Molecular Ecology (2006) **15**, 1699–1711 doi: 10.1111/j.1365-294X.2006.02881.x

Multiple introductions promote range expansion of the mollusc *Cyclope neritea* (Nassariidae) in France: evidence from mitochondrial sequence data

B. SIMON-BOUHET,*† P. GARCIA-MEUNIER† and F. VIARD*

**Evolution et Génétique des Populations Marines, UMR 7144 CNRS-UPMC, Station Biologique de Roscoff, B.P. 74, 29682 Roscoff cedex, France, †Laboratoire de Biologie et Environnement Marins, FRE 2727 CNRS-ULR, Avenue Michel Crépeau, 17042 La Rochelle, France*

Abstract

Since the 1970s, the nassariid gastropod *Cyclope neritea* has been extending its range north along the French Atlantic coasts from the Iberian Peninsula. This may be due to natural spread because of the recent warming of the northeastern Atlantic. However, human-mediated introductions related to shellfish culture may also be a probable explanation for this sudden range expansion. To examine these two hypotheses, we carried out a comprehensive study based on mitochondrial gene sequences (cytochrome oxidase I) of the five recently colonized French bays as well as 14 populations located in the recognized native range of the species. From a total of 594 individuals, we observed 29 haplotypes to split into three divergent clades. In the native range, we observed a low molecular diversity, strong genetic structure and agreement between geography and gene genealogies. Along the French coasts, we observed the opposite: high genetic diversity and low genetic structure. Our results show that recurrent human-mediated introductions from several geographical areas in the native range may be a source for the French Atlantic populations. However, despite the low dispersal ability of *C. neritea*, the isolation-by-distance pattern in France suggested that this gastropod may have been present (although unnoticed) on the French Atlantic coasts before the 1970s. As *C. neritea* shows characteristics of a cryptogenic species, the classification of Atlantic populations as either native or introduced is not straightforward. Cryptogenic species should be studied further to determine the status of new populations close to their recognized native range.

Keywords: biological introductions, cryptogenic species, *Cyclope neritea*, cytochrome oxidase I, marine gastropod, range limit

Received 16 September 2005; revision accepted 14 December 2005

Introduction

In marine habitats and especially coastal areas, the rate of biological invasions has increased considerably in the last few decades, altering community dynamics (Dukes & Mooney 1999; Grosholz 2002). Nevertheless, the processes underlying the sustainable settlement of nonindigeneous species (NIS) in coastal habitats are still not fully known (Carlton 1996; Grosholz 2002). Introduction processes are obviously highly related to the increase in transport in ship ballast waters as well as to intentional or accidental releases for aquaculture and fisheries (Carlton & Geller 1993; Cohen & Carlton 1998). The sustainable settlement of NIS in coastal habitats may also benefit from other facets of global change. Fast environmental changes such as habitat or climatic changes may indeed alter native communities and weaken the competitive ability of indigenous species, thus facilitating the settlement of introduced species (Dukes & Mooney 1999; Carlton 2000; Sax & Brown 2000). For example, Stachowicz et al. (2002) recently demonstrated that changes in the maximum and minimum environmental temperatures may have promoted the establishment of two invasive ascidians at the expense of native species. Temperature is known to be a critical factor influencing the

Correspondence: Benoît Simon-Bouhet, Fax: +33 546500294; E-mail: bsimonbo@univ-lr.fr, bsimonbo@sb-roscoff.fr

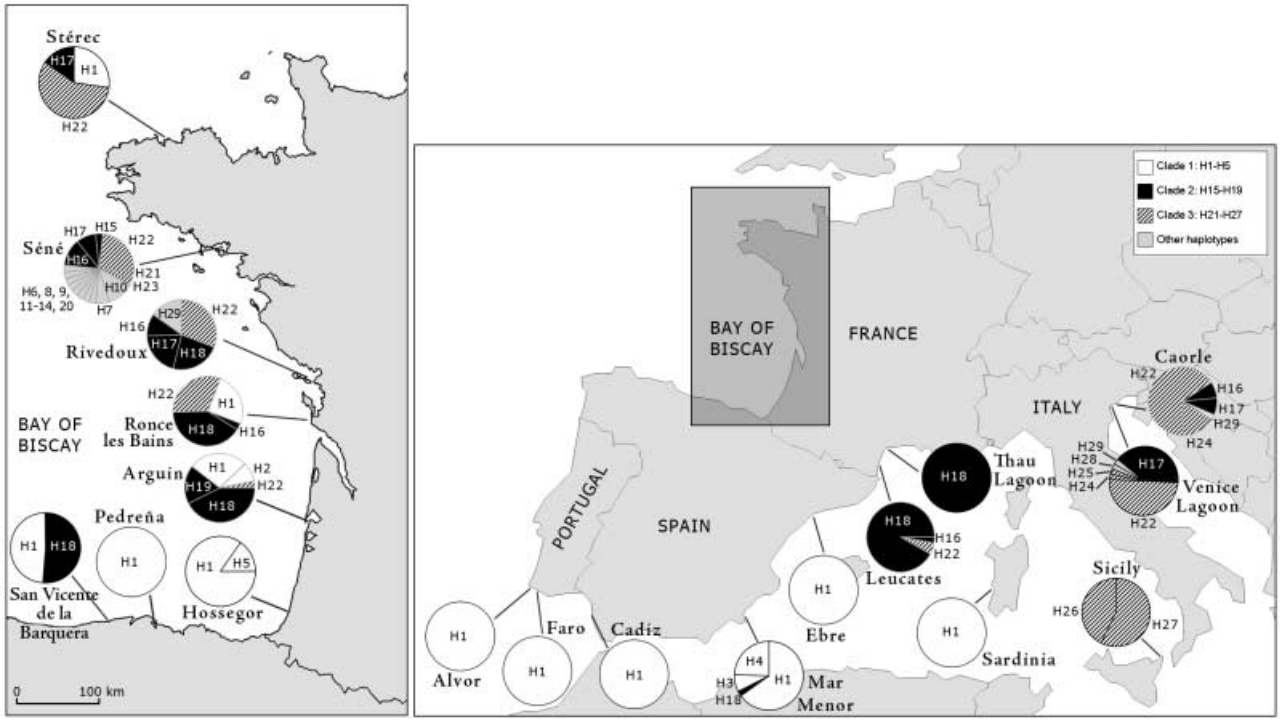

Fig. 1 Location of the study populations and haplotype frequencies at the population level. Native populations are shown on the right whereas FAEC populations and populations of unknown status are shown on the detailed map of the Bay of Biscay (left side). For each population, haplotype names are indicated and colour/figures (see box on the upper-right corner) in pie charts refer to the defined clades in Fig. 2.

natural distribution of animal and plant species: a wide array of species colonized western and northern Europe from the Iberian Peninsula after the ice ages (Comes & Kadereit 1998; Taberlet et al. 1998; or Gomez & Lunt, in press) and species located on both sides of a biogeographic boundary can extend or restrict their natural range in response to local changes in environmental conditions (Gaylord & Gaines 2000). Natural processes might thus interact with human-mediated introductions.

