Journal of Biogeography (J. Biogeogr.) (2009) **36**, 39–54

ORIGINAL ARTICLE

Phylogeography of an estuarine mysid, *Neomysis integer* (Crustacea, Mysida), along the north-east Atlantic coasts

Thomas Remerie[1,3]*, Andy Vierstraete[2,3], Peter H. H. Weekers[2,3], Jacques R. Vanfleteren[2,3] and Ann Vanreusel[1,3]

[1]*Marine Biology Section, Biology Department, Ghent University, Gent, Belgium,* [2]*Biology Department, Ghent University, Gent, Belgium and* [3]*CeMoFE, Centrum voor Moleculaire Fylogenie en Evolutie, Ghent University, Gent, Belgium*

*Correspondence: Thomas Remerie, Marine Biology Section, Biology Department, Ghent University, Krijgslaan 281-S8, 9000 Gent, Belgium.
E-mail: thomas.remerie@ugent.be

ABSTRACT

Aim The brackish water mysid, *Neomysis integer*, is one of the most common mysid species along the coasts of the north-east Atlantic. In the present study, the phylogeographical patterns were examined throughout the distribution range of *N. integer*. In particular, the latitudinal trends in genetic diversity and the distribution of genetic variation were examined in order to elucidate the imprints of the Pleistocene glaciations.

Location North-east Atlantic coasts from the Baltic Sea to the south of Spain.

Methods A total of 461 specimens from 11 populations were analysed by means of single-stranded conformation polymorphism analysis combined with DNA sequencing of a fragment of the mitochondrial cytochrome *c* oxidase I gene. The genetic structure was examined by using a progression of phylogenetic, demographic and population genetic analyses to elucidate not only the geographical structure, but also the evolutionary history producing that structure.

Results The levels of genetic diversity were relatively uniform throughout the distribution range, with the exception of a decline at the northern and southern edges of distribution. A high heterogeneity was observed between the populations analysed (global $\Phi_{ST} = 0.787$). This is caused by the disparate distribution of the cytochrome oxidase I haplotypes, with several population-specific haplotypes. A clear genetic break (2.4% sequence divergence) occurred between the southernmost Guadalquivir population and all other populations.

Main conclusions The present study corroborates the expectations of the genetic patterns typically observed in an estuarine species. The within-population variability was low, whereas a significant (moderate to high) divergence was observed between populations. Phylogeographical analysis revealed that northern populations within the English Channel, North Sea and Baltic Sea are characterized by several widespread haplotypes, while the Irish population and all sites south of the Bay of Biscay consist solely of unique haplotypes. This pattern, combined with the relative high levels of genetic diversity, could be indicative for the presence of a glacial refugium in the English Channel region. Under this scenario *N. integer* must have survived the Last Glacial Maximum in the palaeoriver system present in that region.

Keywords
Brackish water, COI, Crustacea, estuary, glacial refugium, mtDNA, mysid, north-east Atlantic Ocean, phylogeography.

INTRODUCTION

In contrast to terrestrial and freshwater systems, interest in phylogeographical studies of marine and estuarine taxa in Europe has grown only recently (Wilke & Pfenninger, 2002; Coyer *et al.*, 2003; Luttikhuizen *et al.*, 2003; Gysels *et al.*, 2004a,b; Olsen *et al.*, 2004). Environmental perturbations and the transformation of the geography of northern Europe

doi:10.1111/j.1365-2699.2008.01970.x

during the Pleistocene glaciations are thought to have had a major impact on the phylogeographical patterns in extant species. Range compression and expansion caused by glacial events were the most conspicuous biogeographical features (Avise *et al.*, 1998; Taberlet *et al.*, 1998; Hewitt, 2000). In addition to these historical changes, current population dynamics (such as contemporary gene flow), which are related to specific life-history traits (such as dispersal capacity, existence or absence of pelagic larvae), might also affect the distribution of genetic variation. Consequently, species of low dispersal capability are expected to bear more signatures of the Pleistocene glaciations in their genetic composition due to reduced gene flow and increased population differentiation, while it is not unlikely that historical patterns have been erased in species of high gene flow.

Data on terrestrial and freshwater biota provide convincing evidence that the southern European regions, in particular the Iberian and Italian peninsulas, and Balkan and Pontic (Black Sea) regions, served as refugia during glacial events, harbouring the greatest amount of genetic diversity (Taberlet *et al.*, 1998; Hewitt, 1999). Post-glacial range expansion of limited and/or genetically homogeneous numbers of colonists, possibly in combination with bottlenecks, is probably the cause of the usually lower degree of diversity at higher latitudes (Hewitt, 1996, 2000, 2001; Ibrahim *et al.*, 1996). However, the picture in marine species is less clear. Some marine taxa, including gobies, copepods and seaweed (Edmands, 2001; Coyer *et al.*, 2003; Gysels *et al.*, 2004a), show a correlation between higher latitude and reduced diversity, while others do not (Marko, 2004; Olsen *et al.*, 2004). In addition, contemporaneous populations of several species in refugial regions (such as the Iberian Peninsula) might be impoverished due to the post-glacial warming starting from *c.* 11.5 ka. This event may have constituted a strong selective force for refugial populations resulting in a (southward) decline in genetic diversity (Dahlgren *et al.*, 2000; Coyer *et al.*, 2003).

The present distribution of molecular variation in many Western European taxa can be explained by a northward dispersal from a southern Iberian refugium (Taberlet *et al.*, 1998; Hewitt, 1999). Although this pattern has been confirmed for several marine taxa along the north-eastern Atlantic (Garcia-Marin *et al.*, 1999; Consuegra *et al.*, 2002), other studies provide evidence for additional 'northern' glacial refugia (Verspoor *et al.*, 1999; Breton *et al.*, 2003; Coyer *et al.*, 2003; Luttikhuizen *et al.*, 2003; Gysels *et al.*, 2004a; Jolly *et al.*, 2005; Provan *et al.*, 2005). Hence, populations of these species should have survived range compression during glacial periods and post-glacial expansions in unglaciated areas.

In the present study, we examined the phylogeographical structure of the brackish water mysid *Neomysis integer* (Leach, 1814) along the north-east Atlantic coasts. *N. integer* is one of the most common mysids around the north-eastern Atlantic coasts (from the Baltic Sea to the southern Iberian coasts), and it is believed to be a keystone species in the marine ecosystems of these regions (Mees *et al.*, 1994, 1995; Fockedey & Mees, 1999; Hostens & Mees, 1999). It is a euryhaline and eurythermic species that dominates the hyperbenthic fauna of brackish water environments in western European estuaries (Mees *et al.*, 1995; Cunha *et al.*, 1999). Apart from evidence of migration over small geographical scales, as vertical dial migrations, tidal migration and seasonal migrations within an estuary (Mauchline, 1980; Mees *et al.*, 1993), nothing is known about the dispersal capacities of *N. integer* over larger scales, such as between neighbouring estuaries or even over larger distances. However, its life-history traits might suggest that long-range dispersal events are rare. Like all mysids, *N. integer* is a brooder, and hence lacks a planktonic dispersal stage so that the actual dispersal should take place only through the movement of juveniles or adults. In addition, *N. integer* lives in discrete brackish water, estuarine habitats and is rarely encountered in marine waters (Mauchline, 1971). This suggests that *N. integer* could be a promising candidate for elucidating the phylogeographical patterns of estuarine invertebrates along the north-east Atlantic coasts. A pilot study on the genetic differentiation of *N. integer* throughout its distribution range based on a limited number of cytochrome *b* mtDNA sequences supports the hypothesis of limited gene flow resulting in the genetic differentiation of populations (Remerie *et al.*, 2006). However, the limited number of individuals analysed in this study precluded the authors from drawing robust phylogeographical conclusions.

The present study was designed to explore in more detail the phylogeographical structure and patterns of molecular diversity and contemporary gene flow throughout the whole distribution range of *N. integer*. In particular, we hypothesize that (1) the levels of genetic diversity throughout the distribution range of *N. integer* vary latitudinally, with a northwards decline in formerly glaciated regions; (2) historical processes, such as glacial vicariant events and post-glacial expansions, are responsible for deep levels of population structuring and the current distribution of mitochondrial haplotypes along the distribution range of *N. integer*; and (3) the levels of genetic exchange that currently take place among European populations of *N. integer* are restricted because of its assumed low dispersal capacities and discrete estuarine habitat, resulting in a significant divergence among populations. The genetic structure of *N. integer* was examined using a progression of phylogenetic, demographic and population genetic analyses of mitochondrial cytochrome oxidase I (COI) sequences. Such an approach has proven useful in elucidating not only geographical structure, but also the evolutionary history producing that structure (Althoff & Pellmyr, 2002).

