; with a final step of 30 min at 60°C. To minimise genotyping error, PCR amplification of each individual was repeated at least 3 times, and 3 known genotype individuals were run on each plate for control purposes. All PCR products were run on an ABI 3730XL Genetic Analyser 16 capillary system (Applied Biosystems) and sized with an internal lane standard (GENESCAN 500 LIZ, Applied Biosystems) using the program GENEMARKER v 1.6 (softgenetics).

Data analysis

Mitochondrial sequences

Resulting sequences were checked for base call quality and ambiguity using CHROMASPRO 150 (version 1.7.6) (Technelysium) and subsequently aligned using 'CLUSTAL W' implemented in BIOEDIT (version 7.1.9; Hall 1999), alongside sequences representing all haplotypes described by Strasser & Barber (2009). Summary sample diversity statistics, including haplotype (h) and nucleotide diversity (π) and average number of nucleotide differences (k) between haplotypes, were estimated using DNASP v5.1 (Rozas et al. 2003). To assess the level of genetic structuring both on macro- (i.e. trans-Atlantic) and micro- (North America and Europe) geographical levels, sample pairwise Φ_{st} estimators and associated p-values were calculated in ARLEQUIN 3.01. Patterns of genetic subdivision both at macro- and micro-geographical scales were evaluated using analysis of molecular variance (AMOVA), also using ARLEQUIN 3.01. Statistical significance of the AMOVA was tested by permutation (10 000). To examine the relationship among resulting mtDNA haplotypes and to provide additional insights into the phylogeography of *M. arenaria* on both sides of the Atlantic, a haplotypic network was constructed using the median-joining method (Bandelt et al. 1999), implemented in Network 4.5.0.2 (Fluxus Technology: fluxus-engineering.com). For this analysis, sequencing data from Strasser & Barber (2009) were also included for comparison.

Microsatellites

Within-sample statistics, including total number of alleles and mean number of alleles, expected and observed heterozygosity and allele richness (Kalinowski 2004), were estimated using diveRsity (Keenan et al. 2013). The statistical significance of observed differences in the measures of allelic richness between samples (pairwise tests) was assessed using a sign test (Wilcoxon signed-rank tests) across loci as suggested by Kalinowski (2005).

Deviations from Hardy-Weinberg equilibrium (HWE) were assessed using exact probability tests implemented in GENEPOP 3.4 (Raymond & Rousset 1995). Significance levels for multiple comparisons were adjusted using standard Bonferroni correction (Rice 1989). Because null alleles can result in an underestimation of within-population genetic variation, loci that deviated from HWE, with heterozygote deficiencies, were evaluated for the existence of null alleles following an approach implemented in MICRO-CHECKER 2.2.3 (van Oosterhout et al. 2004).

Genetic divergence among samples was compared using F_{ST} (Weir & Cockerham 1984), estimated using diveRsity (Keenan et al. 2013), and statistical signifi-

cance was assessed by bootstrapping loci (5000). Patterns of nuclear genetic subdivision observed both at macro- and micro-geographical scales were evaluated using analysis of molecular variance (AMOVA) also using ARLEQUIN 3.01 with 10 000 permutations.

To further investigate the presence and patterns of population sub-structuring within the sample, an unrooted neighbour-joining (NJ) tree was created based on genetic distances of Nei et al. (1983). Genetic distances and bootstrap values (10 000) were calculated using Populations 1.2.3.1. A tree was then drawn using Figtree (http://tree.bio.ed.ac.uk/software/figtree/).

As an additional approach to examine the patterns of population genetic structure within the data, the Bayesian analytical framework implemented in the program STRUCTURE (Pritchard et al. 2000) was used. This framework allows for the identification of genetic partitioning (i.e. genetic sub-structuring) within a sample for which no *a priori* hypothesis of population structuring exists. Thus, given a sample set consisting of genotypic data from a number of specimens, the analytical framework estimates the minimum number of Mendelian populations (i.e. independent genetic entities) that best explains the data. Here, the STRUCTURE analysis was carried out using a hierarchical approach, which was primarily intended to identify major genetic evolutionary lineages within the data and subsequently refining these to the population level. This analysis was carried out as follows: (1) using the complete data set, STRUCTURE runs were set for k varying from 1 to 10 (20 interactions for each k value), with 100 000 burn-in iterations, followed by 100 000 MCMC iterations for each independent run. STRUCTURE analysis was carried out using the admixture model with correlated allele frequencies. Group matching among independent run interactions was carried out using CLUMP v.1.1.2b (Jakobsson & Rosenberg 2007) employing the 'greedy' algorithm with random input orders and 20 000 repeats. In each case, the optimal k was determined using the ad hoc method of Evanno et al. (2005). (2) Following this initial run, in order to identify further genetic sub-structuring within the data (i.e. sub-groups within the main groups), STRUCTURE was independently run with genotypic data from each of the main groups separately. In each instance, the same approach/parameters adopted for the whole data set were used. If additional sub-structuring was identified, STRUCTURE was again independently run against samples representing each sub-group using the same approach/parameters. This iterative/hierarchical approach was repeated until no further sub-structuring was evident from the data. Results of this hierarchical analysis were summarised in graphical format using STRUCTURE plots. Finally, as an alternative non-parametric approach to assess the existence of population sub-structuring within the data, a factorial component analysis was carried out using GENETIX 4.05 (Belkhir et al. 1996–2004). The results of this analysis were summarised using a 3-dimensional plot generated in MATLAB 2.04 (Mathworks) using the scatter 3-function.