The Bay of Biscay (northeastern Atlantic, Fig. 1) is an interesting area to study this phenomenon because the mean temperature of this geographical area has increased by 0.6–0.8 °C per decade during the last 30 years (IPCC 2001). This rapid and recent warming has promoted the displacement of mollusc species within this bay (e.g. *Macoma balthica*, Hummel *et al*. 2000). Moreover, there are numerous and important shellfish culture sites present along the coasts of the Bay of Biscay and the exchange of commercial species between distant shell-farming sites have been shown to be an efficient way of primary and secondary introductions of NIS, including molluscs (e.g. *Crepidula fornicata*, see References in Dupont *et al*. 2003), in new biotas (Wolff & Reise 2002). Therefore, changes in temperature and the transport of oysters may both induce community changes in the Bay of Biscay. The recent expansion of the distribution range of the mollusc *Cyclope neritea* makes this species an appropriate model to examine the role played by these two processes. This nassariid gastropod is native to the Mediterranean and Black seas, and to the Atlantic coasts of Morocco, southern Spain and Portugal. Until the 1950s, only discontinuous and ephemeral populations were recorded along the northern Atlantic coasts of Spain and the southern French Basque coasts, with the Basque country being defined as the northern edge of its natural range (for a detailed description, see Sauriau 1991). Since 1976, new populations have been recorded gradually moving from south to north along the French Atlantic coasts: first in Arcachon Bay, then in 1983–1984, in the Marennes-Oléron Bay, the Île de Ré, and the Gulf of Morbihan, and more recently from 2000, along English Channel coasts (see References in Bachelet *et al*. 2004). *C. neritea* is rather uncommon because new populations of this species appeared just at the edge of its previously recognized natural range. On the one hand, the gradual appearance of *C. neritea* populations from south to north suggests a natural spread towards north during the past 30 years due to environmental changes. On the other hand, a natural spread over such a short period of time seems unlikely given the limited dispersal ability of this species. As a direct developer, *C. neritea* has no planktonic larval stage (Gomoiu 1964) and is restricted to a very particular substrate (i.e. sheltered spots, characterized by sandy-muddy substrates). Therefore, the discontinuous habitat due to the rocky shores along the southern French Atlantic coasts may slow down natural migrations (Sauriau

1991). Moreover, all the newly established French populations of *C. neritea* were found near oysters beds. Therefore, the appearance of *C. neritea* along the French Atlantic and English Channel (FAEC) coasts may be due to human-mediated introductions via oyster exchanges between Mediterranean and Atlantic shellfish culture sites (Boulhic & Tardy 1986; Pigeot 1988; Le Roux *et al*. 1988; Sauriau 1991).

In a previous study, Bachelet *et al*. (2004) focused on Arcachon Bay, which was the first site in which *C. neritea* was observed outside of its recognized native range. Mitochondrial analyses revealed the coexistence of three highly divergent lineages in this population, reinforcing the idea of accidental introductions due to human-mediated activities. However, based on this single population, the authors could not differentiate between a single introduction from a highly diversified source population or multiple and recurrent introductions from more poorly diversified but highly divergent source populations. Moreover, by analysing only one population, Bachelet *et al*. (2004) could not determine the colonization history of *C. neritea* along the French coasts. Here, we consider: (i) whether the population inhabiting Arcachon Bay exhibits a pattern congruent with a single-source or a multiple-source introduction scenario. As accidental NIS introductions along the FAEC coasts are mainly related to oyster exchanges (Goulletquer *et al*. 2002), we focused our native range sampling on areas where important shellfish cultures occur (i.e. the Iberian Penninsula, the French Mediterranean coasts and the Northern Adriatic Sea); (ii) whether the recently established FAEC populations all share a common mechanism for their establishment and and whether the FAEC populations are genetically related. In particular, we tested for a genetic isolation pattern between the known primary site of introduction (i.e. Arcachon Bay, colonized in 1976) and the more recently and gradually colonized sites along the French coasts; and (iii) whether we can identify the status (either native or introduced) of the populations located at the edge of the recognized natural range of *C. neritea* (i.e. the northern part of the Iberian Peninsula and French Basque Country).

Mitochondrial gene studies were shown to be reliable for phylogeography studies of native species as well as for analysing the mechanisms involved in the settlement of introduced species (Kolbe *et al*. 2004; Voisin *et al*. 2005). Moreover, because the effective population size for mitochondrial genes is four-times lower than for nuclear markers, any genetic drift effects are emphasized and isolation by distance patterns should be easier to detect (Diaz-Almela *et al*. 2004). Therefore, we carried out a comprehensive sampling of *C. neritea* along the FAEC coasts (five populations) as well as 11 populations from the native range and three populations located at the edge of the previously recognized range of *C. neritea*. Classical population genetics analyses were combined with intraspecific phylogenetic approaches by analysing a fragment of the cytochrome oxidase I mitochondrial gene (*COI*, 533 bp) in 594 individuals.

Materials and methods

Sampling

A survey and exhaustive sampling of French Atlantic and English Channel (FAEC) coasts was carried out between March 2002 and June 2003 in each bay in which *Cyclope neritea* has been recorded since the 1970s (Table 1, Fig. 1). We believe this sampling fits exactly the real distribution of the species in its current FAEC range. We collected five FAEC populations: (i) Arguin, the primary site of introduction in 1983 in Arcachon Bay (which was previously studied by Bachelet *et al*. 2004), (ii) Ronce les Bains (Marennes-Oléron Bay, Sauriau 1989), Rivedoux (Île de Ré, Tardy *et al*. 1985) and Séné (Gulf of Morbihan, Le Roux *et al*. 1988), three bays in which *C. neritea* was first observed in 1983–84; and (iii) Stérec (Morlaix Bay, English Channel), the most northerly population, known on French coasts since the end of the 1990s. We also sampled three populations located at the northern edge of the recognized natural range (French Basque Country) between December 2003 and March 2004: (i) Hossegor (French Basque Country), where *C. neritea* went unnoticed until our sampling and located at 50 km North of Socoa where the species was first recorded in 1950 (Kisch 1950) and (ii) two northern Spanish populations, Pedreña and San Vicente de la Barquera. We also sampled 11 populations between March and December 2004 in the recognized native range of the species to allow comparisons with the FAEC populations (Table 1; Fig. 1): (i) along the Atlantic coasts of Portugal and Spain (3 populations), (ii) along the Mediterranean coasts of Spain, France and Italy (6 populations) and (iii) along the coasts of the Adriatic Sea (Italy, 2 populations). Thau Lagoon (France) and Faro (Portugal) had been previously studied by Bachelet *et al*. (2004). For each population, we collected adult specimens, which were stored in 95% ethanol before DNA extraction.

DNA extraction, amplification and sequencing

We followed the protocol detailed in Bachelet *et al*. (2004). Briefly, total DNA was extracted from less than 15 mg of foot muscle using the DNeasy Tissue Kit according to the manufacturer's protocol (QIAGEN). Amplification of the *COI* gene was carried out using specific primers and the polymerase chain reaction (PCR) products were directly sequenced using ABI PRISM BigDye Terminators version 3.0 Cycle Sequencing Kit following the manufacturer's protocol (Applied Biosystems). Both strands were sequenced for each individual using an ABI PRISM-3100 Automated DNA Sequencer (PerkinElmer Applied Biosystems) and sequence data were aligned using CLUSTAL W (Thompson *et al*. 1994).