MATERIALS AND METHODS

Sampling

A total of 461 specimens were collected from 11 locations comprising eight estuaries, one coastal site (Tvärminne, TV), one oligo-saline lagoon (Kilkieran Lake, KILK) and one estuary–coastal lagoon system (Ria de Aveiro, RdA) (Fig. 1). This sampling scheme covers most of the current distribution

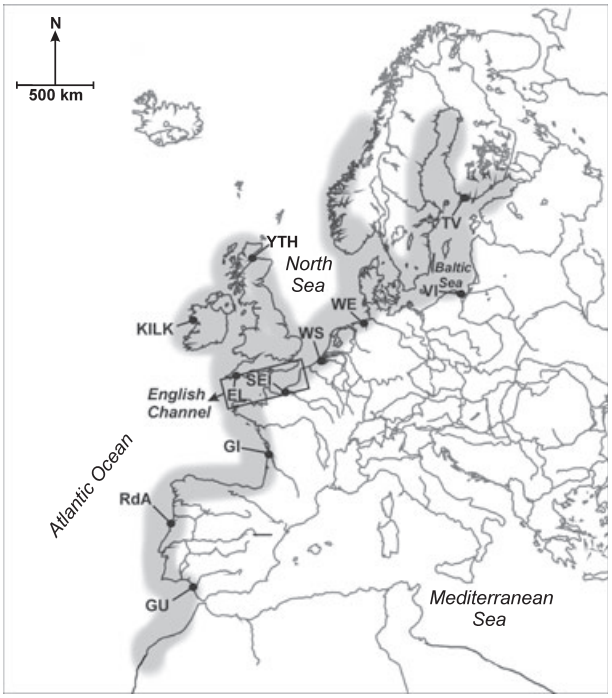

Figure 1 Geographical location of the sampling sites of *Neomysis integer*. The shaded area represents the distribution range of *N. integer*. For details on the sampling locations and abbreviations see Table 1.

range of *N. integer*. Samples from each site were collected with a hand net or a hyperbenthic sledge (mesh size 1 mm) and collections were made between 1999 and 2002 (Table 1). After collection, samples were stored in ethanol (70–95%) or acetone (Fukatsu, 1999) at 4°C.

DNA extraction, PCR, single-stranded conformation analysis and sequencing

DNA was extracted using a modified cetyl trimethyl ammonium bromide (CTAB) protocol (Kocher *et al.*, 1989). Mysid tissue of single individuals was crushed using a beadbeater and afterwards incubated for a minimum of 3 h at 60°C in 500 μL CTAB buffer [2% (w/v) CTAB, 1.4 M NaCl, 0.2% (v/v) mercaptoethanol, 20 mM EDTA, 100 mM Tris/HCl pH 8] with 6 μL proteinase K (1 mg per 100 μL). After overnight incubation at 37°C, DNA was extracted with phenol/chloroform/isoamylalcohol (25 : 24 : 1 pH 8) and chloroform: isoamylalcohol (24 : 1). Finally, DNA was isopropanol-precipitated and rehydrated in 25 μL sterile water. A 651 bp fragment of the COI gene was amplified using the universal primers LCO1490 and HCO2198 (Folmer *et al.*, 1994). The conditions for the COI amplifications with a reaction volume of 25 μL were: 10× PCR buffer with $(NH_4)SO_4$ included (MBI Fermentas, St Leon-Rot, Germany), 2 mM $MgCl_2$, 0.2 mM dNTP, 1 μM forward and reverse primers and 1.25 units *Taq* polymerase. The following thermocycle profile was used: denaturation of template DNA at 94°C for 2 min, followed by a touchdown PCR (annealing temperature decrease of 1°C per cycle) of four cycles (30 s at 94°C, annealing at 59°C for 50 s, extension at 72°C for 90 s), followed by 40 cycles of 30 s at 94°C, 50 s at 55°C and 2 min at 72°C, followed by a final extension of 5 min at 72°C. PCR products were purified with exonuclease I (10 U μL^{-1}; Amersham Biosciences, Amersham, UK) and shrimp alkaline phosphatase (1 U μL^{-1}; Amersham Biosciences). Purified products were cycle-sequenced in both directions using BigDye Terminator Mix (PE Applied Biosystems, Foster City, CA, USA) and the following conditions: 25 cycles of 96°C for 30 s, 50°C for 15 s and 60°C for 4 min. Cycle-sequencing products were electrophoresed on a Perkin-Elmer (Foster City, CA, USA) ABI Prism 377 DNA sequencer.

For the single-strand conformation polymorphism (SSCP) analysis (Orita *et al.*, 1989), two sets of internal primers were designed within the 651 bp COI fragment, generating two COI fragments of size < 250 bp. This was done to ensure a high mutation detection resolution of the SSCP technique within the COI fragment. The sensitivity of SSCP is generally inversely proportional to the size of the fragment; single base-pair differences are resolved 99% of the time for a 100 to 300 bp fragment, only 80% for 400 bp fragments (Sunnucks *et al.*, 2000). The position of both fragments within the COI gene was chosen based on the variability observed in a small pilot study

Table 1 Geographical location, sampling date, sample size (N), number of haplotypes (N_h), diversity measures (h, haplotype diversity; π, nucleotide diversity) and percentage of private haplotypes (%pH) for the 11 samples of *Neomysis integer*.

Sampling location	Code	Latitude	Longitude	Date	N	N_h	h (SD)	π (SD)	%pH
Tvärminne	TV	59°51′ N	23°12′ E	Jun 1999	41	3	0.0963 (0.0624)	0.00065 (0.00078)	33.3
Vistula	VI	54°21′ N	18°56′ E	Feb 2001	41	1	0	0	0
Weser	WE	53°25′ N	08°30′ E	Aug 2000	39	3	0.5263 (0.0688)	0.00534 (0.00330)	0
Ythan	YTH	57°18′ N	02°00′ W	Aug 2000	39	4	0.5803 (0.0430)	0.00147 (0.00129)	75
Westerschelde	WS	51°25′ N	04°00′ E	Feb 2001	60	6	0.4689 (0.0652)	0.00335 (0.00227)	50
East Looe	EL	50°24′ N	04°26′ W	Sep 2002	36	5	0.3048 (0.0970)	0.00194 (0.00155)	80
Kilkieran Lake	KILK	53°20′ N	09°44′ W	Oct 2002	43	5	0.2957 (0.0875)	0.00078 (0.00087)	100
Seine	SEI	48°26′ N	00°10′ E	May 2001	48	4	0.4193 (0.0810)	0.00329 (0.00225)	25
Gironde	GI	45°33′ N	00°55′ E	May 2001	44	3	0.4894 (0.0500)	0.00472 (0.00298)	100
Ria de Aveiro	RdA	40°41′ N	08°45′ W	May 2001	30	5	0.6115 (0.0510)	0.00272 (0.00198)	100
Guadalquivir	GU	36°55′ N	06°17′ W	May 2001	40	5	0.2359 (0.0880)	0.00128 (0.00118)	100

of 10 COI sequences of 651 bp from different sampling sites. The amplification of the two COI fragments (COI part 1 and COI part 2, 215 and 232 bp, respectively) used the primer sequences: LCO1490 (Folmer *et al.*, 1994) and COIR3 (5′-GAG GGA AAG CTA TAT CTG GAG C-3′), COIF2 (5′-TTT AGC AGG GGC TTC CTC TA-3′) and HCO2198 (Folmer *et al.*, 1994). Conditions for PCR were as described previously, but with an annealing temperature of 56 instead of 55°C. SSCP analysis were performed using 0.5-mm-thick non-denaturing polyacrylamide gels (250 × 110 mm) (total acrylamide concentration, T = 12.5%; degree of cross-linking, C = 2%). Ideal running conditions for the SSCP analysis of both COI fragments were assessed by using different conditions (electrophoresis temperature and gel composition) and comparing the banding patterns (SSCP phenotypes) of all gels. Electrophoresis at a constant power of 8 W at 5°C for the COI part 1 fragment and at 12°C for COI part 2 for 3.5 h using polyacrylamide gels with T = 12.5% and C = 2% gave the best resolution. Bands were visualized with a DNA silver-staining kit (Amersham Biosciences) and scored by their relative mobility. The full COI fragment (651 bp), including the two COI fragments used in the SSCP analysis, of individuals from each mobility class was sequenced in both directions using the LCO1490 and HCO2198 primers (Folmer *et al.*, 1994) under the conditions described previously. In order to asses the reliability of SSCP for detecting nucleotide substitutions, at least two individuals of each mobility class were sequenced, with the exception of haplotypes found in only one individual. A total of 71 individuals were sequenced after SSCP analysis (> 15% of all individuals screened with SSCP). Sequences were submitted to Genbank under accession nos AJ852563–AJ852596.

Sequence alignment

Alignment of the sequence data was produced with the ClustalX program (ver. 1.74; Thompson *et al.*, 1997). All polymorphisms were double-checked on the chromatograms. The 204 bp between the two COI SSCP fragments (COI part 1 and part 2), not screened with SSCP, were excluded from the 651 bp alignment block, resulting in a combined alignment of 447 bp (first 228 bp = COI part 1, next 219 bp = COI part 2) which was used in the further phylogeographical and population genetic data analysis.

Population genetic and phylogeographical analysis

Nucleotide diversity (π, the mean number of differences between all pairs of haplotypes) and haplotype diversity (h, the probability that two randomly chosen haplotypes are different in a sample) and its standard deviation (SD) were calculated for each population using the program ARLEQUIN ver. 2.0 (Schneider *et al.*, 2000).

The geographical differentiation of haplotypes was quantified using a hierarchical analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) using ARLEQUIN 2.0 (Schneider *et al.*, 2000). The significance of variance components and Φ-statistic analogues was tested by multiple (1000) random permutations. Pairwise Φ_{ST} values were calculated based on Tamura & Nei (1993) genetic distances using the gamma value obtained in MODELTEST ver. 3.06 (Posada & Crandall, 1998). Their significance was tested by multiple (1000) random permutations. When necessary, corrections for multiple tests were applied using sequential Bonferroni correction (Rice, 1989). The initial AMOVA did not assign populations to groups, while in the second AMOVA localities were grouped into four regions to test the relative variation among populations relative to geographical regions.

An isolation-by-distance (IBD) test was conducted using a reduced major axis regression and Mantel permutation tests of genetic distances between all sampling locations (Φ_{ST}) against minimum coastline distance between all pairs of sampling sites. This linear–linear regression was chosen as it produced the highest r^2 value. The strength and statistical significance of associations were tested using the program IBD ver. 1.52 (Bohonak, 2002).