RESULTS

MtDNA *COX1* region

The alignment of all *COX1* gene sequences (i.e. 84 generated in the present study in addition to the 212 sequences from Strasser & Barber 2009) revealed 34 haplotypes, of which 27 were previously described by Strasser & Barber (2009) and 7 were novel (KU720416, KU720417, KU720418, KU720419, KU720420, KU720421, KU720422, KU720423, KU720424, KU720425, KU720426, KU720427, KU720428, KU720429, KU720430; Table 3). Haplotype A was the only one to be shared among all samples on both sides of the Atlantic and in the Pacific. It was predominant in all except for the samples from Ireland (Bannow Bay, BB; Flaxfort, FF). As reported by Strasser & Barber (2009), with few exceptions, most haplotypes were restricted to particular sampling locations, occurring once only. Among the notable exceptions was haplotype B, which was relatively common throughout the sampling range but appears to be particularly common in Europe, and haplotype E, which, with 1 exception (the Prince Edward Island, Canada, sample, C), was restricted to European samples. Of the 7 novel haplotypes described in the present study, 5 were restricted to North America and 2 (AB and AC) were restricted to Europe.

The median-joining network depicting the genetic relationship among haplotypes revealed a star-shaped phylogeny (Fig. 2). In most cases, haplotypes differed from each other by a single mutational step stemming from haplotype A.

Within-sample variability

Haplotype diversity (h) ranged from a 0 value in Wales (W) and Nova Scotia (NS), which were fixed for haplotype A, to 0.800 and 0.894 in Miles River, Maryland (MLMR) and Bannow Bay, respectively,

Table 3. MtDNA haplotype distribution of the analysed *COX1* region of *Mya arenaria*. Samples are aligned geographically to represent the hypothesised (Strasser & Barber 2009) post-Pleistocene spread of the species. Sampling site abbreviations are defined in Tables 1 & 2. The shaded columns represent European sampling sites, and shaded rows represent haplotypes unique to European sites. Underlined sampling sites represent novel samples collected in the current study. Nhap: no. haplotypes per sample. Other samples are from GenBank (see 'Materials and methods' for details)

mtDNA haplotype	Sample location																			
	MLEB	MLMR	EMD	MMD	NY	WMA	MMA	BMA	QMA	ME	NB	NS	C	NSE	N	BB	FF	W	NEP	All
A	9	4	6	12	11	19	15	17	20	18	8	20	15	8	8	3	4	12	13	222
B				1	1	1					1			3	1	3	2		1	14
C			1							1									2	4
D							1				1								2	4
E													1	4	2	3	4			14
F		2			1	1	1													5
G																			2	2
H											1									1
I										1										1
J		1								1										2
K									1											1
L									1											1
M								1												1
N								1												1
O							1													1
P							1													1
Q		1					1													2
R						1														1
S	1					1														2
T						1														1
U						1														1
V					1															1
W					1															1
X				1																1
Y				1																1
Z			1																	1
AA		1	1																	2
AB																1				1
AC															1					1
AD	1																			1
AE	1																			1
AF		1																		1
AG		1																		1
AH		1																		1
Total	12	12	9	15	15	25	20	19	22	21	11	20	16	15	12	10	10	12	20	296
Nhap	4	8	4	4	5	7	6	3	3	4	4	1	2	3	4	4	3	1	5	

with an average of 0.422 (Table 2). Nucleotide diversity (π) was relatively low for all localities, ranging from 0.00023 in Prince Edward Island to 0.00245 in Miles River. The values for the average number of nucleotide differences (k) ranged from 0 in Wales and 0.125 in Canada to 1.303 in the Miles River sample (Table 2).

Among-sample variability

After standard Bonferroni correction, pairwise inter-population Φ_{st} comparisons revealed certain samples of *Mya arenaria* to be statistically significantly different from each other (Table 4), including the Irish samples with Welsh and both Maryland (MLEB, MLMR) *M. arenaria*, and the Prince Edward Island sample with the Flaxfort and Miles River individuals. The Netherlands (N) sample was not significantly different from any of the other areas using this method. In the hierarchical AMOVA, 90% of mtDNA *COX1* variation was found within samples, whereas 7.9 and 2.1% of total haplotype frequency variation (both values significantly different from 0) were represented by variation between European and North American samples and among samples within regions, respectively (Table S2 in Supplement 1).

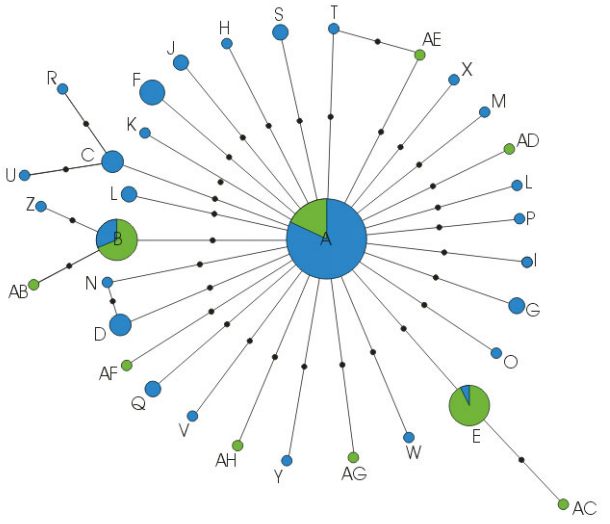

Fig. 2. Median-joining network of *Mya arenaria* mtDNA haplotypes derived from *COX1* sequencing data. The circle area is proportional to the number of individuals characterised by a particular haplotype. The small black dots in the connecting lines represent mutational steps. Colours represent geographic locations where haplotypes were found (blue: North America, green: Europe). Reticulations for unresolved links (e.g. T-A-AE and N-A-D) among haplotypes are also shown. Haplotype designations (A to AH) are as used in Table 3