Table 1 Geographic locations of the 19 study populations. Numbers of individuals (N) are indicated together with genetic diversity estimates (N_H, number of haplotypes; S, number of segregating sites; H_E, haplotypic diversity; π, nucleotidic diversity)

Population	Location	Country (Sea-Ocean)	N	N_H	S	H_E	$\pi \times 10^2$
Recently established populations (set 1)							
Stérec	Morlaix Bay, N. Brittany	France (EC)	33	3	18	0.589	1.371
Séné	Gulf of Morbihan, S. Brittany	France (A)	32	16	14	0.911	1.014
Rivedoux	Île de Ré	France (A)	39	5	21	0.796	1.503
Ronce les bains	Marenne-Oléron Bay	France (A)	32	4	19	0.696	1.498
Arguin	Arcachon Bay	France (A)	33	5	19	0.725	1.234
Mean for recently established populations			33.8	6.6	18.2	0.743	1.324
(Standard deviation)			(2.9)	(5.3)	(2.6)	(0.120)	(0.205)
Total over all recently established populations			169	21	27	0.833	1.518
Populations of unknown status (set 2)							
Hossegor	Basque Country	France (A)	33	2	1	0.265	0.050
Pedreña	Santander Bay, Cantabrica	Spain (A)	3	1	0	0.000	0.000
San Vicente de la Barquera	Cantabrica	Spain (A)	37	2	12	0.514	1.156
Mean for populations with unknown status			24.3	1.7	4.3	0.260	0.402
(Standard deviation)			(18.6)	(0.6)	(6.7)	(0.257)	(0.653)
Total over all populations with unknown status			73	3	13	0.484	0.903
Native populations (set 3)							
Alvor	Algarve	Portugal (A)	39	1	0	0.000	0.000
Faro	Ria Formosa, Algarve	Portugal (A)	31	1	0	0.000	0.000
Cadiz	Andalucia	Spain (A)	36	1	0	0.000	0.000
Mar Menor	Murcia	Spain (MS)	37	4	13	0.527	0.216
Ebre	Cataluña	Spain (MS)	33	1	0	0.000	0.000
Leucates	Languedoc-Roussillon	France (MS)	37	3	13	0.156	0.255
Thau Lagoon	Languedoc-Roussillon	France (MS)	29	1	0	0.000	0.000
Sardinia	Golfo di Oristano	Italy (MS)	25	1	0	0.000	0.000
Sicily	Punta del Faro (strait of Messina)	Italy (MS)	7	2	2	0.571	0.214
Venice Lagoon	Veneto	Italy (AS)	40	6	25	0.603	1.251
Caorle	Veneto	Italy (AS)	38	5	21	0.543	0.769
Mean for native populations			32	2.4	6.7	0.218	0.246
(Standard deviation)			(9.5)	(1.9)	(9.5)	(0.276)	(0.406)
Total over all native populations			352	13	35	0.661	1.378
Total			594	29	36	0.734	1.479

EC, English Channel; A, Atlantic; MS, Mediterranean Sea; AS, Adriatic Sea.
Note: populations were sampled by the authors except for Sardinia, Sicily and Venice Lagoon for which samples were provided by Jeroen Jansen, Francesco Patti and Davide Tagliapietra, respectively.

Sequence analysis

For each population, we examined the genetic diversity by calculating the number of mitochondrial haplotypes, N_H, and polymorphic sites, S, the nucleotide diversity, θ (i.e. average number of nucleotide differences between pairs of sequences), and the gene diversity, H_E (Nei 1987), using DNASP version 3.53 software (Rozas & Rozas 1999). We defined three population sets (Table 1): recently established populations (i.e. populations recorded since the 1970s along the FAEC coasts, set 1), populations of unknown status (i.e. populations located at the edge of the recognized native range of the species, set 2) and native populations (i.e. populations from the recognized native range of the species, set 3). For each set, we calculated the mean (and associated standard deviations) and overall N_H, S, H_E and θ-values.

We investigated the overall or regional population genetic structure by performing analyses of molecular variance (AMOVA, Excoffier *et al*. 1992) as implemented in ARLEQUIN version 2.0. software (Schneider *et al*. 2001). Fixation indices (Wright 1951) analogous to F_{ST}, F_{CT} and F_{SC} parameters, were calculated to analyse the genetic differentiation between populations over the whole native range (Φ_{ST}), among groups of populations (regions) within the native range (Φ_{CT}) and among populations within groups (regions; Φ_{SC}) as defined in Excoffier *et al*. (1992). Φ-statistics take into account both the haplotype frequencies and the molecular distances between haplotypes. We assessed the statistical

significance of the $\hat{\Phi}$ indices using a nonparametric permutation procedure implemented in ARLEQUIN version 2.0. We used the isolation by distance (IBD) model to examine the genetic relationships between FAEC populations. IBD was determined by plotting genetic distance (i.e. $\hat{F}_{ST}/(1-\hat{F}_{ST})$, Rousset 1997) against linear geographic distance (in kilometres) measured following the coastline. These distances were calculated for each FAEC population pair and we used a Mantel test to check a positive correlation between matrices of pairwise geographic and genetic distances using IBD 1.5 software (Bohonak 2002).

Given the pattern that was observed by Bachelet *et al*. (2004) in Arguin, we had to discriminate between two introduction processes: (i) a single introduction event from a highly diversified source population or (ii) multiple introductions from slightly diversified but highly divergent source populations. Therefore, as well as molecular diversity, we also analysed the haplotype distribution across populations after determining their genealogical relationships by a haplotypic network (Cassens *et al*. 2003). The network was constructed using a median-joining algorithm implemented in NETWORK 4.0.0.0 software (Bandelt *et al*. 1999). We then applied a maximum parsimony algorithm to simplify the complex branching pattern and to represent all the most parsimonious intraspecific phylogenies. We illustrated similarities between FAEC and native populations by building a population tree using MEGA version 3.0 software (Kumar *et al*. 2004): pairwise genetic distances matrices were calculated using the Kimura 2-parameters model (Kimura 1980) and phylogenetic reconstructions were carried out using a neighbour-joining algorithm (Saitou & Nei 1987). Finally, for FAEC populations, we calculated the distribution of the observed number of pairwise nucleotide site differences (i.e. mismatch distribution) in DNASP version 3.53 to examine the mixing of evolutionary divergent lineages at the population level. Given the low resolving power of Tajima's *D*-test (Galtier *et al*. 2000), we carried out tests for departure from mutation–drift equilibrium under an infinite-site model following the method of Depaulis & Veuille (1998). This test (*K*-test) is based on the comparison between the observed number of haplotypes (N_H) and the expected number under equilibrium (N_{EH}) given the sample size (N) and the observed number of segregating sites (S). It is particularly relevant to our study because it can detect recent population admixture (Hamblin & Veuille 1999). We used a computer program provided by M. H. Muller (2002) to compute confidence intervals and one-tailed P values for N_H.

Results

Molecular diversity over the dataset

A high molecular diversity was found over the whole study although not evenly distributed across population sets. We analysed a 533-bp fragment of the *COI* gene and observed only synonymous substitutions (Appendix). Over the whole data set ($N = 594$ individuals), we identified 36 polymorphic sites defining 29 haplotypes (Table 1, Appendix; GenBank Accession nos: AY789970–AY789991 and DQ087210–DQ087216). The haplotype distribution at the population level is shown in Fig. 1. Over the whole data set, we found three haplotypes (H1, H18 and H22) at a high frequency, $f = 0.443$, 0.200 and 0.163, respectively (sum = 0.806) compared to the others ($0.002 < f < 0.059$). We observed large differences in gene and molecular diversities for the three population sets (Table 1): gene and molecular diversity indices were higher in recently established populations (set 1) than in populations of unknown status (set 2) or from the native range (set 3). For example, H_E ranged from 0.589 to 0.911 (mean = 0.743) in set 1, whereas it ranged from 0 to 0.514 (mean = 0.260) in set 2 and from 0 to 0.603 (mean = 0.218) in set 3. We observed the same for N_H, S and π. All populations presented high and similar polymorphism levels in set 1, whereas the other two sets showed monomorphic populations (7 out of 14). Molecular and genetic diversities were not distributed similarly across populations according to the set: overall and mean values for diversity indices were closer in set 1 than in sets 2 and 3. For example, overall value for π for set 3 was six times higher than mean value (1.38×10^2 and 0.25×10^2 respectively) whereas overall and mean values were almost identical for set 1 (1.52×10^2 and 1.32×10^2 respectively). Finally, in the recently established populations, only Séné differed in the nucleotidic and haplotypic diversity estimates: this population showed the highest number of haplotypes and highest haplotypic diversity ($N_H = 16$, $H_E = 0.911$) whereas the number of segregating sites and nucleotide diversity ($S = 14$, $\pi = 0.01024$) were the lowest of the FAEC populations.