Prior to estimating divergence times, the hypothesis of equal substitution rates among lineages was tested using a likelihood ratio test in PAUP* ver. 4.0b8 (Swofford, 1998). As the molecular clock hypothesis was not significantly rejected (P = 0.0626), the time of divergence of population pairs (T) was estimated based on the mean nucleotide divergence between populations corrected for within-group variation (Nei, 1987; equation 10.21) (net nucleotide divergence corrected for ancestral polymorphisms). Because there is no fossil record and no geological or climatic event that might be useful in calibrating a clock of mtDNA divergence specifically for mysids, we used a general molecular clock for crustacean COI mtDNA of 1.4–2.6% of nucleotide divergence Myr^{-1}. These estimates of mutation rates were adopted from several calibrations for crustacean taxa thought to have been divided by the Isthmus of Panama (e.g. snapping shrimp, *Alpheus* sp., Knowlton *et al.*, 1993) and crab species thought to have been subdivided since the trans-Arctic interchange (*Sesarama* sp., Schubart *et al.*, 1998).

A nested clade analysis (NCA, Templeton *et al.*, 1995; Templeton, 1998, 2004) was performed to test for associations between haplotypes and geography. It aims at separating patterns of population history and gene flow. A haplotype network was constructed to visualize evolution among haplotypes, haplotype frequency and geographical representation. This haplotype network was created using a parsimony criterion in the program TCS ver. 1.18 (Clement *et al.*, 2000) with a connection limit set at 95%. Ambiguities and loops in the haplotype network were resolved following the criteria suggested by Crandall & Templeton (1993). The TCS haplotype network was nested into clades using the nesting rules given by Templeton *et al.* (1987) and Crandall (1996). Subsequently, an exact permutational contingency test was conducted for each clade and a χ^2-statistic was calculated from the contingency tables (clades vs. geographical locations) by treating sample locations as categorical variables. The statistical significance of the clade distance (D_c), which measures how geographically

widespread the individuals from a clade are, and nested clade distance (D_n), which measures the distance of a clade from the geographical centre of their nesting category, was calculated by comparison with a null distribution (no geographical association of clades and clade dispersal distances are not significantly different from random) derived from 10,000 random permutations of clades against sampling locations using the program GEODIS ver. 2.2 (Posada *et al.*, 2000). The interpretation of the observed distance patterns was conducted using Templeton's (2004) inference key.

Historical population dynamics, such as demographic expansions, were estimated by computing the distribution of pairwise differences (mismatch distribution) and its parameters using ARLEQUIN ver. 2.0 (Schneider *et al.*, 2000). Mismatch distributions and Rogers' (1995) parameters of mismatch distribution (τ, θ_0, θ_1) were assessed using Monte Carlo simulations of 1000 random samples. Additionally, Tajima's *D* (Tajima, 1989) and Fu's F_s (Fu, 1997) neutrality tests were used to infer the nature of sequence evolution (e.g. rapid selection or neutral) and probable historical population movements. Significant negative values of Tajima's *D* are expected to occur in cases of recent population expansion (Slatkin & Hudson, 1991; Fu, 1997; Knowles *et al.*, 1999) or a selective sweep (Maruyama & Birky, 1991; Fu, 1997; Filatov *et al.*, 2000). Whereas significant negative Fu's F_s values are indicative of an excess of rare alleles, which might be caused by a recent population expansion (Fu, 1997), positive Tajima's *D* and Fu's F_s values indicate balancing selection or population structure.

RESULTS

SSCP and sequencing results

The SSCP technique distinguished 19 haplotypes within the COI-1 fragment and 20 haplotypes within the COI-2 fragment. The combined information led to the identification of 34 haplotypes among the 461 specimens analysed from 11 samples. The reliability of the SSCP was high, as the sequence analysis showed little additional variation not detected by SSCP (see Appendix S1 in Supporting Information). DNA sequencing detected a total of 35 polymorphic positions (7.8%) among the 34 different haplotypes. Most polymorphisms were due to single nucleotide changes, and constitute 29 transitions and seven transversions; 20 characters were parsimony-informative. Eight nucleotide changes resulted in an amino acid change. Pairwise DNA differences between haplotypes ranged from 0.22% (a single substitution) to 2.68% (12 base substitutions) nucleotide divergence.

Seventeen unique haplotypes (50%) were observed (Tables 1 and 2). The haplotypes strongly segregated with geographical origin. Only four haplotypes (which were present in 53% of all individuals analysed) were shared between various sampling sites, while the remaining 13 haplotypes were population-specific. Interestingly, the shared haplotypes were observed only in sampling locations north of the English Channel, with the exception of the Irish population (KILK), which possessed only private haplotypes. The most common haplotype (Df) was observed in five different locations and was present in 29.2% of individuals analysed. All sampling sites except two (VI and WE) possessed private haplotypes, and four locations (KILK, GI, RdA and GU) consisted solely of private haplotypes.

Intraspecific evolution

Mean nucleotide diversity (π) among all locations was 0.002323, ranging from 0 to 0.00534 (Table 1). Lowest levels of nucleotide diversity were observed in the Baltic Sea (VI and TV) and in the Irish population (KILK). Highest levels were observed in the Weser Estuary (WE). Values of haplotype diversity (*h*) ranged from 0 to 0.6115 and had an average of 0.3662. The lowest *h* values were observed within the Baltic Sea (TV and VI), while haplotype diversity in the Portuguese sample of Aveiro (RdA) and the North Sea samples from the Ythan (YTH) and Weser (WE) were almost twice the average. No significant correlation was observed between haplotype diversity and latitude ($P = 0.49$) and between nucleotide diversity and latitude ($P = 0.55$).

The parsimony network is shown in Fig. 2. Two loops were excluded from the network: (1) a connection of the singleton haplotype Je with the frequent, central haplotype Ee was preferred rather than a central position of this haplotype; and (2) a connection of both singletons Ra and Rh with the haplotype Ha from the same population was favoured. The centre of the network consisted of haplotypes observed in the North Sea (YTH) and English Channel (EL and SEI). Most haplotypes were relatively closely related to each other, with the exception of all haplotypes observed in the Guadalquivir population (GU) (Cc, Kc, Lc, Mp, Cq). The haplotypes formed a divergent subgroup, separated by at least six mutational steps from the central haplotypes (1.3% of uncorrected genetic divergence). Likewise, the haplotypes from the Gironde population (GI) were more divergent from the central haplotypes (0.5–0.9% of uncorrected genetic divergence). Other subgroups within the haplotype network involved clustering of haplotypes from the same sample, as observed for the haplotypes of the Kilkieran Lake (KILK) and the haplotypes (except haplotype Id) of the Portuguese Ria de Aveiro population (RdA).

Spatial genetic structure

Statistically significant differences were observed in haplotype frequencies among all samples (global test) and among all pairs of samples ($P < 0.0009$ in all cases), except the two Baltic samples ($P = 0.489$). The AMOVA showed that the global Φ_{ST} value across all samples amounted to 0.7867 ($P < 0.001$), indicating a significant differentiation between samples. Additional significant geographical structuring was tested by grouping the different samples in four geographical regions: Baltic Sea (TV and VI), North Sea (WE, WS and YTH), English Channel (SEI and EL) and a fourth group including the KILK, GI and RdA samples. The GU sample was excluded to avoid

Table 2 Distribution of COI haplotypes (columns) among populations (rows) of *Neomysis integer*.

	TV	VI	WE	WS	YTH	SEI	EL	KILK	GI	RdA	GU	Total
Df	39	41	10	42		3						135
Ee	1		25	1		36						63
Ga								36				36
Cc											35	35
Ha							30					30
Fh					20	7	3					29
Aa									28			28
Fg			4	13								17
Fa					16							16
Bb									15			15
Fd										14		14
Id										13		13
Ja								4				4
D'i				2								2
Fj						2						2
Kc											2	2
Fk					2							2
Fl										1		1
Fm										1		1
D'n				1								1
Eo				1								1
Lc											1	1
Mp											1	1
Cq											1	1
Ih					1							1
Gr								1				1
Je							1					1
Ds	1											1
Na								1				1
Ot									1			1
Pa								1				1
Qd										1		1
Ra							1					1
Rh							1					1

For sampling site abbreviations see Table 1.

distortion of the results because of the large divergence of its haplotypes. A significant amount of variation (29%) was explained by differences among the geographical groups, however, the largest amount of variation was observed at the interpopulation within-groups level (Table 3). This heterogeneity among samples is also evident from the pairwise genetic distances (Table 4). Virtually all pairwise comparisons were significant, with the exception of the pairwise genetic distances between the WE and SEI samples and between both Baltic samples (VI and TV). Highest differentiation was found between the Guadalquivir (GU) and the Vistula (VI) population ($\Phi_{ST} = 0.976$). Genetic distances within the English Channel, North Sea and Baltic Sea ranged from low ($\Phi_{ST} = 0$) and moderate values ($\Phi_{ST} = 0.5$) (comparisons involving geographically closely located samples) to high values ($\Phi_{ST} > 0.8$) (comparisons involving the Baltic and the YTH and EL samples).

The Mantel test showed a highly significant positive correlation between geographical distance and genetic distance (Φ_{ST}) among all samples ($P = 0.008$). This indicates a genetic IBD pattern, with geographical distance explaining 18.8% of the mitochondrial DNA variation found ($r = 0.43346$). At a smaller geographical scale, IBD was detected in the samples of the Baltic Sea, North Sea and English Channel ($r = 0.4572$, $P = 0.0158$), with 21% of the variation in genetic differentiation explained by geographical distance. A plot of the genetic distance vs. geographical distance, showing the pattern of IBD, is shown in Fig. 3. Even when excluding the outliers from the analysis, a significant correlation could still be detected ($P = 0.0014$).

