Population genetic analysis of *Ensis directus* unveils high genetic variation in the introduced range and reveals a new species from the NW Atlantic

Joaquín Vierna[1,2], K. Thomas Jensen[3], Ana M. González-Tizón[1], Andrés Martínez-Lage[1]

[1] Department of Molecular and Cell Biology, Evolutionary Biology Group (GIBE), Universidade da Coruña, A Fraga 10, 15008 A Coruña, Spain

[2] AllGenetics, Edificio de Servicios Centrales de Investigación, Campus de Elviña s/n, 15008 A Coruña, Spain

[3] Marine Ecology, Department of Bioscience, Aarhus Universitet, Ole Worms Allé 1, Building 1135, 8000 Aarhus C, Denmark

Marine Biology Volume 159, Issue 10, pages 2209–2227, October 2012
Received 9 March 2012, Accepted 28 June 2012, Published online 31 July 2012

The final publication is available at Springer via https://doi.org/10.1007/s00227-012-2006-6.

How to cite:

Vierna J, Jensen KT, González-Tizón A, Martínez-Lage A (2012) Population genetic analysis of *Ensis directus* unveils high genetic variation in the introduced range and reveals a new species from the NW Atlantic. Mar Biol 159: 2209-2227.

Abstract

We report current genetic variation of populations of the razor shell *Ensis directus* (Conrad 1843) (Mollusca: Bivalvia: Pharidae) in native (North American) and introduced (European) ranges using nuclear and mitochondrial sequence-based markers. We expected less variation within the introduced range, especially considering the frequent mass mortality events observed in Europe since the species was recorded for the first time in 1978. However, we found higher variation in Europe. The possible significance of temporal fluctuations of genetic variation, limited effect of random genetic drift, and multiple introductions are discussed. Interestingly, the multiple-introduction hypothesis contrasts with the gradual colonisation of European coastal waters but is supported by trained clustering analysis and by the intensity of transatlantic shipping. Genetic and morphometric evidence strongly supports that examined individuals from a supposed *E. directus* population from Newfoundland (Canada) belong to a separate species. This new *Ensis* is formally described here and named *E. terranovensis* n.sp.

Keywords: COI, ITS, adenine nucleotide translocase, razor clam, *Ensis terranovensis*

Abbreviations

MNCN	Museo Nacional de Ciencias Naturales, Madrid, Spain
MNHN	Muséum national d'Histoire naturelle, Paris, France
ZMUC	Zoologisk Museum—Københavns Universitet, Copenhagen, Denmark
CMNML	Canadian Museum of Nature, Ottawa-Gatineau, Canada

Introduction

The razor shell *Ensis directus* (Conrad 1843) syn. *E. americanus* (Binney, in Gould and Binney 1870) (Mollusca: Bivalvia: Pharidae) is native to Atlantic North America, and it was introduced into European coastal waters in the late 1970s probably as larvae in ballast water from a ship crossing the Atlantic (von Cosel et al. 1982). The first European specimens were observed in the German Bight near the mouth of the river Elbe estuary in 1979 (von Cosel et al. 1982; Cosel 2009). Since, it has colonised the continental coastline of North Europa from Normandy to southern Norway. From 1989 onwards, it has also spread along the North Sea coast of England until the Humber Estuary (Cosel 2009), and there have been several observations along the Channel coast. It was detected in 2002 in South Wales, near Milford Haven (Paul Dansey personal communication), and very recently, in Liverpool Bay (Dansey 2011), and in the Cantabrian Sea (Arias and Anadón 2012). Furthermore, it has also been observed in high numbers in the North Sea until a depth of about 30 m (Mühlenhardt-Siegel et al. 1983; Cosel 2009).

Ensis directus has become an integral part of the recipient ecosystem as a consumer of phytoplankton and as a prey to various fish and water birds (Tulp et al. 2010). However, no studies have been published so far that support a suppression of native species by *E. directus* in European waters. Quite the opposite, both Armonies and Reise (1999) and Dannheim and Rumohr (2011) concluded that the newcomer might have favoured the appearance of some other species in the areas in which it occurs. Nowadays, *E. directus* is a commercial species in Europe (see Marine Stewardship Council 2012). Similarly, in the north eastern US and eastern Canada, there is an increasing interest in *E. directus* fisheries and valuable work was carried out to investigate the aquaculture potential of the species in Maine (Maine Sea Grant 2012) and in Nova Scotia (Kenchington et al. 1998).