Genealogical relationships and geographic distribution of haplotypes

Gene lineages were clearly split across regions within the native range. The most likely (i.e. parsimonious) genealogies between haplotypes are shown in Fig. 2. We found that the three most frequent haplotypes (i.e. H1, H18 and H22) belonged to three main clades (clade 1: H1–H5, clade 2: H15–H19 and clade 3: H21–H27). The genetic distances between these clades were large (11–20 mutational steps) compared to the genetic distances within clades (1–6 mutational steps). We found that only two haplotypes, H28 and H29, were distant from these three clades (9–15 mutational steps according to the cluster). The 10 other haplotypes, found at low frequencies, were close to each other and represented an unresolved torso in the middle of the network. The overall structure of the network was

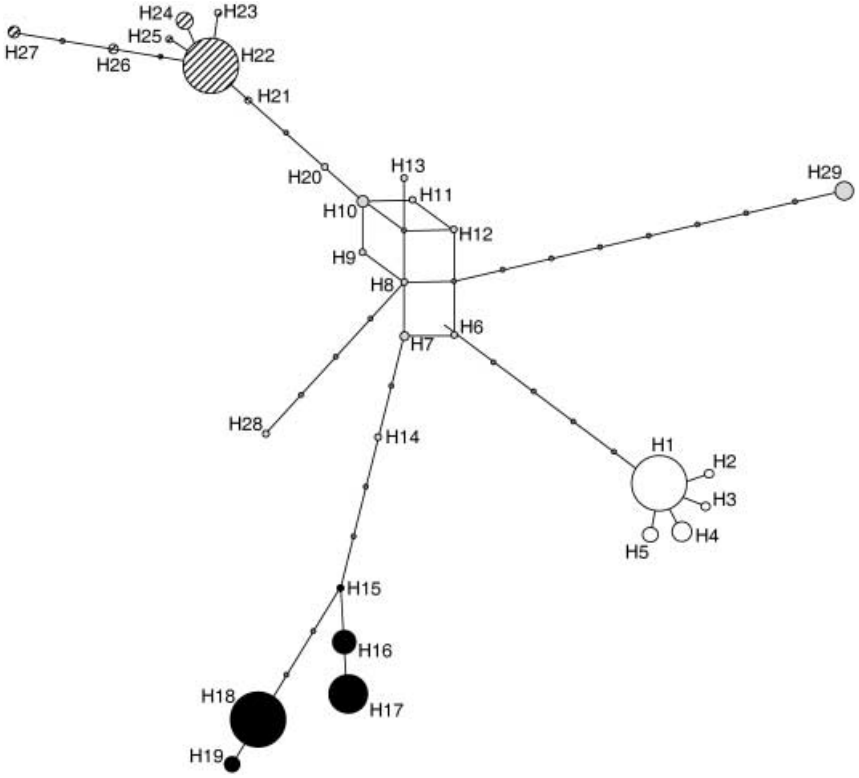

Fig. 2 Haplotypic network showing the phylogenetic relationships between haplotypes. Circle sizes are proportional to haplotype frequencies over the whole data set. The three main clades are identified by different colours (plain black, plain white and black and white hatched circles). (●): inferred mutational step.

consistent with the geographic location of the native populations (Fig. 1). Within the native range, we could clearly identify three subgroups of populations corresponding to three geographic regions: (i) Southwestern populations (i.e. Iberian Peninsula plus Sardinia) were characterized almost exclusively by haplotypes belonging to clade 1; (ii) French Mediterranean populations were mainly characterized by haplotypes belonging to clade 2; and (iii) the Eastern populations (Sicily and Adriatic Sea populations) were mainly characterized by haplotypes belonging to clade 3. This apparent genetic clustering within the natural range was statistically significant according to AMOVA: the overall high and significant genetic structure ($\Phi_{ST} = 0.869$, $P < 10^{-3}$) was due to a large genetic differentiation between these three regions ($\Phi_{CT} = 0.846$, $P < 10^{-3}$) compared to the low genetic differentiation among the populations within these regions ($\Phi_{SC} = 0.152$, $P < 10^{-3}$). The correspondence between genetic and geographic structures is well illustrated by a neighbour-joining tree (Fig. 3), in which the three subgroups are clearly separated. We found that the populations of unknown status were closely related to the Iberian Peninsula populations. Conversely, the recently established populations were distributed over the whole tree suggesting several origins for the FAEC populations.

Genetic pattern along the French Atlantic and English Channel coasts

Contrary to the pattern observed among native populations, we found the genetic differentiation between the FAEC populations to be low ($\Phi_{ST} = 0.135$) but significant and associated with a high population genetic diversity. The five recently established populations exhibited haplotypes belonging to two or three of the clades as shown by the network analysis (Figs 1 and 2). Mismatch distribution analyses corroborated this co-occurrence of individuals having evolutionary divergent haplotypes (Fig. 4). Arguin, Ronces, Rivedoux and Stérec all showed at least two major peaks characterizing two or more groups of genetically divergent haplotypes co-occurring at the population level. In these four populations, the observed number of haplotypes (N_H) was significantly lower than expected under mutation–drift equilibrium (N_{EH}, *K*-tests in Fig. 4), suggesting a recent admixture (Hamblin & Veuille 1999). However, in Séné we observed a larger number of haplotypes that differed from each other by a large range of mutational steps (from 1 to 14), generating a more uniform mismatch distribution. Moreover, in this population was observed a larger number of haplotypes than expected (*K*-test, Fig. 4), suggesting a recent population expansion

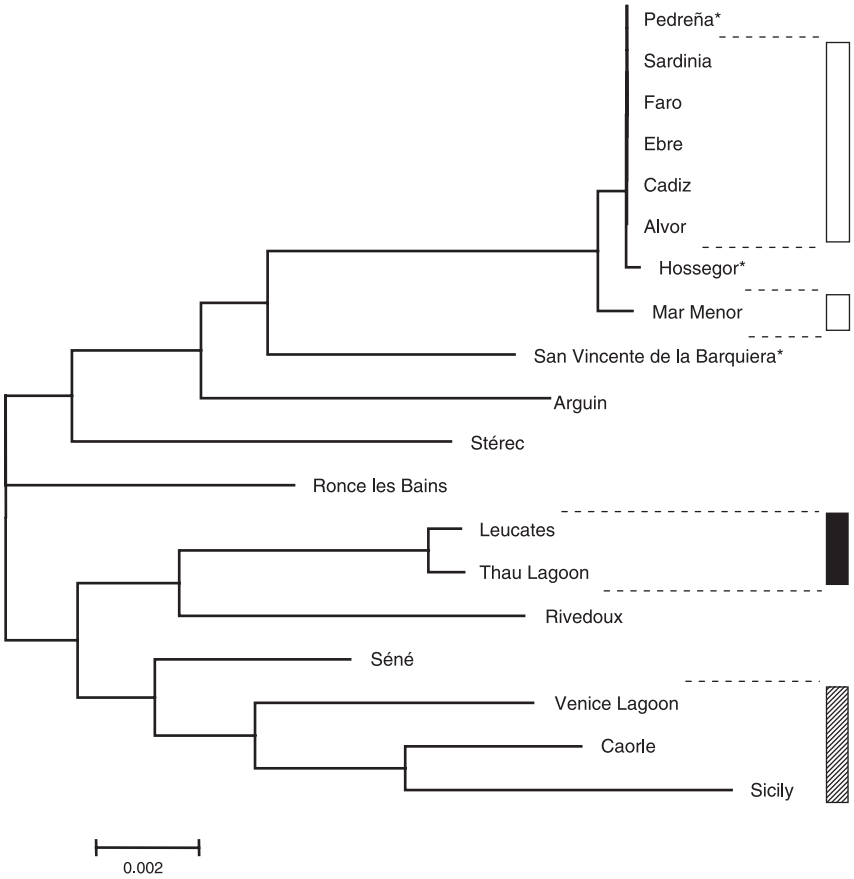

Fig. 3 Population neighbour-joining tree. Genetic distances between populations were estimated using the Kimura 2-parameters model. Coloured rectangles on the right of the tree show the main clade (as defined in Fig. 2) found in each native population. '*': populations of unknown status.

or ancient balanced polymorphism (Depaulis & Veuille 1998). Finally, within the FAEC populations, we observed an increase in the pairwise F_{ST} estimates when we considered more distant populations, which resulted in a significant pattern of isolation by distance ($P = 0.016$, Fig. 5).