Nested clade analysis

The 34 haplotypes fitted into 17 one-step clades, eight two-step clades and three three-step clades (Fig. 4). The nested contingency analysis detected significant associations between haplotypes and geography within eight nested haplotype clades (Table 5). The NCA detected signals of contemporary

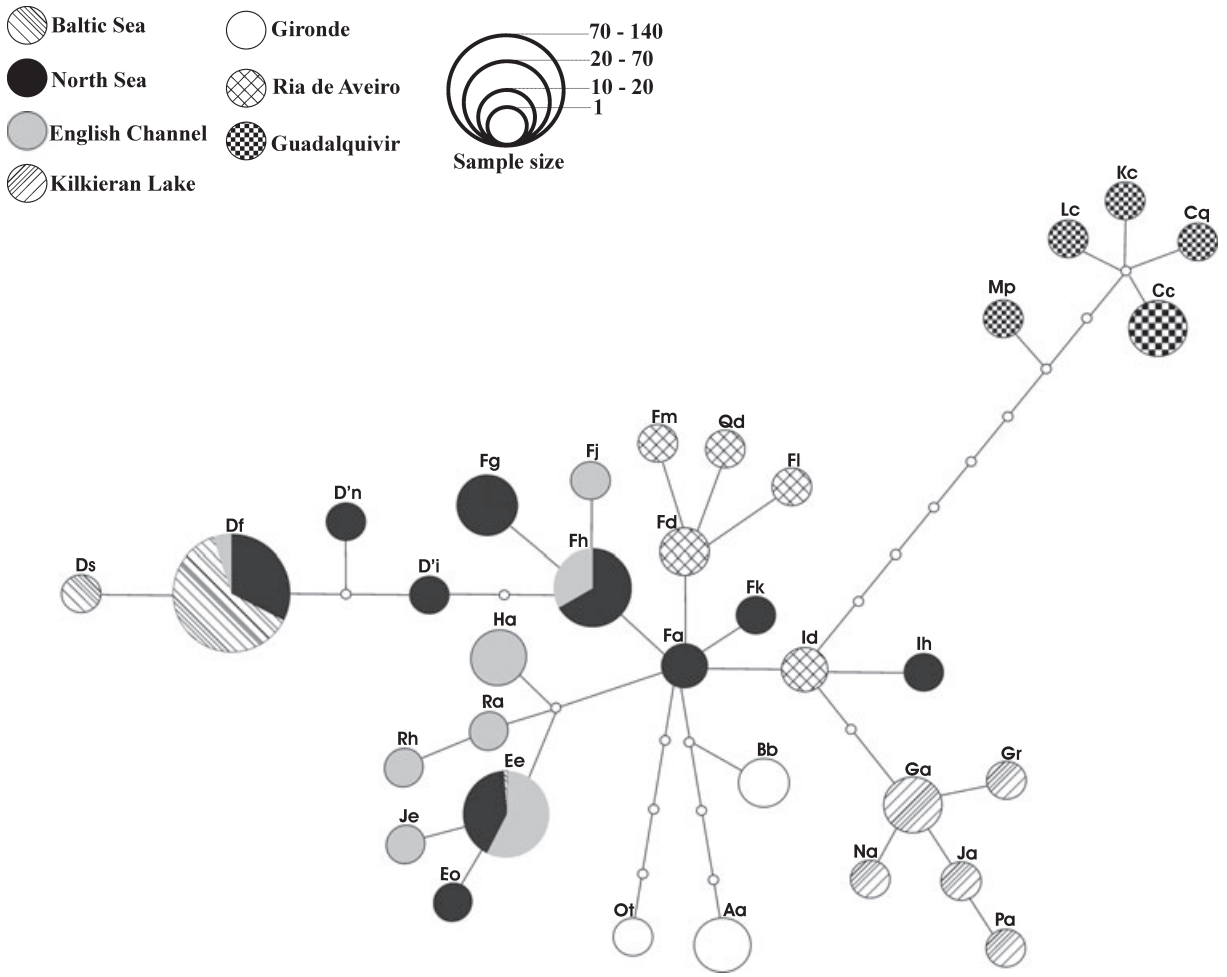

Figure 2 The 95% plausible parsimony network showing the mutational relationships among the COI haplotypes of *Neomysis integer*. Each line in the network represents a single mutational change; haplotypes are represented by circles. The size of each circle is proportional to the frequency of occurrence of the haplotype; circles are shaded according to the geographical occurrence of the haplotype. Small empty circles indicate missing haplotypes.

Table 3 Results of the hierarchical analysis of molecular variance (AMOVA).

Analysis	Source of variation	d.f.	Percentage total variance	Fixation indices	*P*-value
All samples	Among populations	10	78.67	$\Phi_{ST} = 0.7867$	< 0.001
	Within populations	453	21.33		
Without GU sample	Among populations	9	71.39	$\Phi_{ST} = 0.7139$	< 0.001
	Within populations	414	28.61		
Four groups (TV, VI) (WE, WS, YTH) (SEI, EL) (KILK, GI, RdA)					
	Among groups	3	27.45	$\Phi_{CT} = 0.2745$	0.024
	Among populations/Within groups	6	45.21	$\Phi_{SC} = 0.6232$	< 0.001
	Within populations	411	27.34	$\Phi_{ST} = 0.72602$	< 0.001

For sample codes see Table 1.

processes such as restricted gene flow with IBD at lower-level nesting groups (1-4 and 1-9). The distribution of these two one-step clades (1-4 and 1-9) seems to be restricted to the English Channel, North Sea and Baltic Sea samples (Fig. 5b). IBD was also detected at two higher-nesting groups (2-7 and 3-2); however, in these cases an effect of past fragmentation and/or range expansion could not be ruled out due to an inadequate sampling scheme (Table 5). At higher-nesting

Table 4 Pairwise Φ_{ST} values among the 11 samples of *Neomysis integer*.

	TV	VI	WE	WS	YTH	SEI	EL	KILK	GI	RdA	GU
TV	–										
VI	0.000$^{n.s.}$	–									
WE	0.622	0.666	–								
WS	0.127	0.166	0.454	–							
YTH	0.816	0.874	0.381	0.521	–						
SEI	0.778	0.814	0.041$^{n.s.}$	0.624	0.517	–					
EL	0.882	0.916	0.440	0.726	0.692	0.505	–				
KILK	0.949	0.972	0.755	0.830	0.867	0.817	0.885	–			
GI	0.806	0.829	0.544	0.710	0.639	0.602	0.649	0.810	–		
Rda	0.833	0.874	0.452	0.641	0.457	0.551	0.663	0.790	0.612	–	
GU	0.963	0.976	0.842	0.901	0.932	0.879	0.916	0.961	0.871	0.911	–

n.s., Not significant, $P > 0.05$.

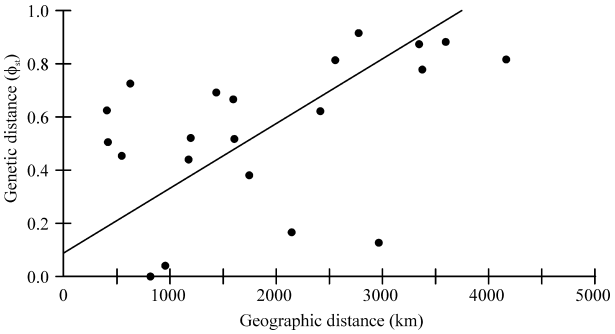

Figure 3 Pairwise genetic distances (Φ_{ST}) between the Baltic Sea, North Sea and English Channel samples of *Neomysis integer* plotted as function of the geographical distance (minimal coastline distance) between the samples ($n = 7$). The slope had a value of 2.43×10^{-4} and R^2 of 0.209 with 95% confidence intervals for the slope (1.469×10^{-4}, 3.3125×10^{-4}) and R^2 (0.007, 0.554).

groups (two-step and three-step clades) several historical demographic events, such as past population fragmentation and range expansion, could be inferred. A contiguous range expansion from the English Channel throughout the North Sea into the Baltic Sea is inferred for clade 2-4, which is restricted to the samples of the English Channel, North Sea and to a lesser extent the Baltic Sea (Fig. 5c). Similarly, the signals of past fragmentation and range expansion detected in nested clade 2-3 are caused by the restricted distribution of interior clade 1-5 (haplotypes Fa and Fk, found in the YTH sample) and the tip clade 1-6 (restricted to the RdA sample).

At the level of the total cladogram, the haplotype network shows a grouping into three clades (3-1, 3-2 and 3-3) that have a different geographical distribution (Fig. 5d). Clade 3-3 is restricted to the southern GU sample, clade 3-1 seems to have a northern distribution, dominating the Baltic and North Sea samples, whereas clade 3-2 dominates the more southern samples (RdA, GI, SEI, EL and KILK). Zones of geographical overlap of both clades are found in the English Channel (EL and SEI samples) and the North Sea (YTH and WE samples). The NCA revealed that the pattern at the level of the total cladogram might be caused by a range expansion and past fragmentation. But, again due to the lack of northern Iberian samples and intermediate samples between the Portuguese Ria de Aveiro sample (RdA) and the southern Guadalquivir sample (GU), a long-distance colonization process might have caused a similar pattern.