Microsatellites

Intra-sample parameters

We genotyped 320 *M. arenaria* from 8 geographical sites across the species range at 10 microsatellite loci (Supplement 2 at www.int-res.com/articles/suppl/m549p099_supp.xls). No linkage disequilibrium was observed between locus pairs, and all 10 loci were polymorphic. The 3 loci that deviated from HWE with heterozygote deficiencies (Ma 02, Ma 06 and Ma 15) were positively identified as possessing null alleles in MICRO-CHECKER 2.2.3 (van Oosterhout et al. 2004). Since the apparent presence of null alleles at these 3 loci had a negligible effect on the estimation of overall F_{ST} (uncorrected value: 0.058; corrected value: 0.057), all 10 loci were used in calculation of pairwise F_{ST} values (Table 5). The average number of alleles per sample varied from 7.2 in the Welsh sample (W), to 24.8 in the Miles River, Maryland, sample (MLMR) (Tables 1 & S3, the latter in Supplement 1), with the total average number of alleles per locus of 15, and a total of 355 alleles being detected overall. Across all loci sampled, alleles unique to an area were present in small numbers in all sites except the Welsh sample (W) (Table S4 in Supplement 1).

Table 4. Pairwise population Φst values for mtDNA sequence data, showing population pairwise Φst values below the diagonal and associated p-values above. The values with a '+' sign above the diagonal are significant after standard Bonferroni correction. Site abbreviations are as given in Tables 1 and 2

	MLEB	MLMR	EMD	MMD	NY	WMA	MMA	BMA	QMA	ME	NB	NS	C	NSE	N	BB	FF	W	NEP
MLEB		–	–	+	–	–	–	+	+	–	–	+	+	–	–	+	+	–	+
MLMR	0.017		–	–	–	–	–	–	–	–	–	+	–	–	–	+	+	–	–
EMD	−0.008	0.008		–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–
MMD	0.035	0.018	−0.042		–	–	–	–	–	–	–	–	–	–	–	+	+	–	–
NY	0.000	−0.001	−0.010	−0.016		–	–	–	–	–	–	+	–	–	–	+	+	–	–
WMA	0.023	−0.017	−0.020	0.001	−0.015		–	–	–	–	–	–	–	–	+	+	+	–	–
MMA	0.003	−0.003	0.013	0.015	−0.013	−0.001		–	–	–	–	–	–	+	+	+	+	–	–
BMA	0.049	0.014	0.047	0.033	0.007	0.001	−0.008		–	–	–	–	–	+	+	+	+	–	–
QMA	0.061	0.021	0.060	0.040	0.012	0.005	0.003	0.001		–	–	–	–	+	+	+	+	–	+
ME	0.033	0.005	0.012	0.024	0.003	−0.010	0.001	−0.001	0.001		–	–	–	+	+	+	+	–	–
NB	0.013	0.000	−0.024	−0.028	−0.023	−0.010	−0.018	0.006	0.028	0.008		+	–	–	–	+	+	–	–
NS	0.073	0.045	0.097	0.060	0.020	0.000	0.000	0.003	−0.004	−0.002	0.058		–	+	+	+	+	–	–
C	0.043	0.014	0.047	0.032	0.003	−0.006	−0.008	−0.003	−0.003	−0.007	0.021	0.014		–	–	+	+	–	–
NSE	0.061	0.078	−0.008	−0.021	0.022	0.043	0.071	0.128	0.144	0.098	0.000	0.208	–		–	–	–	–	–
N	0.061	0.078	0.043	0.058	0.059	0.062	0.071	0.128	0.144	0.098	0.051	0.208	0.057	0.013		–	–	–	+
BB	0.146	0.202	0.094	0.128	0.166	0.193	0.214	0.293	0.319	0.259	0.132	0.375	0.258?	0.001	0.038		–	+	+
FF	0.130	0.191	0.111	0.152	0.167	0.177	0.190	0.283	0.308	0.239	0.146	0.385	0.226	0.049	−0.043	−0.044		+	+
W	0.023	0.000	0.034	0.017	−0.016	−0.025	−0.028	−0.026	−0.031	−0.029	0.008	0.000	−0.019	0.136	0.136	0.280	0.290		–
NEP	0.044	0.021	−0.015	0.017	0.013	−0.001	0.009	0.022	0.039	0.013	−0.023	0.045	0.028	0.052	0.075	0.184	0.175	0.010	

Table 5. F_{ST} estimates for 10 microsatellite loci, after correction for the presence of null alleles. Site abbreviations are as given in Table 1. Pairwise F_{ST} values are shown

	MLEB	MLMR	C	N	BB	FF	W
MLMR	0.0017[a]						
C	0.0381	0.0379					
N	0.0316	0.0377	0.0648				
BB	0.0701	0.0742	0.0923	0.0513			
FF	0.0763	0.0771	0.0996	0.0549	0.0165		
W	0.0910	0.0832	0.1131	0.0975	0.0909	0.1009	
O	0.0174	0.0182	0.0486	0.0355	0.0773	0.0816	0.0980

[a]Value is not statistically significantly different after standard Bonferroni correction

Observed and expected within-sample heterozygosity values varied from 0.678 (Flaxfort, FF) to 0.914 (Oregon, NEP), and from 0.706 (Wales, W) to 0.901 (Maryland samples, MLEB and MLMR) (Tables 1 & S3, the latter in Supplement 1). The highest percentage of unique alleles present in an area was found in the Oregon sample (NEP) of *M. arenaria*, and the lowest in the Welsh sample (W) (Table S4). Genetic variability as measured by allelic richness (R_s) ranged from 11.1 in the 2 Maryland sites (MLEB and MLMR), to 5.6 in the Welsh sample (W), with a mean allelic richness of 8.41. Results from the Wilcoxon signed-rank tests for differences in allele diversity (i.e. allele richness) among samples indicates significant statistical outcomes between all pairwise comparison involving samples from North America and Europe ($p < 0.01$ in all instances). Thus, samples from North America consistently displayed significantly higher levels of allele diversity in comparison to European samples. There were no significant differences in allele richness among pairwise comparison involving North American samples only. Among the European samples, however, the Netherlands sample was found to display significantly higher levels of allele diversity ($p < 0.01$) in comparison to all other European samples, which in turn, displayed similar levels of allele diversity (i.e. not statistically significant).