Discussion

Recurrent introductions from multiple sources

Although the overall level of mitochondrial polymorphism observed in our study of *Cyclope neritea* was comparable with other studies of the *COI* sequences below the species level [e.g. *Scylla serrata*, Gopurenko *et al*. (1999); *Patelloida profunda*, Kirkendale & Meyer (2004); *Littorina* sp., Wilding *et al*. (2000)], we observed a marked difference when comparing native and recently established populations. We noticed a low within-population genetic diversity associated with a strong genetic structure in the native range of *C. neritea*, whereas the FAEC populations exhibited a large within-population genetic diversity associated with a moderate genetic structure. With the exception of Séné, our study revealed that FAEC populations did not originate from a unique introduction event from a highly diversified native population but primarily from recurrent introductions from multiple sources. More specifically, in these FAEC populations, we observed large haplotypic and nucleotidic diversities and the co-occurrence of two or three groups of highly divergent haplotypes. Multipeak mismatch curves and significant *K*-tests confirmed that these four FAEC populations are not in demographic equilibrium and that admixture events (i.e. mixing of individuals from populations that have evolved independently) have occurred. None of the 11 native populations exhibited a mixture of the haplotypes belonging to the three clades identified in most of the FAEC populations (Fig. 1). Moreover, we found all native populations were much less diversified than the recently established populations, with the most polymorphic native population (Venice Lagoon) being less diversified than any of the recently established populations. In conclusion, the native populations showed a phylogeographic pattern with a strong consistency between their geographic locations and their genetic identity, which was confirmed by AMOVA and the neighbour-joining tree. At the population level, the recently established range has an increased genetic diversity compared to the native range, and, on a more regional scale, the strong genetic structure of the natural range had been erased in the recently colonized areas.

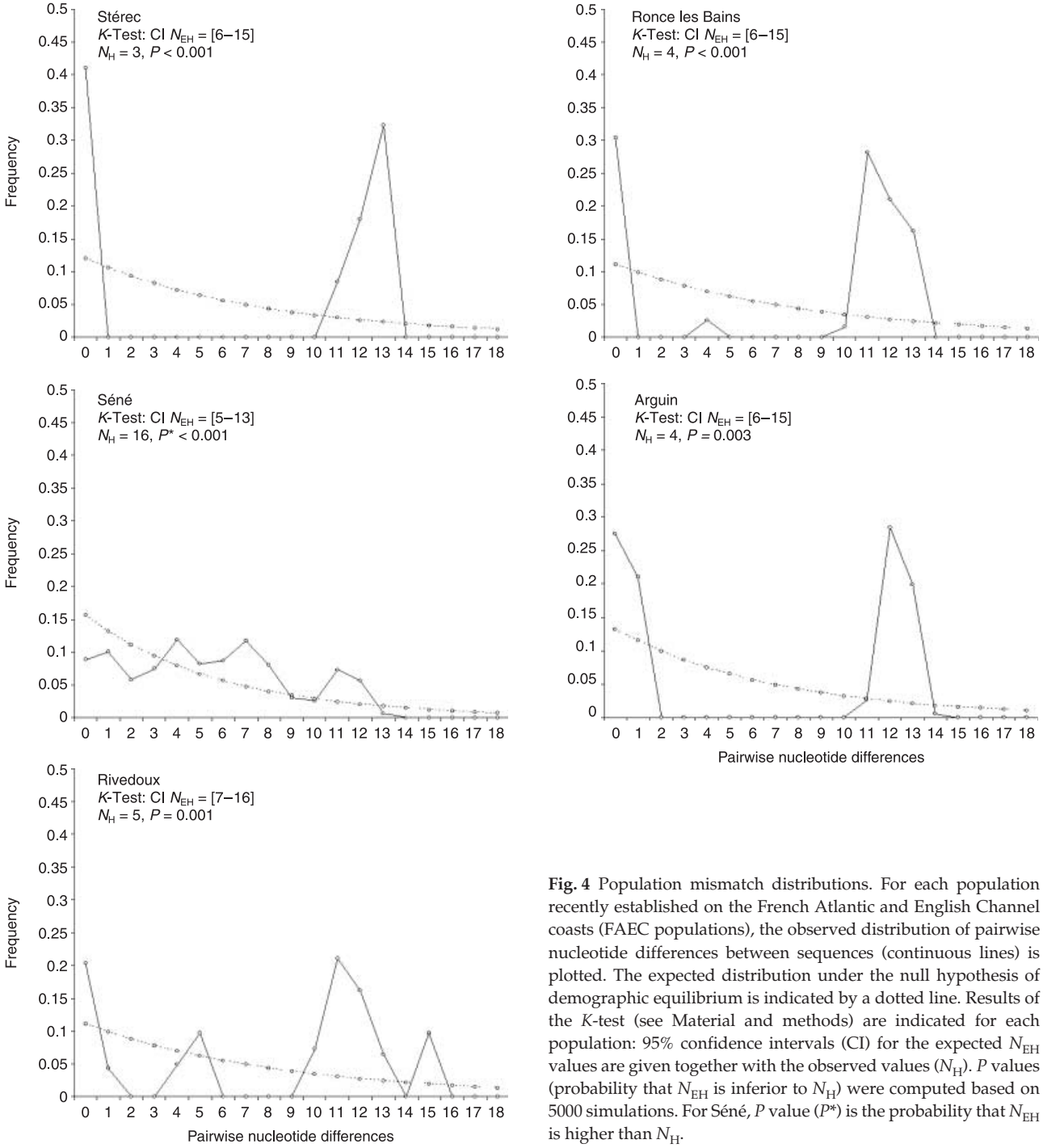

Fig. 4 Population mismatch distributions. For each population recently established on the French Atlantic and English Channel coasts (FAEC populations), the observed distribution of pairwise nucleotide differences between sequences (continuous lines) is plotted. The expected distribution under the null hypothesis of demographic equilibrium is indicated by a dotted line. Results of the *K*-test (see Material and methods) are indicated for each population: 95% confidence intervals (CI) for the expected N_{EH} values are given together with the observed values (N_H). *P* values (probability that N_{EH} is inferior to N_H) were computed based on 5000 simulations. For Séné, *P* value (*P**) is the probability that N_{EH} is higher than N_H.