Demographic history

The overall mismatch distribution was clearly multimodal (distributions not shown), hence a fit to the sudden expansion model of Rogers (1995) was significantly rejected [$P(SSD_{obs}) = 0.04$]. Likewise, signals of a post-glacial population expansion could not be detected in the North Sea/English Channel region [$P(SSD_{obs}) = 0.02$]. In addition both Tajima's D and Fu's F_s values were not significantly different from zero, which further supports the hypothesis of a stable population structure (Table 6). Only the mismatch distributions from the Irish (KILK) and the southern GU sample seemed to fit the distribution underlying the sudden expansion model of Rogers (1995) [$P(SSD_{obs}) = 0.46$ and $P(SSD_{obs}) = 0.07$, respectively]. However, the results of Tajima's and Fu's neutrality tests were not in all cases congruent with the mismatch distributions, and hence did not always support the model of sudden population expansion. In the case of the Guadalquivir population (GU), Fu's F_s value was negative but non-significant, while for the Irish (KILK) population, Tajima's D was marginally non-significant ($P = 0.051$).

DISCUSSION

The overall phylogeographical structure of *N. integer* is characterized by a large heterogeneity between the populations, as indicated by the AMOVA analysis. There is a clear geographical clustering of the haplotypes, showing a completely different pattern in the sites north and south of the English Channel. Samples from the areas north of the English Channel share several common haplotypes, while the southern samples consist solely of population-specific and unique

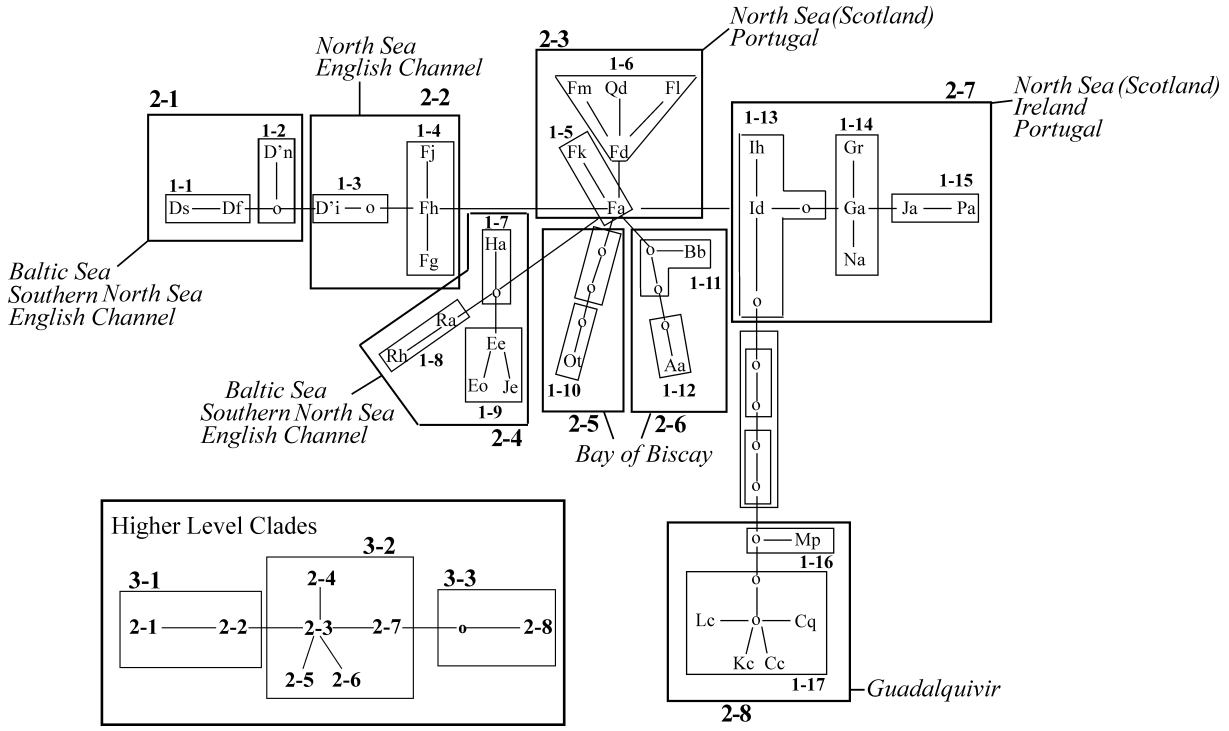

Figure 4 Haplotype network among COI haplotypes of *Neomysis integer* with nesting design used in the nested clade analysis. The geographical distribution of the two-step clades is indicated.

Table 5 Nested contingency analysis of geographical associations with phylogeographical inferences from the nested clade analysis (Templeton, 1998).

Clade	χ^2	Probability	Inference	Comments
1-1	2.4178	0.4210		
1-4	57.3481	0.0000*	1-2-3-4-No	Restricted gene flow with isolation-by-distance
1-9	96.9841	0.0010*	1-2-11-17-4-No	Restricted gene flow with isolation-by-distance
1-13	14.0000	0.0760		
2-1	2.2021	1.0000		
2-2	6.1714	0.2650		
2-3	34.0000	0.0000*	1-2-11-12-13-14-Yes	Range expansion, long distance colonization and/or past fragmentation (sampling design inadequate)
2-4	92.6135	0.0000*	1-2-11-12-No	Contiguous range expansion
2-7	56.0000	0.0000*	1-2-3-5-15-16-18-No	Past fragmentation, range expansion or isolation-by-distance (inadequate sampling scheme)
3-1	125.1341	0.0000*	1-2-11-12-No	Contiguous range expansion
3-2	609.6668	0.0000*	1-2-3-4-9-10-No	Past fragmentation or isolation-by-distance (inadequate sampling scheme)
Total	772.3222	0.0000*	1-2-11-13-13-14-Yes	Range expansion, long distance colonization and/or past fragmentation (sampling design inadequate)

*Significant at 5% level, derived from 10,000 random permutations.
For nesting design, see Fig. 4.

haplotypes (Fig. 5a; Tables 1 and 2). At a higher-nesting clade level in the haplotype network, the clades 3-1, 3-2 and 3-3 show a different geographical distribution (Fig. 5d). While clade 3-3 is restricted to the southern Guadalquivir population (GU), clades 3-1 and 3-2 have a larger distribution, with the first occurring predominantly in populations north of the English Channel, and the latter found in more southerly populations and around the UK. Both clades are only observed sympatrically in the English Channel and North Sea populations. Based on the high number of haplotypes and the interior position within the network, clade 3-2 is assumed to be more ancestral than the others.

Clade 3-3 is characterized by a large divergence of its haplotypes from all other haplotypes (1.3–2.68%). This points

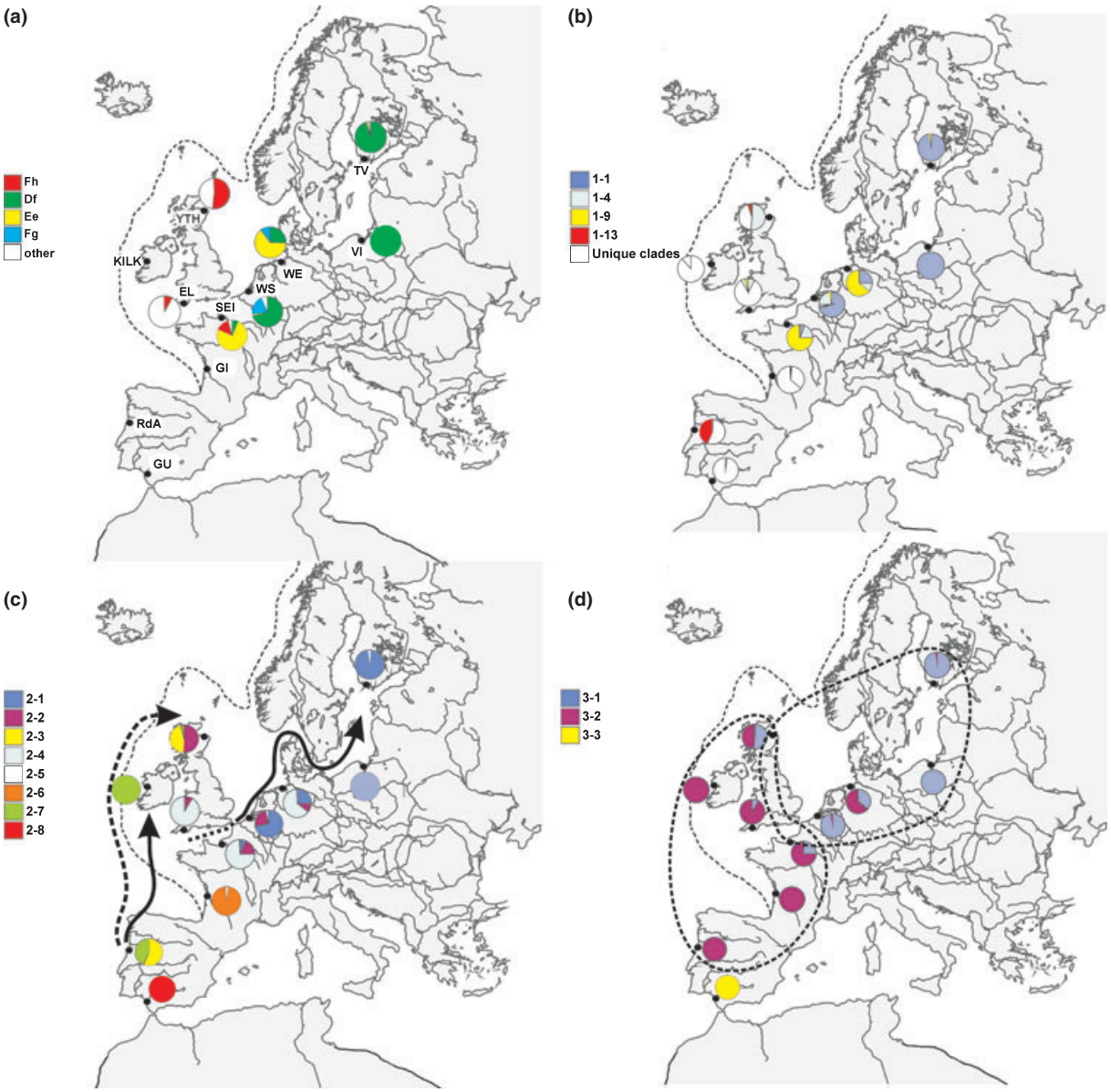

Figure 5 Geographical distribution of 0-, one-, two- and three-step clades. Dashed line, shoreline during the Last Glacial Maximum (18 ka) (redrawn from Frenzel *et al.*, 1992). Distribution of: (a) shared 0-step clades (haplotypes); (b) one-step clades identified in the nested clade analysis (see nested haplotype network in Fig. 4); (c) two-step clades (arrows indicate putative colonization routes); (d) three-step clades.