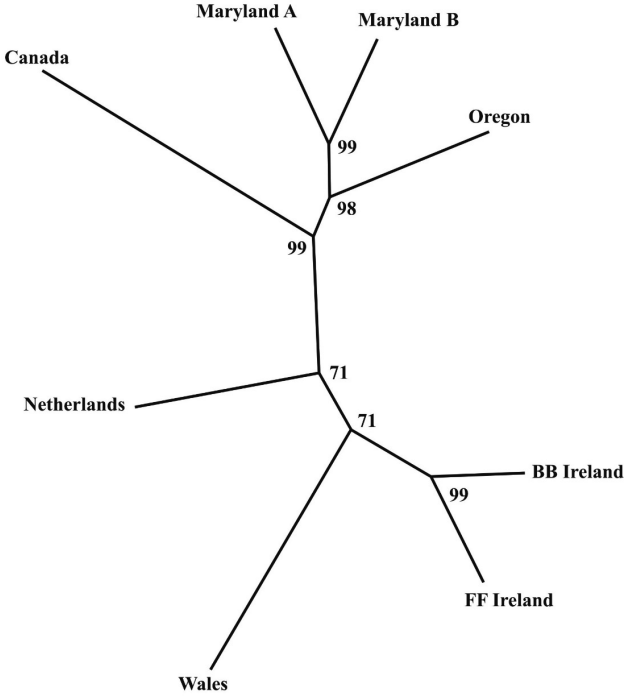

Fig. 3. Unrooted neighbour-joining tree based on genetic distances of Nei et al. (1983) (D_a).Values on nodes represent percentage bootstrap support for groupings (out of 10 000 bootstraps). For abbreviations see Table 2

Inter-sample parameters

Genetic differentiation over the 8 softshell clam samples was significant at an overall value of 0.0224 (95% CI 0.0140–0.0341). Pairwise F_{ST} values ranged from the single not statistically significant 0.0017 between the 2 Maryland sites (MLEB, MLMR), to 0.113 and 0.101 between Welsh (W) and Canadian (C), and Welsh (W) and Flaxfort (FF) *M. arenaria*, respectively (Table 5). All samples of *M. arenaria* were significantly different from each other, except those from the 2 Maryland sites.

The population structuring suggested by the pairwise F_{ST} analysis was further supported in the un-rooted NJ tree based on inter-sample D_a values (measure of genetic distance), and showed the European and North American samples to be clearly separated, with the Netherlands (N) sample being in an intermediate position (Fig. 3). Samples of *M. arenaria* from the 2 Irish sites, Bannow Bay (BB) and Flaxfort (FF), clustered together on the NJ tree, as did the 2 samples from Maryland. The hierarchical STRUCTURE plots (Fig. 4) provide further support to these inter-sample relationships. Samples from the 2 continents separated at the first level (i.e. Level 0 in Fig. 4). Significant within-continent variation was also identified at the second hierarchical level (Level 1, groups 1 and 2), with no further obvious differentiation between the 2 Maryland or Irish samples (i.e. third hierarchical level, Level 2). The significant level of pop-