As suggested earlier for Arguin (Bachelet *et al*. 2004), the results of our study shows that *C. neritea* populations recorded along the French coasts have been introduced since the 1970s and also strongly suggests that the introduction involved recurrent introduction events from several poorly diversified but highly divergent native populations. This is consistent with many genetic studies of terrestrial, freshwater and marine nonindigenous species (NIS) that failed to show the expected decrease of genetic polymorphism for numerous introduced species [Poaceae, Novak *et al*. (1993), brown mussel, Holland (2001), thiarid snail, Facon *et al*. (2003), slipper limpet, Dupont *et al*. (2003), Cuban lizard, Kolbe *et al*. (2004) or Japanese alga, Voisin *et al*. (2005)]. As observed by Kolbe *et al*. (2004), when native populations

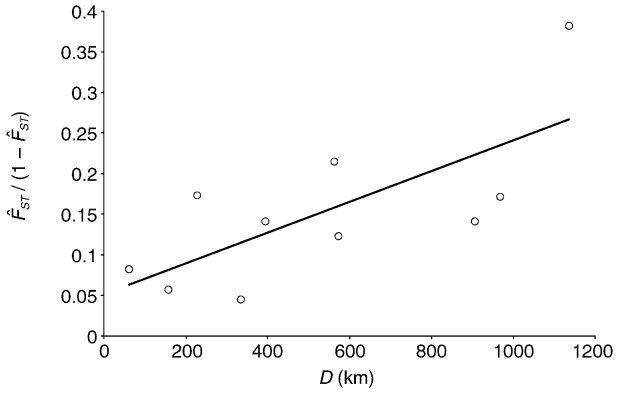

Fig. 5 Isolation by distance among the five French Atlantic and English Channel (FAEC) populations. The genetic distance $\hat{F}_{ST}/(1-\hat{F}_{ST})$ for each population pair is plotted against geographic distance (D) in kilometres (see Materials and methods). The P value associated with a Mantel test for a positive correlation between genetic and geographic distances is given.

are monomorphic, the occurrence of multiple introductions can promote the settlement of NIS by transforming among-population variation in native ranges to within-population variation in introduced areas.

Alternative vector of introduction for Séné

Oyster exchanges, which are known to be responsible for numerous NIS introductions along the Atlantic coasts of Europe (Goulletquer *et al*. 2002; Wolff & Reise 2002), may have played a major role in the colonization of the FAEC coasts by *C. neritea* by introducing, near oyster beds, many individuals from various genetically differentiated populations. Such mechanisms and vectors have been suggested for other coastal marine introductions (see for instance *Ocinebrellus inornatus*, Martel *et al*. 2004). However, other shellfish cultures may also have contributed to the successful settlement of *C. neritea* along the FAEC coasts. This is exemplified by the particular genetic pattern of Séné, which showed: (i) a higher number of haplotypes than expected under the equilibrium hypothesis (*K*-test, Fig. 4); (ii) a position close to the populations from the Adriatic Sea in the neighbour-joining tree (Fig. 3); and (iii) the only nonsignificant genetic differentiation (55 pairwise Φ_{ST} comparisons) between the five recently established and the 11 native populations of our data set (i.e. between Séné and Venice Lagoon $\Phi_{ST} = 0.045$, $P = 0.072$; data not shown). A limited number of genetically diversified source populations, one of which having recently undergone a large demographic expansion, would seem a more probable explanation than the numerous introductions from less diversified but highly divergent populations that we observed in the other FAEC populations. This is consistent with the frequent exchanges of Japanese Manila clams, *Tapes philippinarum*, occurring between the Venice Lagoon and shellfish culture sites of the Gulf of Morbihan, where Séné is located [J.-F. Auvray (SATMAR), personal communication]. This may be the principal or even the unique vector for the introduction of *C. neritea* in this area. This result illustrates the difficulty in generalizing the identification of a vector of introduction for a given NIS over its full range of introduction. As recently demonstrated for other marine NIS (e.g. the alga *Undaria pinnatifida*; Voisin *et al*. 2005), comprehensive studies of introduced populations are necessary for an inventory of all the factors/vectors that may promote the successful settlement of NIS. Nevertheless, all the French populations are characterized by at least one common feature: a large genetic diversity compared to populations from the native range. This has also been demonstrated in numerous successful NIS, suggesting the importance of propagule pressure for invasion success (Williamson 1996; Sax & Brown 2000).

Combination of natural spread and human-mediated introductions along the FAEC coasts

Human-mediated introductions associated with shellfish (i.e. oysters and/or Manila clams) transport are likely to have played a major role in the primary colonization process by *C. neritea* of the FAEC coasts. Recently established populations gradually appeared from south and moved northward and an isolation-by-distance pattern could be clearly seen between these populations (Fig. 5). This IBD may be explained either by secondary introductions between shellfish culture sites or by a natural spread from either the Iberian Peninsula or the primary site of introduction (i.e. Arguin). As we have no precise monitoring data about oyster exchanges between shellfish culture sites and contamination by *C. neritea*, we cannot definitively rule out secondary introductions. However, human-mediated exchanges between French shellfish culture sites occur randomly between very distant and noncontinuous locations. In particular, all French oysters farms (from the English Channel, Atlantic and even Mediterranean culture sites) rely on oyster spats only produced in the bays of Arcachon and Marennes-Oléron (Goulletquer & Heral 1997). If random human-mediated dispersal was the only process affecting *C. neritea* populations, the haplotypes should have been randomly mixed across the populations resulting in the absence of the apparent step-by-step process revealed by the IBD pattern. Given the low dispersal ability of the species, it is appealing to suggest a natural spread to explain the observed IBD although it is unlikely that such a pattern could have been produced over only 20–30 years. Nevertheless, natural spread could have occurred over a longer time period if the *C. neritea* populations had been present at low densities, and consequently unnoticed, along the FAEC coasts long before 1976. The patchy distribution of *C. neritea*

populations, even in its natural range, makes it difficult to detect even when population densities are not low.

Cyclope neritea can be classified as a cryptogenic species as defined by Carlton (1996). Such 'not demonstrably native or introduced' species probably represent a large part of the newly recorded species (Carlton 1996). Ancestral *C. neritea* populations may have been present long before 1976 either because of natural migrations from the most northern native populations or because of ancient and unrecognized accidental introductions. For example, the oyster *Ostrea edulis* has been exploited in France since the early 20th century, with exchanges occurring between the Thau lagoon and the Bay of Marennes-Oléron (Hinard & Lambert 1928). Sax & Brown (2000) recently demonstrated that species that suddenly become abundant and widespread often do so only after having failed to establish from earlier multiple natural or accidental introductions. The recent sustainable settlement of large and detectable *C. neritea* populations (from 400 to > 1600 ind./m² in Rivedoux, Tardy *et al*. 1985) may have been promoted by two factors: (i) an accelerated rate of accidental introductions in the 1970s due to intensified oyster transfers between Mediterranean and Atlantic shellfish farms and (ii) the rapid increase of temperature in the Bay of Biscay. By weakening native communities, both factors (Dukes & Mooney 1999) may have favoured the settlement of introduced individuals and migrants coming from the south, thus creating a demographic reinforcement of cryptic *C. neritea* populations along the French Atlantic coasts.