Table 6 Tests of neutrality within the pooled samples of the major geographical regions of *Neomysis integer*.

	Samples	Tajima's *D*	*P*-value	Fu's F_s	*P*-value
All samples		−0.3415	0.3940	−5.8146	0.1120
Baltic Sea	VI and TV	−2.0403	0.0000	−1.7188	0.0950
North Sea	WE, WS and YTH	1.0602	0.8760	0.3368	0.6120
English Channel	EL and SEI	0.0436	0.6000	−0.0280	0.5310
Ireland	KILK	−1.4587	0.0510	−3.1766	0.0080
Bay of Biscay	GI	−0.4632	0.5790	4.8800	0.9750
Portugal	RdA	−0.0995	0.5130	−0.1123	0.4760
Gulf of Cadiz	GU	−1.8069	0.0080	−1.8330	0.0800

to a long-lasting isolation of the Guadalquivir population (GU), as suggested previously based on a limited number of mtDNA cyt *b* sequences (Remerie *et al.*, 2006). Such isolation is not unexpected as the genetic divergence may be enhanced near the edge of a species' distribution range due to selective sweeps and higher genetic drift in these populations (Lesica & Allendorf, 1995; Hewitt, 2000; Coyer *et al.*, 2003; Hoffman & Blouin, 2004). The below-average densities of *N. integer* in the Guadalquivir estuary (GU) indicate that these southern habitats may not be optimal for the species and that they live near the tolerance limit (Drake *et al.*, 2002). The comparatively low levels of genetic diversity and the rejection of the hypothesis of neutral variation with Tajima's *D*-test further support the hypothesis of a selective sweep in this population (Table 6). Populations within clades 3-1 and 3-2 are less diverged; however, an approximation of the divergence times of the isolated Bay of Biscay and Irish populations clearly seem to pre-date the timing of the Late Glacial Maximum (LGM; 18 ka) when using a 'traditional' molecular clock of 1.4–2.6% Myr^{-1} (GI vs. others, 100–350 ka; KILK vs. others, 130–180 ka). However, these pre-LGM divergences need to be interpreted with caution as the molecular clock calibrations used are derived from unrelated taxa (*Alpheus* sp. and *Sesarama* sp., see Materials and Methods) and are based on old events (Isthmus of Panama vicariance, 3 Ma). Recent studies have shown that the use of these 'old' clocks might be inappropriate for calculating divergences on much more recent time-scales because of the so-called time dependency of molecular clocks (Ho *et al.*, 2005; Ho & Larson, 2006). Elevated rates of molecular change on post-glacial time-scales have already been proposed in some fish taxa, and even in circumarctic mysid crustaceans (Audzijonyte & Väinölä, 2006; Burridge *et al.*, 2008). Hence, the observed pre-LGM divergence timings could be overestimated and the divergences might be concordant with post-glacial time-scales when using an appropriate, species-specific molecular clock.

In contrast to our expectations based on the general trend of declining genetic diversity from south to north in continental Europe and North America ('northern purity–southern richness' model; see Hewitt, 1996, 2000, 2004; Avise, 2000), no distinct gradient in molecular diversity was observed in the present study, considering that the average haplotype diversity of the North Sea populations was only slightly lower than in the Iberian RdA population (Table 1). However, it must be noted that species-specific attributes such as colonizing ability (related to dispersal capacities and physiological tolerance) largely influence the general predictions of latitudinal trends in genetic diversity (Taberlet *et al.*, 1998). Computer simulations have shown that a slow post-glacial colonization ('phalanx' or diffusive expansions) tends to retain the levels of genetic diversity (Nichols & Hewitt, 1994; Ibrahim *et al.*, 1996). This slow process involves a high proportion of individuals dispersing over short distances in a continuous front, a pattern that does not involve a subsampling of the genetic diversity of source populations through founder events and where the effective population sizes remain much larger than in the case of a fast pioneer colonization process (Nichols & Hewitt, 1994). Although *N. integer* has a predominant estuarine distribution, low densities (annual average of 12 individuals per 100 m^2) have been reported in coastal waters (Beyst *et al.*, 2001), especially during winter months, when floodwater discharge is higher and *N. integer* has an increased salinity tolerance under the lower temperatures (Kinne, 1955; Vlasblom & Elgershuizen, 1977). Hence, such a gradual range expansion and colonization of estuaries might be expected for this species.

Considering scenarios of continuous post-glacial range expansion from southern refugia, as assumed for many terrestrial species in Europe (Taberlet *et al.*, 1998; Hewitt, 2000) and a variety of marine species along the north-east Atlantic (Garcia-Marin *et al.*, 1999; Consuegra *et al.*, 2002; Gysels *et al.*, 2004a), it is surprising that not a single haplotype was shared between *N. integer* populations in glaciated and unglaciated regions. According to the present data, it seems that the southern populations of *N. integer* (e.g. GI, RdA and GU) did not participate in the most recent range expansion to northern areas after the LGM. In contrast, NCA revealed evidence of a range expansion from the Iberian coast to northern European regions only at higher clade levels (Fig. 5c, putative range expansion from the Iberian to the Irish and northern UK coasts). In addition, the higher genetic diversity and heterogeneity in glaciated areas (English Channel and North Sea) than expected for populations that have been repeatedly affected by founder-flush cycles during Pleistocene glaciations is suggestive of the hypothesis of range persistence during the LGM. This could imply that during the Holsteinian or Eemian interglacials (400–370 and 120–100 ka, respectively), mysids from a southern refugium, located along the Iberian coasts, colonized northern Europe. However, instead of being pushed back south during the following glaciations (Saalian and Weichselian); some of these populations might have survived the subsequent glaciations in secondary northern refugia.

A growing number of phylogeographical surveys in marine taxa along the north-east Atlantic coasts have recognized the existence of northern glacial refugia, probably located in a palaeoriver system within the English Channel (the Hurd Deep depression, see Lericolais *et al.*, 1995) and/or on ice-free coasts in south-western Ireland (Wilke & Davis, 2000; Consuegra *et al.*, 2002; Breton *et al.*, 2003; Coyer *et al.*, 2003; Luttikhuizen *et al.*, 2003; Gysels *et al.*, 2004a,b; Provan *et al.*, 2005; Chevolot *et al.*, 2006; Jolly *et al.*, 2006; Hoarau *et al.*, 2007). A similar scenario was suggested previously for *N. integer* (Remerie *et al.*, 2006) based on the distribution of cyt *b* haplotypes, but, due to the limited sample size, no firm conclusions could be drawn on the levels of genetic diversity and hence alternative hypotheses could not be ruled out. These preliminary results are corroborated by the present study, as *N. integer* populations in the English Channel/North Sea region are characterized by a relatively high genetic diversity and several widespread haplotypes (e.g. Df, Ee, Fh) that are not detected elsewhere. Such a pattern is characteristic of a glacial

refugium for *N. integer* in that region, given that it was able to withstand the lower temperatures during the LGM (summer sea surface temperature 5–6°C; Meland *et al.*, 2005).

The differential distribution of the (common) haplotypes (e.g. Df, Fh) suggest that the post-glacial recolonization of northern Europe took place in different phases and via different routes after the LGM. Before the opening of the Dover Strait (around 7.5 ka), a first wave of colonizers was forced to go around the UK, while after the establishment of the connection between the English Channel and North Sea, a second colonization wave could take the route through the Dover Strait. The higher degree of genetic structuring and the high number of private haplotypes of the UK and Irish populations (YTH, KILK, EL) are indicative of this older colonization event, whereas the high frequency of the interior haplotype Df in populations of the continental coasts of the English Channel (SEI), North Sea (WS and WE) and the Baltic Sea (TV and VI) point to a colonization by a more recent sweep from the same refugium.

In addition, the geographically structured pattern of recolonization may be largely a result of the poor dispersal ability of *N. integer* and the structuring nature of the estuarine habitat. Estuaries are believed to be sources of diversification by adaptive divergence and spatial isolation (Kelly *et al.*, 2006). The seasonal oscillations of salinity, turbidity, temperature and other environmental variables, combined with the discontinuity and isolation of this habitat, all contribute to the genetic structuring of estuarine species, enhancing the effects of microevolutionary processes such as gene flow, drift and selection (Bilton *et al.*, 2002; Virgilio & Abbiati, 2004). Unlike the high census size of *N. integer* populations (average densities up to 160 individuals m^{-2}; Mees *et al.*, 1994), extremely low estimates of genetic effective population size relative to census population size have been proposed for this species (Remerie, 2005). Hence, genetic diversity might be lost at a high rate through drift at each generation, therefore we can assume that these populations are strongly affected by genetic drift but show little influence of gene flow. Occasional gene flow between neighbouring populations could be caused by stochastic events, such as exceptional rainfall or floods. Under these circumstances, gene flow might occur between proximate estuaries by plumes of floodwater discharge that extend offshore (see the hydrodynamic model of Lacroix *et al.*, 2004, in the southern North Sea). This mechanism has been observed in several estuarine and brackish water species with disjunct distributions (Maltagliati, 2002; Burridge *et al.*, 2004). The results of the AMOVA analysis, the high pairwise Φ_{ST} values and the detection of an IBD pattern further corroborate the very limited degree of gene flow between adjacent *N. integer* populations. However, in the case of both Baltic samples, a non-significant pairwise Φ_{ST} value was observed. Homogeneity of the Baltic samples could be caused by the recent colonization of the Baltic (last 8000 years, after the opening of the Danish Straits) resulting in a migration–drift balance that has not yet attained equilibrium. Alternatively, high rates of gene flow within the Baltic Sea could be linked with the specific environmental characteristics of the Baltic Sea. The water in the Baltic Sea is brackish, with an average salinity lower than 10 practical salinity units; this could result in a higher connectivity between suitable habitats for *N. integer* leading to higher rates of gene flow within the Baltic. In addition, the inference of restricted gene flow at different levels in the NCA is also consistent with high levels of population subdivision (Table 5). Surprisingly, two clades analysed in the NCA suggested long-distance dispersal among geographical regions as one of the possible inferences. It appears more likely to us that this inference is due to the inadequate sampling scheme, rather than a biological reality.