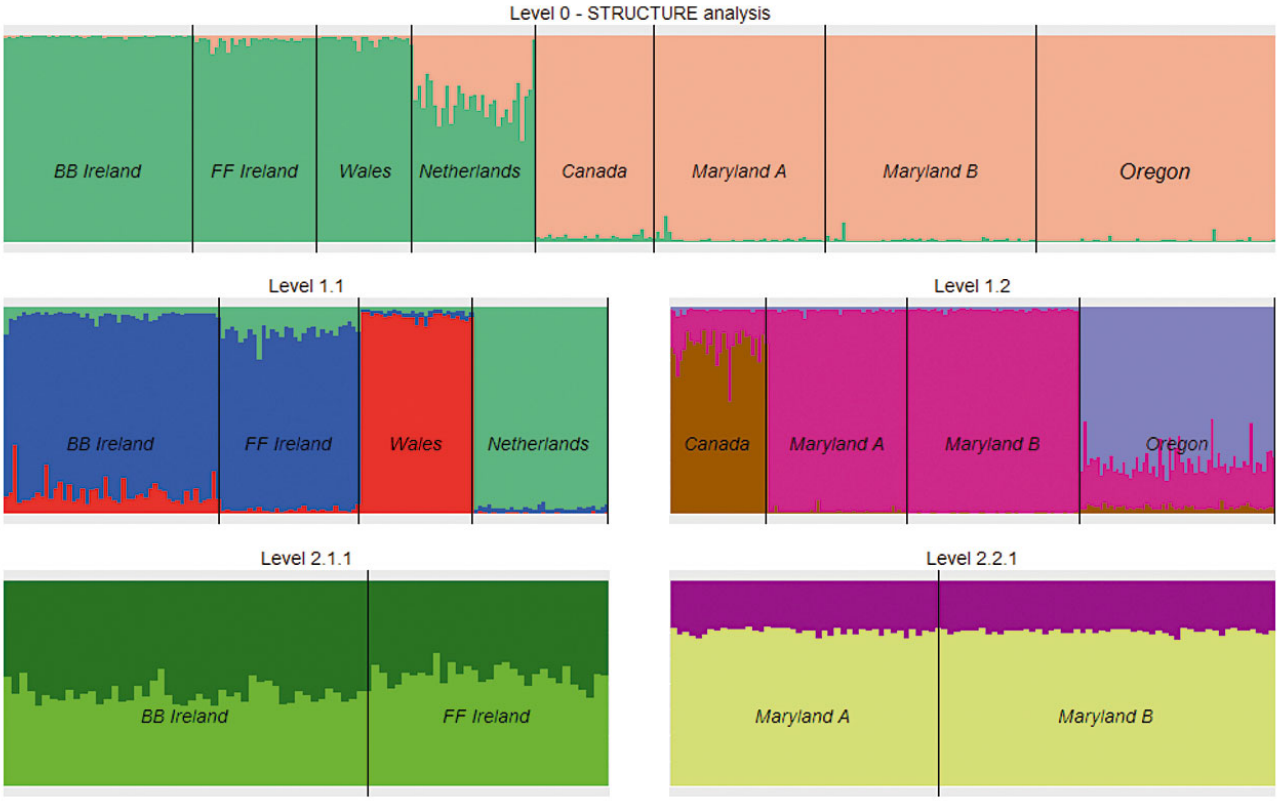

Fig. 4. Estimated population structure from hierarchical STRUCTURE analyses for the full data set (Level 0) and independent STRUCTURE runs subsequently carried out on samples representing the major identified groups (i.e. hierarchical Levels 1 and 2 and respective subgroups). In each instance (i.e. hierarchical level and group), each particular individual is represented by a thin vertical line, which is divided into *k* coloured segments that represent the individual's estimated membership fractions in *k* clusters. Black thicker lines separate individuals from different samples. For abbreviations see Table 2

ulation genetic structuring observed with pairwise F_{ST} estimates and confirmed from both NJ and STRUCTURE analyses was further supported by the independent results of factorial component analysis (Fig. S1 in Supplement 1 at www.int-res.com/articles/suppl/m549p099_supp.pdf), which shows clear separation among samples from Europe and North America in addition to 'within major regions' sub-structuring.

Hierarchical AMOVA analyses corroborate previous analysis by indicating significant differences in genetic partitioning of *M. arenaria* at all levels of geographic organisation (with regional partitioning as used above for *COX1* data), i.e. between the 2 regions (Europe and North America), among populations within regions and within populations (Table S5). As with mtDNA *COX1* AMOVA results, most variation was due to differentiation within populations (88.29%), while 5.48% and 6.23% of total variation were represented by variation among Europe and North American samples, and among populations within regions, respectively.

DISCUSSION

As noted in the 'Introduction', previous studies of *Mya arenaria* failed to demonstrate significant inter- and intra-continental differences between samples. Conversely, in the present study, virtually all samples were significantly different for pairwise multi-locus microsatellite F_{ST} values, except those in close proximity in Chesapeake Bay (Maryland) (Table 5). The greatest difference was between samples from either side of the Atlantic (see AMOVA results in Table S4). This suggests that all but 2 of these samples should be regarded as largely distinct populations, with little contemporary gene flow between them.

In the present investigation, using a combination of 2 independent molecular methods in this species was novel. Another innovation with these markers, which have previously proven to be discriminatory in eastern North America (Strasser & Barber 2009, St-Onge et al. 2011), was the inclusion of more than 1 European sample. This allows consideration of aspects such as possible colonisation modes and routes, and

population variability within Europe. Both molecular methods indicate major or exclusive colonisation from eastern North America, in that European samples shared the same common mitochondrial haplotypes or common microsatellite alleles with eastern North American samples. In addition, North American samples displayed significantly higher levels of genetic diversity in comparison to European samples in a pattern that is consistent with source populations. Thus, we suggest that, based on current data, eastern North American populations were the major source of post-Pleistocene colonisers to northern Europe.

Routes of European colonisation

Previous authors (Petersen et al. 1992, Strasser 1999) have suggested that *M. arenaria* was introduced into Europe by Vikings returning to Europe in pre-Columbian times. Alternatively, natural colonisation could have occurred by larval drift from west to east via northern coastal locations. Future sampling of *M. arenaria* in the more northerly parts of the contemporary range in eastern North America and western Europe, and in transitional regions such as Iceland and Greenland, might clarify the possible routes of colonisation, as has been the case in molecular investigations of *Arctica islandica* (Dahlgren et al. 2000), *Cerastoderma edule* (Krakau et al. 2012), *Macoma balthica* (Luttikhuizen et al. 2003) and *Littorina saxatilis* (Panova et al. 2011).

Regardless of whether European colonisation was human mediated or natural, data obtained in the current study would suggest that colonisation occurred from the north of the eastern North American range. Analysis of the relationship at microsatellite loci between the *M. arenaria* samples suggests a closer relationship between the Prince Edward Island sample (north of the eastern North American range) and the Netherlands sample (North Sea, Europe), than for samples from farther south in the eastern North American range (Maryland, USA) (Fig. 3).

Assuming that variability decreases from the point of first colonisation (Hewitt 2000), levels of microsatellite variability (i.e. allele diversity) and genetic similarity observed in the 4 European samples may give some indication of the sequence of north-western European colonisation. The average level of genetic diversity was significantly higher in the North Sea (Netherlands) sample than in other European samples (from the Celtic Sea, southern Ireland and Wales; Table 1). Following the theory proposed by Hewitt (2000), this suggests that the North Sea could have been the area of first European colonisation from which other areas would have subsequently been colonised. While no significant differences were observed among the other European samples, it is interesting to note that the sample from Wales displayed the lowest level of genetic diversity numerically, thus suggesting that this area may have been the last to be colonised. The resulting NJ tree (Fig. 3) and the results from the STRUCTURE analysis (Fig. 4) support this interpretation, as do pairwise F_{ST} values for microsatellites (Table 5). Indeed, the F_{ST} estimates might seem to suggest a 'stepping stone' sequence of progressive colonisation, similar to that proposed by Allendorf & Luikart (2008).