Populations at the edge: introduction vs. expansion

The three populations located at the edge of the previously recognized native range of the species, namely San Vicente de la Barquera (SVB), Pedreña and Hossegor, were classified as populations of unknown status (Table 1). Although located at the northern edge of its distributional range, this area is classified as part of the natural range of *C. neritea*, as its presence has been regularly reported in the past (Sauriau 1991). Moreover, shellfish have been cultured in this area for a long time and the populations sampled were all located near oyster farming sites (*C. neritea* populations were not found elsewhere, B. Simon-Bouhet, personal observation). Therefore, these populations could be either natural or introduced, justifying the classification of *C. neritea* as a cryptogenic species. Some of our results support a native status for these study populations because, for example, the three populations were genetically close to other Iberian Peninsula populations (Fig. 3). However, the three study populations did not share exactly the same characteristics. Hossegor, which is located at the very edge of the natural range showed features that we observed in the set of native populations (set 3, Table 1): (i) the haplotypic and nucleotidic diversities were closer to the values observed in set 3 than in recently established populations and (ii) the two haplotypes (H1 and H5) found in this population belonged to the same clade as native populations (clade 1; Figs 1 and 2). Conversely, the SVB population showed features that we observed in most of the FAEC populations: high genetic diversity and the occurrence of two highly divergent haplotypes, namely H1 and H18 (found in clade 1 and 2, respectively) that were never found together in native populations. In Pedreña, the low sampling size prevents us from carrying out fine analysis. We found many more empty shells than in all other populations and we caught only three living individuals, suggesting that in the past this population may have been large but is now declining. Our sampling along the Iberian Peninsula was based on a survey made by Sauriau (1991) whose records for the occurrence of *C. neritea* since 1920 were reviewed. However, *C. neritea* was not found in most of the sites cited by Sauriau (1991) showing both the instability of populations and the cryptic nature of *C. neritea* in this area. The appearance and disappearance of *C. neritea* populations is well known in this geographical area (Kisch 1950; Morton 1960) and may be partly due to the closeness of a biogeographic boundary between northern temperate species and southern subtropical species (Glémarec 1979). Such natural barriers are characterized by very peculiar patterns for species distribution, hybridization between species (e.g. Bierne *et al*. 2003) or genetic differentiation at the population level (Luttikhuizen *et al*. 2003; Jolly *et al*. 2005). A more detailed study of *C. neritea* populations in this area would be useful, especially because the warming of the Bay of Biscay may remove selective constraints, leading to reinforcement of introduced populations by natural migrations northwards.

In conclusion, our study showed that the introduced status of *C. neritea* along the FAEC coasts may be associated with several vectors (e.g. oysters or Manila clam) and four out of five populations exhibited a genetic pattern consistent with introductions from multiple sources. Recurrent introductions and admixtures of several evolutionary lineages could have been the starting point for a rapid demographic expansion in sites where this cryptogenic species may have been present at low densities in the past. Bachelet *et al*. (2004) suggested biological features that could make *C. neritea* a particularly competitive gastropod. The temperature increase in the Bay of Biscay may also facilitate its long-term settlement by allowing larger scale migrations northwards, reinforcing the potential for this gastropod to become a real invasive species. Genetic and/or ecological surveys are clearly needed to study those NIS located at the edge of their natural range.

Acknowledgements

The authors are grateful to Jacques Pigeot, Davide Tagliapietra, Francesco Patti and Jeroen Jansen for their help for collecting

samples. We acknowledge Morgan Perennou from the Ouest-Genopole®/Genomer platform at Station Biologique de Roscoff for technical assistance, Marie-Hélène Muller for providing a program she developed for simulating gene genealogies and Carolyn Engel and Claire Daguin for critical reading and improvements of the manuscript. This project is part of the INVABIO program of the Ministère de l'Ecologie et du Développement Durable (project n°D4E/SRP/01115). F.V. acknowledges supports from the European Network of Excellence 'Marine Genomics' (contract n°505403), the CNRS (ATIP program) and the Région Bretagne (Program Accueil & Emergence). B. S.-B. benefited from a PhD grant from the Ministère de l'Education Nationale et de la Recherche.

References

Bachelet G, Simon-Bouhet B, Desclaux C *et al.* (2004) Invasion of the eastern Bay of Biscay by the nassariid gastropod *Cyclope neritea*: origin and effects on the resident fauna. *Marine Ecology Progress Series*, **276**, 147–159.

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Bierne N, Borsa P, Daguin C *et al.* (2003) Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *M. galloprovincialis*. *Molecular Ecology*, **12**, 447–461.

Bohonak AJ (2002) IBD (ISOLATION BY DISTANCE): a program for analyses of isolation by distance. *Journal of Heredity*, **93**, 153–154.

Boulhic M, Tardy J (1986) *Cyclope neritea* (Linné, 1758): Gastéropode Prosobranche nouvellement implanté sur les côtes de Charent-Maritime; éthologie, strucutre de la population. *Haliotis*, **15**, 41–49.

Carlton JT (1996) Biological invasions and cryptogenic species. *Ecology*, **77**, 1653–1655.

Carlton JT (2000) Global change and biological invasions in the oceans. In: *Invasive Species in A Changing World* (ed. Mooney HA, Hobbs RJ), pp. 31–53. Island Press, Washington, D.C.

Carlton JT, Geller JB (1993) Ecological roulette: the global transport of non-indigeneous marine organisms. *Science*, **261**, 78–82.

Cassens I, Van Waerebeek K, Best PB *et al.* (2003) The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of network methods and rooting procedures. *Molecular Ecology*, **12**, 1781–1792.

Cohen AN, Carlton JT (1998) Accelerating invasion rate in a highly invaded estuary. *Science*, **279**, 555–558.

Comes HP, Kadereit JW (1998) The effect of quaternary climatic changes on plant distribution and evolution. *Trends in Plant Sciences*, **3**, 432–438.

Depaulis F, Veuille M (1998) Neutrality tests based on the distribution of haplotypes under an infinite-site model. *Molecular Biology and Evolution*, **15**, 1788–1790.

Diaz-Almela E, Boudry P, Launey S, Bonhomme F, Lapegue S (2004) Reduced female gene flow in the European flat oyster *Ostrea edulis*. *Journal of Heredity*, **95**, 510–516.

Dukes JS, Mooney HA (1999) Does global change increase the success of biological invaders? *Trends in Ecology & Evolution*, **14**, 135–139.

Dupont L, Jollivet D, Viard F (2003) High genetic diversity and ephemeral drift effects in a successful introduced mollusc (*Crepidula fornicata*: Gastropoda). *Marine Ecology Progress Series*, **235**, 183–195.

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

Facon B, Pointier J-P, Glaubrecht M *et al.* (2003) A molecular phylogeography approach to biological invasions of the New World by parthenogenetic thiarid snails. *Molecular Ecology*, **12**, 3027–3039.

Galtier N, Depaulis F, Barton NH (2000) Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. *Genetics*, **155**, 981–987.

Gaylord B, Gaines SD (2000) Temperature or transport? Range limits in marine species mediated solely by flow. *American Naturalist*, **155**, 769–789.

Glémarec M (1979) Les fluctuations temporelles des peuplements benthiques liées aux fluctuations climatiques. *Oceanologica Acta*, **2**, 365–371.

Gomez A, Lunt DH (in press) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Penninsula. In: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity* (eds Weiss S, Ferrand N). Kluwer Academic Publishers, Dordrecht, The Netherlands.

Gomoiu M (1964) Biologishes Studium der Arten *Nassa reticulata* (L.) und *Cyclonassa neritea* (L.) im schwarzen Meer (rumänisher Küstenbereich). *Revue Roumaine de Biologie, Série de Zoologie*, **9**, 39–49.

Gopurenko D, Hughes JM, Keenan CP (1999) Mitochondrial DNA evidence for rapid colonisation of the Indo-West Pacific by the mudcrab *Scylla serrata*. *Marine Biology*, **134**, 227–233.

Goulletquer P, Heral M (1997) Marine molluscan production trends in France: from fisheries to aquaculture. In: *The History, Present Condition and Future of the Molluscan Fisheries of North and Central America and Europe Volume 3, Europe* (eds Mackenzie CL Jr, Burrell VG Jr, Rosenfield A, Hobart WL), pp. 137–164. NOAA technical report NMFS, Seatle, Washington.

Goulletquer P, Bachelet G, Sauriau P-G, Noel P (2002) Open Atlantic coast of Europe — a century of introduced species into French waters. In: *Invasive Aquatic Species of Europe. Distribution, Impacts and Management* (eds Leppäkoski E, Gollasch S, Olenin S), pp. 276–290. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Grosholz E (2002) Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology & Evolution*, **17**, 22–27.