The dispersal of *E. directus* is facilitated by a pelagic larval life (2–4 weeks) during which it may reach a distance of 125 km downstream from its source population (Armonies 2001) (estimate based on its dispersal from 1979). Human-facilitated spread within its new ecosystem could be a further dispersal mechanism that may have contributed to its present European distribution. According to Armonies (2001), the temporal course of dispersal along the coastline fits the hypothesis of a single introduction into the North Sea but the possibility of multiple imports of the species cannot be excluded, and this may be tested genetically.

One important factor impacting the genetic variation of the European populations could be the often observed mass mortalities in local *E. directus* populations that seem to be a characteristic feature of the species. For instance, thousands of dead or dying razor shells have been observed several times along the Danish shoreline (North Sea, Limfjorden, and Kattegat areas) (Freudendahl et al. 2010). Such events may reduce levels of genetic variation either because of selective mortality or through random eradication of genotypes.

In population genetics, the analysis of several loci has become a must, as studies based on nucleotide variation at a single locus provide insufficient information of genetic patterns. In this sense, the selection of loci with different evolutionary histories (e.g. cytoplasmic vs nuclear genes) is important to distinguish among factors affecting genetic variation (e.g. low variation due to a selective sweep or a population bottleneck). But, the selection of markers is restricted to availability of suitable primers in non-model organisms.

In this work, we have studied nucleotide variation at four sequence-based molecular markers, including mitochondrial, nuclear multi-copy, and nuclear single-copy regions in order to: (1) compare current genetic variation in native and introduced ranges to assess to what extent potential bottlenecks and mass mortality events in Europe have impacted diversity, (2) obtain information about the possible origin of European individuals, and (3) study population structure in the native range of the species. Besides, we serendipitously detected a very divergent population from Conception Bay, Newfoundland (Canada) that is proposed to be a

new *Ensis* species based on genetic and morphometric evidence. This new *Ensis* species from the NW Atlantic is formally described here.

Materials and methods

Specimens and lab procedures

We studied a set of 148 razor shells from native and introduced sites (see Table 1; Fig. 1 for details). Razor shells were preserved in 100 % ethanol, and they were identified in the lab as *E. directus* according to shell morphology (Cosel 2009). Identifications of some of the specimens (Table 1) were confirmed by Rudo von Cosel (Muséum national d'Histoire naturelle, Paris, France) and included in his article on the taxonomy of Atlantic *Ensis* (Cosel 2009). One individual from the related species *E. minor* Dall, 1899 (collected off Christmas Bay, Texas, USA) was included as outgroup in some of the analyses.

Table 1. Sites sampled in this work.

Sampling site	Coordinates	Collected by	Years	Depth	Identified by	Museum code
Sillerslev	56°42′20″N, 8°47′20″E	K. T. Jensen	2005	1–2 m	K. T. Jensen	
Sundsøre	56°42′25″N, 9°10′31″E	K. T. Jensen	2005	1–2 m	K. T. Jensen	
Juvre Deep	55°11′45″N, 8°25′56″E	K. T. Jensen	2005	Intertidal (<2 m)	K. T. Jensen	
Vester Vedsted	55°16′29″N, 8°37′45″E	K. T. Jensen	2008	Intertidal (<2 m)	J. Vierna	
The Wash	52°56′21″N, 0°24′53″E	D. Palmer	2007	1–2 m	R. von Cosel, J. Vierna	
Katwijk	52°11′57″N, 4°24′41″E	J. Goud	2008	Thrown on beach after storm (ca. 1–3 m)	J. Goud, J. Vierna	RMNH.MOL.102103
Dunkerque	51°02′03″N, 2°16′37″E	J. M. Dewarumez	2009	Intertidal (<2 m)	J. Vierna	
Cobscook Bay	44°54′35″N, 67°4′13″W	T. Sheehan (Gulf of Maine)	2008	1.8 m at high mean water (intertidal)	J. Vierna	MNCN 15.07/11733
Shinnecock Bay	40°52′07″N, 72°28′02″W	S. T. Tettelbach	2008	0.5–3 m	J. Vierna	MNCN 15.07/11734
Long Pond	47°30′54″N, 52°58′30″W	P. Sargent, R. O'Donnell	2007	6–11 m	R. von Cosel, J. Vierna	Several codes (see 'Taxonomy')

RMNH, NCB Naturalis, Leiden (The Netherlands). *MNCN* Museo Natural de Ciencias Naturales, Madrid (Spain).