The situation is far less clear cut for the mitochondrial data, possibly because of the lower variability observed in *COX1* sequences and the use of substantially lower sample sizes. Mitochondrial DNA pairwise Φ_{st} estimates for Canada compared with 3 of the European samples (the Netherlands, Bannow Bay and Flaxfort) provided the same trend of increasing Φ_{st} estimates as observed for the microsatellites (Table 5). In contrast, the pairwise mitochondrial DNA Φ_{st} between Canada and Wales was 0, while an F_{ST} value of 0.113 for microsatellite data from the same samples was higher than those between the Canadian and other European samples (Table 5). The examination of genetic variability (*h* values) for *COX1* also showed a complex pattern (Table 2), where *h* values of the Irish samples were particularly high. However, these results must be interpreted with caution because of the small sample size involved. One possibly hypothesis is that at *COX1*, the Irish samples display evidence of a separate colonisation event, either from eastern North America or from a possible southern European refugium (Maggs et al. 2008)

Possible European refugia?

The concept that European refugia, if they existed, should have occurred at the south of the range assumes that coastal marine species spread northwards post-glacially (Hewitt 2000). Many marine species, including the northern quahog *Mercenaria mercenaria* (Baker et al. 2008), American lobster *Homarus americanus* (Kenchington et al. 2009), the marine gastropod *Acanthinucella spirata* (Hellberg et al. 2001) and the dog whelk *Nucella ostrina* (Marko 2004), appear to conform to the model proposed by Hewitt (2000). However, it has been proposed that several marine species instead had northern peri-

glacial refugia, on the basis of high genetic variability in contemporary northern populations. These include the ocean quahog *Arctica islandica* (Dahlgren et al. 2000), *Macoma balthica* (Luttikhuizen et al. 2003), *Cerastoderma edule* (Krakau et al. 2012), *Pagurus longicarpus* (Young et al. 2002) and *Littorina saxatilis* (Panova et al. 2011). In addition, peri-glacial refugia have been suggested for many other marine species including *Carcinus maenas* (Roman & Palumbi 2004), some seaweed species (Hoarau et al. 2007, Olsen et al. 2010, Coyer et al. 2011), the common shrimp *Crangon crangon* (Luttikhuizen et al. 2008), the genus *Patella* (Sa-Pinto et al. 2005) and tubiferous polychaetes (Jolly et al. 2006).

It is interesting to note that the molecular data produced by St-Onge et al. (2013) suggest a southern refugium for *M. arenaria* in eastern North America. In the present study, it was not feasible to determine whether there has also been a limited contribution from 1 or more refugia on European shores, if the species did persist in Europe during the Pleistocene. While some haplotypes were almost or totally exclusive to Europe (see Irish samples in Table 3), the majority were only 1 mutational step removed from haplotypes common in eastern North America (see Fig. 2). For microsatellites, only a minority of rare alleles was exclusive to European samples. More extensive European sampling, particularly farther south in the range, will be necessary to address the question of possible refugia.

Genetic variability may indicate population size at introduction

The level of genetic variability at the microsatellite loci, observed in the Pacific North American sample, is informative in the context of size of introductions. In this case, there is documented evidence of recent human introduction from the southern part of the eastern North American range, when a large number of *M. arenaria* were introduced to areas through the commercial oyster industry, and dedicated plantings were made to initiate a commercial fishery (Hanna 1966, Carlton, 1989, Strasser 1999). The sample analysed in the present study from Oregon on the Pacific coast had a level of microsatellite variability similar to that from Maryland on the Atlantic coast, giving genetic confirmation of a large introduction from an eastern North American area. In contrast, the European clam samples showed substantially lower microsatellite variability, suggesting a lower number of individuals introduced to this area (see Table 1).

The large inter- and intra-continental difference demonstrated in the current study may be indicative of intermittent movements of small numbers of individuals from North America to Europe. The results from the analysis of *COX1* sequences are less clear than those of microsatellites, which provided better support for discrimination among samples within continents in comparison to the single mitochondrial region. Interestingly, however, both mitochondrial DNA and microsatellites provide almost equal support to the inter-region (continent) differences.

What is clearly demonstrated in our results is that the use of different marker types provides more detailed insights into the mode of colonisation and contemporary population structure. We therefore suggest that a genomic approach be applied to the investigation of *M. arenaria* in the future, perhaps by identifying and screening a large number of single nucleotide polymorphisms (Beaumont & Balding 2004). This, in addition to modern statistical approaches based on Bayesian hypothesis inference (e.g. Cornuet et al. 2014), may further clarify the ecology and previous movements of what is now an ecologically important and abundant soft-sediment infaunal species in northwestern Europe. It would also provide genomic information that would be of use in *M. arenaria* aquaculture, which is widespread and commercially important in North America, and is to be developed in Europe.

Acknowledgements. This project was partly funded by the Ireland Wales Programme 2007–2013 INTERREG 4A European Regional Development Fund. Many thanks to Chris Dungan, John Chapman, Dante Mateo, Pauline Kamermans, Ruth Stockdale, Allen Whitaker, Eileen Dillane and Jens and Nettan Carlsson for providing *Mya arenaria* samples and practical help. C.R.B. and P.M. were supported by the Beaufort Marine Research Award in Fish Population Genetics funded by the Irish Government under the Sea Change Programme.