CONCLUSIONS

The present study of mitochondrial COI variation in populations of the mysid *N. integer* revealed a significant differentiation throughout its distribution range, with a complex phylogeographical structure. The heterogeneous distribution of the haplotypes in northern Europe and the lack of a clear latitudinal trend in the levels of genetic diversity point to colonization of these areas prior to the last glaciation and survival in a northern periglacial refugium. This contradicts the general expectations derived from current palaeo-climatological and palaeo-oceanographic models (Frenzel *et al.*, 1992; Meland *et al.*, 2005) that suggest survival during glacial periods in southern refugia located along the Iberian coasts. Although supported by a previous study of *N. integer* and many other marine species, additional analyses of samples from 'critical' areas such as the northern Iberian Peninsula, Bay of Biscay and coasts of Brittany may be useful to validate the current hypothesis. Likewise, additional analyses of unlinked nuclear loci might be needed as the genetic pattern observed at a single (mitochondrial) locus represents just one realization of an evolutionary process with a large stochastic component (Maddison, 1997; Nichols, 2001).

ACKNOWLEDGEMENTS

This research is supported by the FWO research project G029200, the GOA BOF project 'Biodiversity of marine benthic communities along ecological gradients' (contract no. 1205398), the GOA BOF project 'Biogeography and Biodiversity of the Sea – BBSea' (contract no. 01GZ0705) from Ghent University (Flemish Government of Education), and the TROPHOS and WESTBANKS projects of the Belgian Federal Science Policy Office (BELSPO, contract nos. EV/02/25A and SD/BN/01A). The authors acknowledge the support by the MarBEF Network of Excellence which is funded by the Sustainable Development, Global Change and Ecosystem Programme of the European Community's Sixth Framework Programme (contract no. GOCE-CT-2003-505446). This publication is contribution number MPS-08024 of MarBEF. We would like to thank R. Väinölä, A. Audzijonyte, B. Malinga, J. Witt, M. Solan, J.-C. Dauvin, B. Sautour, M. Cunha, P. D. Moyano, G. Oliver, S. Roast, S. Shellito, N.

Fockedey, K. Hostens, S. Degraer and T. Gheskiere for their help with collecting and/or providing *Neomysis integer* samples. We thank J. Vanhoudt, B. Hellemans and J. Pinceel for the training in SSCP analysis. We are grateful to E. Gysels and F. Volckaert for comments on an earlier version of the manuscript.

REFERENCES

Althoff, D.M. & Pellmyr, O. (2002) Examining genetic structure in a bogus yucca moth: a sequential approach to phylogeography. *Evolution*, **56,** 1632–1643.

Audzijonyte, A. & Väinölä, R. (2006) Phylogeographic analyses of a circumarctic coastal and boreal lacustrine mysid crustacean, and evidence of fast post-glacial mtDNA rates. *Molecular Ecology*, **15,** 3287–3301.

Avise, J.C. (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge.

Avise, J.C., Walker, D. & Johns, G.C. (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society B: Biological Sciences*, **265,** 1707–1712.

Beyst, B., Buysse, D., Dewicke, A. & Mees, J. (2001) Surf zone hyperbenthos of Belgian sandy beaches: seasonal patterns. *Estuarine, Coastal and Shelf Science*, **53,** 877–895.

Bilton, D.T., Pauls, J. & Bishop, J.D.D. (2002) Dispersal, genetic differentiation and speciation in estuarine organisms. *Estuarine, Coastal and Shelf Sciences*, **55,** 937–952.

Bohonak, A.J. (2002) IBD (isolation by distance): a program for analyses of isolation by distance. *Journal of Heredity*, **93,** 154–155.

Breton, S., Dufresne, F., Desrosiers, G. & Blier, P.U. (2003) Population structure of two northern hemisphere polychaetes, *Neanthes virens* and *Hediste diversicolor* (Nereididae), with different life-history traits. *Marine Biology*, **142,** 707–715.

Burridge, C.P., Hurt, A.C., Farrington, L.W., Coutin, P.C. & Austin, C.M. (2004) Stepping stone gene flow in an estuarine-dwelling sparid from south-east Australia. *Journal of Fish Biology*, **64,** 805–819.

Burridge, C.P., Craw, D., Fletcher, D. & Waters, J.M. (2008) Geological dates and molecular rates: fish DNA sheds light on time dependency. *Molecular Biology and Evolution*, **25,** 624–633.

Chevolot, M., Hoarau, G., Rijnsdorp, A.D., Stam, W.T. & Olsen, J.L. (2006) Phylogeography and population structure of thornback rays (*Raja clavata* L., Rajidae). *Molecular Ecology*, **15,** 3693–3705.

Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9,** 1657–1659.

Consuegra, S., De Leaniz, C.G., Serdio, A., Morales, M.G., Straus, L.G., Knox, D. & Verspoor, E. (2002) Mitochondrial DNA variation in Pleistocene and modern Atlantic salmon from the Iberian glacial refugium. *Molecular Ecology*, **11,** 2037–2048.

Coyer, J.A., Peters, A.F., Stam, W.T. & Olsen, J.L. (2003) Post-ice age recolonization and differentiation of *Fucus serratus* L. (Fucaceae: Phaeophyta) populations in Northern Europe. *Molecular Ecology*, **12,** 1817–1829.

Crandall, K.A. (1996) Multiple interspecies transmissions of human and simian T-cell leukaemia/lymphoma virus type I sequences. *Molecular Biology and Evolution*, **13,** 115–131.

Crandall, K.A. & Templeton, A.R. (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134,** 959–969.

Cunha, M.R., Sorbe, J.C. & Moreira, M.H. (1999) Spatial and seasonal changes of brackish peracaridan assemblages and their relation to some environmental variables in two tidal channels of the Ria de Aveiro (NW Portugal). *Marine Ecology Progress Series*, **190,** 69–87.

Dahlgren, T.G., Weinberg, J.R. & Halanych, K.M. (2000) Phylogeography of the ocean quahog (*Arctica islandica*): influences of paleoclimate on genetic diversity and species range. *Marine Biology*, **137,** 487–495.

Drake, P., Arias, A.M., Baldó, F., Cuesta, J.A., Rodríguez, A., Silva-García, A., Sobrino, A., García-González, D. & Fernández-Delgado, C. (2002) Spatial and temporal variation of the nekton and hyperbenthos from a temperate European estuary with regulated freshwater inflow. *Estuaries*, **25,** 451–468.

Edmands, S. (2001) Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. *Molecular Ecology*, **10,** 1743–1750.

Excoffier, L., Smouse, P. & Quattro, J. (1992) Analyses of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131,** 479–491.

Filatov, D.A., Moneger, F., Negrutiu, I. & Charlesworth, D. (2000) Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution. *Nature*, **404,** 388–390.

Fockedey, N. & Mees, J. (1999) Feeding of the hyperbenthic mysid *Neomysis integer* in the maximum turbidity zone of the Elbe, Westerschelde and Gironde estuaries. *Journal of Marine Systems*, **22,** 207–228.

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. *Molecular Marine Biology and Biotechnology*, **3,** 294–297.

Frenzel, B., Pécsi, M. & Velichko, A.A. (1992) *Atlas of paleoclimates and paleoenvironments of the Northern Hemisphere: Late Pleistocene–Holocene*. Geographic Research Institute, Hungarian Academy of Science, Budapest.

Fu, Y.X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147,** 915–925.

Fukatsu, T. (1999) Acetone preservation: a practical technique for molecular analysis. *Molecular Ecology*, **8,** 1935–1945.

García-Marín, J.L., Utter, F.M. & Pla, C. (1999) Postglacial colonization of brown trout in Europe based on distribution allozyme variants. *Heredity*, **82,** 46–56.

Gysels, E.S., Hellemans, B., Pampoulie, C. & Volckaert, F.A.M. (2004a) Phylogeography of the common goby, *Pomatoschistus microps*, with particular emphasis on the colonization of the Mediterranean and the North Sea. *Molecular Ecology*, **13,** 403–417.

Gysels, E.S., Hellemans, B., Patarnello, T. & Volckaert, F.A.M. (2004b) Current and historic gene flow of the sand goby *Pomatoschistus minutus* on the European Continental Shelf and in the Mediterranean Sea. *Biological Journal of the Linnean Society*, **83,** 561–576.

Hewitt, G.M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58,** 247–276.

Hewitt, G.M. (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68,** 87–112.

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

Hewitt, G.M. (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10,** 537–549.

Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359,** 183–195.

Ho, S.Y.W. & Larson, G. (2006) Molecular clocks: when times are a-changin'. *Trends in Genetics*, **22,** 79–83.