Hamblin MT, Veuille M (1999) Population structure among African and derived populations of *Drosophila simulans*: evidence for ancient subdivision and recent admixture. *Genetics*, **153**, 305–317.

Hinard G, Lambert L (1928) Tableau de l'ostréiculture française (deuxième partie). *Revue des travaux de l'office des pêches maritimes*, **1**, 61–127.

Holland BS (2001) Invasion without a bottleneck: microsatellite variation in natural and invasive populations of the brown mussel *Perna perna* (L). *Marine Biotechnology (New York)*, **3**, 407–415.

Hummel H, Bogaards RH, Bachelet G *et al.* (2000) The respiratory performance and survival of the bivalve *Macoma balthica* (L.) at the southern limit of its distribution area: a translocation experiment. *Journal of Experimental Marine Biology and Ecology*, **251**, 85–102.

IPCC (2001) *Climate change 2001: The scientific basis*. Technical summary of the working group I report (eds Albritton DL, Meira Filho LG). Cambridge University Press, Cambridge, England.

Jolly MT, Jollivet D, Gentil F, Thiebaut 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.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

Kirkendale LA, Meyer CP (2004) Phylogeography of the *Patelloida profunda* group (Gastropoda: Lottidae): diversification in a dispersal-driven marine system. *Molecular Ecology*, **13**, 2749–2762.

Kisch BS (1950) Growth stages of *Cyclope neritea* (L.). *Journal of Conchology*, **23**, 108.

Kolbe JJ, Glor RE, Rodriguez Schettino L *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.

Le Roux A, Maheo R, Dagusan J (1988) Présence et particularités écologiques de *Cyclope neritea* (Linné 1758) (Gastéropode prosobranche) dans le Golfe du Morbihan. *Haliotis*, **18**, 310.

Luttikhuizen PC, Drent J, Baker AJ (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.

Martel C, Viard F, Bourguet D, Garcia-Meunier P (2004) Invasion by the marine gastropod *Ocinebrellus inornatus* in France I. Scenario for the source of introduction. *Journal of Experimental Biology and Ecology*, **305**, 155–170.

Morton J (1960) The habits of *Cyclope neritea*, a style-bearing stenoglossan gastropod. *Proceedings of the Malacological Society of London*, **34**, 96–105.

Muller M-H (2002) *Flux de gènes et domestication: l'histoire évolutive d'un complexe sauvage-cultivé — Approches expérimentales et théorique chez la luzerne pérenne (*Medicago sativa*)*, Biologie, physiologie des organismes et des populations, Thesis, Université des Sciences et Techniques du Languedoc.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Novak SJ, Mack RN, Soltis PS (1993) Genetic variation in *Bromus tectorum* (Poaceae): introduction dynamics in North America. *Canadian Journal of Botany*, **71**, 1441–1448.

Pigeot J (1988) Présence du Gastéropode Nassaridé *Cyclope neritea* (Linné, 1788) dans le basin de Marennes-Oléron. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**, 717–719.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rozas J, Rozas R (1999) DNASP (version 3): an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.

Saitou N, Nei M (1987) The neighbor-joining method: a new method of reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.

Sauriau P-G (1989) *Cyclope neritea* (Linné, 1758) dans le basin de Marennes-Oléron, Bilan des observations de 1984 à 1986. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**, 995–1007.

Sauriau P-G (1991) Spread of *Cyclope neritea* (Mollusca: Gastropoda) along the north-eastern Atlantic coasts in relation to oyster culture and to climatic fluctuations. *Marine Biology*, **109**, 299–309.

Sax DF, Brown JH (2000) The paradox of invasion. *Global Ecology and Biogeography*, **9**, 363–371.

Schneider S, Roessli D, Excoffier L (2001) *ARLEQUIN: A software for population data analysis* Version 2.000. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Switzerland.

Stachowicz JJ, Terwin JR, Whitlatch RB, Osman RW (2002) Linking climate change and biological invasions: ocean warming facilitates nonindigeneous species invasions. *Proceedings of the National Academy of Sciences, USA*, **99**, 15497–15500.

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.

Tardy J, Gaillard J, Portères G (1985) *Cyclope neritea* (Linné, 1788), une espèce de gastéropode nassaridé nouvelle pour les côtes du centre-ouest Atlantique de France. *Annales de la Société des Sciences Naturelles de Charente-Maritime*, **7**, 391–396.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Voisin M, Engel C, Viard F (2005) Differential shuffling of native genetic diversity across introduced region in a brown alga: aquaculture vs. maritime traffic effects. *Proceedings of the National Academy of Sciences, USA*, **102**, 5432–5437.

Wilding CS, Grahame J, Mill PJ (2000) Mitochondrial DNA CoI haplotype variation in sibling species of rough periwinkles. *Heredity*, **85**, 62–74.

Williamson M (1996) *Biological Invasions*. Chapman & Hall, London.

Wolff WJ, Reise K (2002) Oyster imports as a vector for introduction of alien species into northern and western coastal waters. In: *Invasive Aquatic Species of Europe. Distribution, Impact and Management* (eds Leppäkoski E, Gollasch S, Olenin S). Kluwer Academic Publishers, Dordrecht/Boston/London.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

This work is part of B. Simon-Bouhet's doctoral research. His thesis aims at studying the basis of recent changes in the geographical distribution of the gastropod *C. neritea*. P. Garcia-Meunier uses molecular genetics to study the effects of environmental processes on genetic variability in coastal species. F. Viard study historical and contemporary dispersal of marine algae and invertebrates, including introduced species, through population genetics approaches.

Appendix

Haplotype definition (H1 to H29: GenBank accession numbers: AY789970–AY789991 and DQ087210–DQ087216) and number of occurrence (*N*). Only polymorphic sites are indicated and positions are given in number of base pairs.

Haplotypes	*N*	11122222222233333333334444445 22445506912356689912234778890224770 15914034428404462297038747298125069 3
H1	263	AACACAATCGAAGCGCGGGCTATATACGGTTGAGAT
H2	3	----------A----------
H3	3	----------A----------
H4	9	----------T----------
H5	5	-------C----------
H6	1	-----G----------G----C-C----C
H7	2	-----G----------G-G--C-C----C
H8	1	----TG----------G-G--C-C----C
H9	1	----TG----------A-----G-G--C-C----C
H10	4	G---TG----------A-----G-G--C-C----C
H11	1	G---TG----------A-----G----C-C----C
H12	1	G---TG----------G----C-C----C
H13	1	G---TG----G----------G-G--C-C----C
H14	1	-----GG----------G-G--C-C-----
H15	1	----TGG------T-T-------G-G--C-C-----
H16	13	----TGG------T-T------CG-G--C-C-----
H17	35	----TGG------T-T--A---CG-G--C-C-----
H18	119	----TGG------T-T---T---G-G--C-C---GC
H19	6	----TGG------T-T---T---G-GA-C-C---GC
H20	1	G---TG-C----------A-----G-G--C-C----C
H21	1	G---TG-C--GG-----A-----G-G--C-C----C
H22	97	G---TG-C--GG-----A-----G-G--C-CA---C
H23	1	G---TG-C--GG-----A-----G----C-CA---C
H24	7	G--GTG-C--GG-----A-----G-G--C-CA---C
H25	1	G---TG-C--GG----AA-----G-G--C-CA---C
H26	3	G---TG-C--GG----------G-G--C-CA---C
H27	4	G---TG-C-AGGA----------G-G--C-CA---C
H28	1	----TG----------G-G-AC-C-GA--
H29	8	-G--TG--T------T---TCG-GC---CCC----C