LITERATURE CITED

Allendorf FW, Luikart G (2008) Conservation and the genetics of populations. Blackwell, Malden, MA

Arias-Pérez A, Fernández-Tajes J, Gaspar MB, Méndez J (2012) Isolation of microsatellite markers and analysis of genetic diversity among East Atlantic populations of the sword razor shell *Ensis siliqua*: a tool for population management. Biochem Genet 50:397–415

Baker P, Austin JD, Bowen BW, Baker SM (2008) Range-wide population structure and history of the northern quahog (*Mercenaria mercenaria*) inferred from mitochondrial DNA sequence data. ICES J Mar Sci 65:155–163

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. Mol Ecol 13:969–980

Beebee TJC, Rowe G (2008) An introduction to molecular ecology. Oxford University Press, London

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier

Blakeslee AMH, Byres JE (2008) Using parasites to inform ecological history: comparisons among three congeneric marine snails. Ecology 89:1068–1078

Blakeslee AMH, Byers JE, Lesser MP (2008) Solving cryptogenic histories using host and parasite molecular genetics: the resolution of *Littorina littorea's* North American origin. Mol Ecol 17:3684–3696

Bologa AS, Bodeanu N, Petran A, Tiganus V, Zaitsev YP (1995) Major modifications of the Black Sea benthic and planktonic biota in the last three decades. Bull Inst Océanogr (Monaco) Special 15:85–110

Brawley S, Coyer J, Blakeslee A, Hoarau G and others (2009) Historical invasions of the intertidal zone of Atlantic North America associated with distinctive patterns of trade and emigration. Proc Natl Acad Sci USA 106:8239–8344

Briski E, Ghabooli S, Bailey SA, MacIsaac HJ (2012) Invasion risk posed by macroinvertebrates transported in ships' ballast tanks. Biol Invasions 14:1843–1850

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20:1004–1006

Caporale DA, Beal BF, Roxby R, Van Beneden RJ (1997) Population structure of *Mya arenaria* along the New England coastline. Mol Mar Biol Biotechnol 6:33–39

Carlton JT (1989) Man's role in changing the face of the ocean: biological invasions and implications for conservation of near-shore environments. Conserv Biol 3: 265–273

Chapman J (2007) The mysterious periwinkle (*Littorina littorea*): Norse or later introduction from Europe? Arct Stud Centre Newsl 14:22

Chapman JW, Carlton JT, Bellinger MR, Blakeslee AMH (2007) Premature refutation of a human-mediated marine species introduction: the case history of the marine snail *Littorina littorea* in the northwestern Atlantic. Biol Invasions 9:737–750

Chapman JW, Blakeslee AMH, Carlton JT, Bellinger MR (2008) Parsimony dictates a human introduction: on the use of genetic and other data to distinguish between the natural and human-mediated invasion of the European snail *Littorina littorea* in North America. Biol Invasions 10:131–133

Conde A, Novais J, Domínguez J (2010) Southern limit of distribution of the soft-shell clam *Mya arenaria* on the Atlantic East Coast. Biol Invasions 12:429–432

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A and others (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics 30: 1187–1189

Coyer JA, Hoarau G, Van Schaik J, Luijckx P, Olsen JL (2011) Trans-Pacific and trans-Arctic pathways of the intertidal macroalga *Fucus distichus* L. reveal multiple glacial refugia and colonizations from the North Pacific to the North Atlantic. J Biogeogr 38:756–771

Crocetta F, Turolla E (2011) *Mya arenaria* Linné, 1758 (Mollusca: Bivalvia) in the Mediterranean Sea: its distribution revisited. J Biol Res (Thessalon) 16:188–193

Dahlgren TG, Weinberg JR, Halanych KM (2000) Phylogeography of the ocean quahog (*Arctica islandica*): influences of paleoclimate on genetic diversity and species range. Mar Biol 137:487–495

Diniz FM, Iyengar A, Sarmanho da Costa Lima P, Maclean N, Bentzen P (2007) Application of a double-enrichment procedure for microsatellite isolation and the use of tailed primers for high throughput genotyping. Genet Mol Biol 30:380–384

Dupont L, Ellien C, Viard F (2007) Limits to gene flow in the slipper limpet *Crepidula fornicata* as revealed by microsatellite data and a larval dispersal model. Mar Ecol Prog Ser 349:125–138

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

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hanna GD (1966) Introduced mollusks of western North America. Occas Pap Calif Acad Sci 48:1–108

Harley CDG, Anderson KM, Lebreton CAM, MacKay A and others (2013) The introduction of *Littorina littorea* to British Columbia, Canada: potential impacts and the importance of biotic resistance by native predators. Mar Biol 160:1529–1541

Hellberg ME, Balch DP, Roy K (2001) Climate-driven range expansion and morphological evolution in a marine gastropod. Science 292:1707–1710

Hewitt GM (1999) Post-glacial re-colonisation of European biota. Biol J Linn Soc 68:87–112

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405:907–913

Hoarau G, Coyer JA, Veldsink JH, Stam WT, Olsen JL (2007) Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus.* Mol Ecol 16:3606–3616

Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a crossplatform computer program that plans and optimizes multiplex PCR. Biotechniques 46:511–517

Jakobsson M, Rosenberg NA (2017) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Jolly MT, Viard F, Gentil F, Thiébaut E, Jollivet D (2006) Comparative phylogeography of two coastal polychaete tubeworms in the Northeast Atlantic supports shared history and vicariant events. Mol Ecol 15:1841–1855

Kalinowski ST (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling designs. Conserv Genet 5:539–543

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes 5:187–189

Keenan K, McGinnity P, Cross TF, Crozier WW, Prodohl PA (2013) diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 4:782–788

Kenchington EL, Harding EL, Jones MW, Prodohl PA (2009) Pleistocene glaciation events shape genetic structure

across the range of the American lobster, *Homarus americanus.* Mol Ecol 18:1654–1667