Ho, S.Y.W., Phillips, M.J., Cooper, A. & Drummond, A.J. (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22,** 1561–1568.

Hoarau, G., Coyer, J.A., Veldsink, J.H., Stam, W.T. & Olsen, J.L. (2007) Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus*. *Molecular Ecology*, **16,** 3606–3616.

Hoffman, E.A. & Blouin, M.S. (2004) Historical data refute range contraction as cause of low genetic diversity in isolated frog populations. *Molecular Ecology*, **13,** 271–276.

Hostens, K. & Mees, J. (1999) The mysid-feeding guild of demersal fishes in the brackish zone of the Westerschelde estuary. *Journal of Fish Biology*, **55,** 704–719.

Ibrahim, K.M., Nichols, R.A. & Hewitt, G.M. (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77,** 282–291.

Jolly, M.T., Jollivet, D., Gentil, F., E Thiébaut, E. & Viard, F. (2005) Sharp genetic break between Atlantic and English Channel populations of the polychaete *Pectinaria koreni*, along the North coast of France. *Heredity*, **94,** 23–32.

Jolly, M.T., 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. *Molecular Ecology*, **15,** 3481–3482.

Kelly, D.W., MacIsaac, H.J. & Heath, D.D. (2006) Vicariance and dispersal effects on phylogeographic structure and speciation in a widespread estuarine invertebrate. *Evolution*, **60,** 257–267.

Kinne, O. (1955) *Neomysis vulgaris* Thompson, eine autökologisch-biologische Studie. *Biologisches Zentralblatt*, **74,** 160–202.

Knowles, L.L., Futuyma, D.J., Eanes, W.F. & Rannala, B. (1999) Insights into speciation mode from historical demography in the phytophagous beetle *Ophraella*. *Evolution*, **53,** 1846–1856.

Knowlton, N., Weight, L.A., Solorzano, L.A., Mills, D.K. & Berningham, E. (1993) Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. *Science*, **260,** 1629–1632.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X. & Wilson, A.C. (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences USA*, **86,** 6196–6200.

Lacroix, G., Ruddick, K., Ozer, J. & Lancelot, C. (2004) Modelling the impact of the Scheldt and Rhine/Meuse plumes on the salinity distribution in Belgian waters (southern North Sea). *Journal of Sea Research*, **52,** 149–163.

Lericolais, G., Auffret, J.P. & Bourillet, J.F. (1995) The Quaternary Channel River: seismic stratigraphy of its palaeo-valleys and deeps. *Journal of Quaternary Science*, **18,** 245–260.

Lesica, P. & Allendorf, F.W. (1995) When are peripheral populations valuable for conservation? *Conservation Biology*, **9,** 753–760.

Luttikhuizen, P.C., Drent, J. & Baker, A.J. (2003) Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larval dispersal. *Molecular Ecology*, **12,** 2215–2229.

Maddison, W.P. (1997) Gene trees in species trees. *Systematic Biology*, **46,** 523–536.

Maltagliati, F. (2002) Genetic monitoring of brackish-water populations: the Mediterranean toothcarp *Asphanius fasciatus* (Cyprinodontidae) as a model. *Marine Ecology Progress Series*, **235,** 257–262.

Marko, P.B. (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. *Molecular Ecology*, **13,** 597–611.

Maruyama, T. & Birky, C.W., Jr (1991) Effects of periodic selection on gene diversity in organelle genomes and other systems without recombination. *Genetics*, **127,** 449–451.

Mauchline, J. (1971) The biology of *Neomysis integer*. *Journal of the Marine Biological Association of the UK*, **51,** 347–354.

Mauchline, J. (1980) *The biology of mysids and euphausiids*. Advances in Marine Biology, Vol. 18. Academic Press, London.

Mees, J., Cattrijsse, A. & Hamerlynck, O. (1993) Distribution and abundance of shallow water hyperbenthic mysids (Crustacea, Mysidacea) and euphausiids (Crustacea, Euphausiacea) in the Voordelta and the Westerschelde, southwest Netherlands. *Cahier Biologique Marine*, **34,** 165–186.

Mees, J., Abdulkerim, Z. & Hamerlynck, O. (1994) Life history, growth and production of *Neomysis integer* in the Westerschelde estuary (SW Netherlands). *Marine Ecology Progress Series*, **109,** 43–57.

Mees, J., Fockedey, N. & Hamerlynck, O. (1995) Comparative study of the hyperbenthos of three European estuaries. *Hydrobiologia*, **311,** 153–174.

Meland, M.Y., Jansen, E. & Elderfield, H. (2005) Constraints on SST estimates for the northern North Atlantic/Nordic Seas during the LGM. *Quaternary Science Reviews*, **24,** 835–852.

Nei, M. (1987) *Molecular evolutionary genetics*. Columbia University Press, New York.

Nichols, R. (2001) Gene trees and species trees are not the same. *Trends in Ecology and Evolution*, **16,** 358–364.

Nichols, R.A. & Hewitt, G.M. (1994) The genetic consequences of long distance dispersal during colonization. *Heredity*, **72,** 312–317.

Olsen, J.L., Stam, W.T., Coyer, J.A., Reusch, T.B.H., Billingham, M., Boström, C., Calvert, E., Christie, H., Granger, S., La Lumière, R., Milchakova, N., Oudot-Le Secq, M.-P., Procaccini, G., Sanjabi, B., Serrão, E., Veldsink, J., Iddicombe, S. & Wyllie-Echeverria, S. (2004) North Atlantic phylogeography and large-scale population differentiation of the seagrass *Zostera marina* L. *Molecular Ecology*, **13,** 1923–1941.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. & Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences USA*, **86,** 2766–2770.

Posada, D. & Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14,** 817–818.

Posada, D., Crandall, K.A. & Templeton, A.R. (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9,** 487–488.

Provan, J., Wattier, R.A. & Maggs, C.A. (2005) Phylogeography of the red seaweed *Palmaria palmata* reveals a Pleistocene marine glacial refugium in the English Channel. *Molecular Ecology*, **14,** 793–805.

Remerie, T. (2005) *Molecular diversity and population structure of two mysid taxa along European coasts*. PhD Thesis, Ghent University, Ghent.

Remerie, T., Gysels, E., Vierstraete, A., Vanfleteren, J. & Vanreusel, A. (2006) Evidence of genetic differentiation of the brackish water mysid *Neomysis integer* (Crustacea, Mysida) concordant with Pleistocene glaciations. *Vie et Milieu*, **56,** 15–22.

Rice, W.R. (1989) Analyzing tables of statistical tests. *Evolution*, **43,** 223–225.

Rogers, A.R. (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49,** 608–615.

Schneider, S., Roessli, D. & Excoffier, L. (2000) *Arlequin: a software for population genetics data analysis, version 2.00*. University of Geneva, Geneva.

Schubart, C.D., Diesel, R. & Blair Hedges, S. (1998) Rapid evolution to terrestrial life in Jamaican crabs. *Nature*, **393,** 363–365.

Slatkin, M. & Hudson, R.R. (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129,** 555–562.

Sunnucks, P., Wilson, A.C.C., Beheregaray, L.B., Zenger, K., French, J. & Taylor, A.C. (2000) SSCP is not so difficult: the application and utility of single-stranded conformation polymorphisms in evolutionary biology and molecular ecology. *Molecular Ecology*, **9,** 1699–1710.

Swofford, D.L. (1998) *PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4.0*. Sinauer Associates, Sunderland, MA.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G. & Cosson, J.-F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7,** 453–464.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123,** 585–589.

Tamura, K. & Nei, M. (1993) Estimations of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10,** 512–526.

Templeton, A.R. (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7,** 381–397.

Templeton, A.R. (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13,** 789–809.

Templeton, A.R., Boerwinkle, E. & Sing, C.F. (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, **117,** 343–351.

Templeton, A.R., Routman, E. & Phillips, C.A. (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, **140,** 767–782.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & 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.

Verspoor, E., McCarthy, E.M. & Knox, D. (1999) The phylogeography of European Atlantic salmon (*Salmo salar* L.) based on RFLP analysis of the ND1/16S RNA region of the mtDNA. *Biological Journal of the Linnean Society*, **68,** 129–146.

Virgilio, M. & Abbiati, M. (2004) Habitat discontinuity and genetic structure in populations of the estuarine species *Hediste diversicolor* (Polychaeta: Nereididae). *Estuarine, Coastal and Shelf Science*, **61,** 361–367.

Vlasblom, A.G. & Elgershuizen, J.H.B.W. (1977) Survival and oxygen consumption of *Praunus flexuosus* and *Neomysis integer*, and embryonic development of the latter species, in

different temperature and chlorinity combinations. *Netherlands Journal of Sea Research*, **11,** 305–315.

Wilke, T. & Davis, G.M. (2000) Infraspecific mitochondrial sequence diversity in *Hydrobia ulvae* and *Hydrobia ventrosa* (Hydrobiidae: Rissoidea: Gastropoda): do their different life histories affect biogeographic patterns and gene flow? *Biological Journal of the Linnean Society*, **70,** 89–105.

Wilke, T. & Pfenninger, M. (2002) Separating historic events from recurrent processes in cryptic species, phylogeography of mud snails (*Hydrobia* ssp.). *Molecular Ecology*, **11,** 1439–1451.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Variable nucleotide positions of the COI haplotypes observed in *Neomysis integer*.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

BIOSKETCH

Thomas Remerie is a post-doctoral researcher at the Marine Biology Section of Ghent University. This study derives from his doctoral research on the phylogeographical and population genetic structure of different mysid species along the European coasts. His current research interests include the molecular diversity of marine benthic invertebrates.

Editor: David Bellwood