Krakau M, Jacobsen S, Jensen KT, Reise K (2012) The cockle *Cerastoderma edule* at Northeast Atlantic shores: genetic signatures of glacial refugia. Mar Biol 159: 221–230

Krapal AM, Popa OP, Iorgu EI, Costache M, Popa LO (2012) Isolation and characterization of new microsatellite markers for the invasive softshell clam, *Mya arenaria* (L.) (Bivalvia: Myidae). Int J Mol Sci 13:2515–2520

Lasota R, Hummel H, Wolowicz M (2004) Genetic diversity of European populations of the invasive soft-shell clam *Mya arenaria* (Bivalvia). J Mar Biol Assoc UK 84: 1051–1056

Levinton J (1973) Genetic variation in a gradient of environmental variability: marine bivalvia (Mollusca). Science 180:75–76

Luttikhuizen PC, Drent J, Van Delden W, Piersma T (2003) Spatially structured genetic variation in a broadcast spawning bivalve: quantitative vs molecular traits. J Evol Biol 16:260–272

Luttikhuizen PC, Campos J, van Bleijswijk J, Peijnenburg KT, van der Veer HW (2008) Phylogeography of the common shrimp, *Crangon crangon* (L.) across its distribution range. Mol Phylogenet Evol 46:1015–1030

MacNeil FS (1965) Evolution and distribution of the genus *Mya*, and tertiary migrants of mollusca. Geological Survey Professional Paper 483-G. US Government Printing Office, Washington, DC

Maggs CA, Castilho R, Foltz D, Henzler C and others (2008) Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. Ecology 89(Suppl):S108–S122

Marko PB (2004) 'What's larvae got to do with it?' Disparate patterns of post-glacial population structure in two benthic marine gastropods with identical dispersal potential. Mol Ecol 13:597–611

Morgan RP, Block SB, Ulanowicz NI, Buys C (1978) Genetic variation in the soft-shelled clam, *Mya arenaria.* Estuaries 1:255–258

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. J Mol Evol 19: 153–170

Olsen JL, Zechman FW, Hoarau G, Coyer JA, Stam WT, Valero M, Aberg P (2010) The phylogeographic architecture of the fucoid seaweed *Ascophyllum nodosum*: an intertidal 'marine tree' and survivor of more than one glacial-interglacial cycle. J Biogeogr 37:842–856

Panova M, Blakeslee AMH, Miller AW, Mäkinen T, Ruiz GM, Johannesson K, André C (2011) Glacial history of the North Atlantic marine snail, *Littorina saxatilis*, inferred from distribution of mitochondrial DNA lineages. PLoS One 6:e17511

Petersen KS, Rasmussen KL, Heinemeier J, Rud N (1992) Clams before Columbus? Nature 359:679

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

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

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Roman J, Palumbi SR (2004) A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe. Mol Ecol 13:2891–2898

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365–386

Ruiz GM, Carlton JM, Grosholz ED, Hines AH (1997) Global invasions of marine and estuarine habitats by non-indigenous species: mechanisms, extent, and consequences. Am Zool 37:621–632

Sa-Pinto A, Branco M, Harris DJ, Alexandrino P (2005) Phylogeny and phylogeography of the genus *Patella* based on mitochondrial DNA sequence data. J Exp Mar Biol Ecol 325:95–110

Sparagano OAE, Purdom I, Priest FG, Kingston PF (2002) Molecular characterization of the bivalves *Mya arenaria, Mya truncata* and *Hiatella arctica.* J Molluscan Stud 68: 190–191

St-Onge P, Parent E, Sévigny JM, Tremblay R, Rioux-Paré R (2011) Isolation and characterization of eight novel microsatellite markers for the softshell clam (*Mya arenaria*). Mol Ecol Resour Primer Database: mer-11-0122

St-Onge P, Sevigny JM, Strasser C, Tremblay R (2013) Strong population differentiation of softshell clams (*Mya arenaria*) sampled across seven biogeographic marine ecoregions: possible selection and isolation by distance. Mar Biol 160:1065–1081

Strasser M (1999) *Mya arenaria* — an ancient invader of the North Sea coast. Helgol Mar Res 52:309–324

Strasser C, Barber P (2009) Limited genetic variation and structure in softshell clams (*Mya arenaria*) across their native and introduced range. Conserv Genet 10:803–814

Strauch F (1972) Phylogenese, Adaptation und Migration einiger nordischer mariner Molluskengenera (*Neptunea, Panomya, Cyrtodaria* und *Mya*). Abh Senckenb Natforsch Ges 531:1–211

Tozaki T, Mashima S, Hirota K, Miura N, Choi-Miura NH, Tomita M (2001) Characterization of equine microsatellites and microsatellite-linked repetitive elements (eMLREs) by efficient cloning and genotyping methods. DNA Res 8:33–45

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA (2007) Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 35(Suppl 2):W71–W74

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Microchecker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

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

Weston S, Buttner JK (2010) Softshell clam culture: basic biology and general culture considerations. NRAC Publication 201. Northeastern Regional Aquaculture Center, University of Maryland, College Park, MD

Young A, Torres C, Mack J, Cunningham C (2002) Morphological and genetic evidence for vicariance and refugium in Atlantic and Gulf of Mexico populations of the hermit crab *Pagurus longicarpus.* Mar Biol 140:1059–1066

Zaitsev Y, Mamaev V (1997) Marine biological diversity in the Black Sea: a study of change and decline. United Nations Publications, New York, NY

Editorial responsibility: Philippe Borsa, Montpellier, France

Submitted: August 8, 2014; Accepted: February 25, 2016
Proofs received from author(s): April 21, 2016