DNA was extracted from muscle tissue using the NucleoSpin Tissue kit (Macherey–Nagel GmbH and Co. KG). All razor shells were sequenced for a fragment of the mitochondrial cytochrome oxidase subunit I gene (COI), and for the nuclear ribosomal multi-copy region encoding both internal transcribed spacers (ITS1 and ITS2) and the 5.8S ribosomal gene (5.8S). A subset of 70 animals (including the outgroup) were additionally

sequenced for a fragment of a nuclear single-copy region, the adenine nucleotide translocase gene (ANT). Using the 'universal' primers designed by Folmer et al. (1994) and Audzijonyte and Vrijenhoek (2010), we obtained sequences from three individuals that were then used to design three pairs of internal primers in GeneFisher (Giegerich et al. 1996) (COI-*directus*-F, 5′ CAG GTT TAG TTG GAA CTA GG; COI-*directus*-R, 5′GAT CTC CRC CAC CTC T; ANT-*Ensis*-a-F, 5′ AAA CAT GGC CAA CTG CAT CCG AT; ANT-*Ensis*-a-R, 5′ CAA GGA CAT AAA GCC CTC TGC CTT; ANT-*Ensis*-b-F, 5′ TTC CCA ACC CAG GCC TTG; ANT-*Ensis*-b-R, 5′ ATG ATG GTY GTG GCA CAG T). These internal primers were used in all subsequent amplifications. The primers used to amplify the ITS1-5.8S-ITS2 region were those from Heath et al. (1995). Each PCR reaction (25 μL) contained ~25 ng of genomic DNA, 0.625 U of *Taq* DNA polymerase (Roche Diagnostics), 5 nmol of each dNTP (Roche Diagnostics), 20 pmol of each primer, and the buffer recommended by the polymerase supplier. The general reaction conditions were: an initial denaturation step at 94 °C for 3 min followed by 35 cycles of denaturation at 94 °C for 20 s; annealing at the following temperatures (COI-*directus*, 48 °C; ANT-*Ensis*-a, 58 °C; ANT-*Ensis*-b, 54 °C; and ITS1-5.8S-ITS2 region, 59 °C) for 20 s; extension at 72 °C for 30–50 s; and a final extension at 72 °C for 5 min. PCR products were run on agarose gels, stained with ethidium bromide, and imaged under UV light. COI and ANT PCRs yielded one single gel band from the expected sizes, and the amplification products were purified using ExoSAP-IT (USB). COI amplification products were sequenced in both directions using PCR primers, whereas ANT amplicons were sequenced using the ANT-*Ensis*-b-F primer only. ITS1-5.8S-ITS2 PCRs yielded single-band patterns. Nevertheless, since intragenomic variants occur in *E. directus* (Vierna et al. 2010), a cloning step was necessary. Therefore, amplification products were cloned using the TOPO TA Cloning kit (Invitrogen). We selected transformant colonies, checked their insert size by PCR, spread one clone per individual on an LB plate, and let it grow overnight at 37 °C. Plasmids were purified with the QiaPrep Spin Miniprep Kit (Qiagen), and they were sequenced using the M13 Forward and Reverse primers supplied in the cloning kit.

Bioinformatic analyses

The software BioEdit 7.0.9.0 (Hall 1999) and Geneious Pro 5.4.6 (Drummond et al. 2011) were used to examine the electropherograms. Since we expected heterozygote positions in the ANT sequences, we used Geneious Pro 5.4.6 to detect them (peak similarity = 50 %). Afterwards, they were confirmed manually. The 5.8S region was not further considered here because it was almost invariable. Alignments were carried out in ClustalW 2.0 (Larkin et al. 2007) from the MEGA 5.03 package (Tamura et al. 2011). In the case of COI and ANT, they were performed considering the amino acid sequence. ITS1 and ITS2 alignments were manually corrected using the RALEE (RNA ALignment Editor in Emacs) tool (Griffiths-Jones 2005).

The gametic phase of each ANT sequence was obtained in DnaSP 5.10.01 (Librado and Rozas 2009) implementing the algorithm provided in PHASE 2.1 (Stephens et al. 2001; Stephens and Scheet 2005). PHASE simulations were run five times (with a different seed each time) using the MR0 model and assuming recombination. Output probability thresholds were 0.9, and all other parameters were set as default. No differences were appreciated among the five resulting alignments (140 sequences, each), which were inspected for discrepancies by comparing their consensus sequences (constructed in Geneious Pro 5.4.6), and their nucleotide diversity values (obtained from DnaSP 5.10.01).

Diversity analyses were done in DnaSP 5.10.01, after creating several subsets according to sampling sites or geographic areas. In the cases of ITS1 and ITS2 sequences, gap-containing positions were excluded in each subset. However, alignment gaps were considered in the haplotype/sequence-type data files that we created in DnaSP 5.10.01 (which are compatible with the Arlequin 3.5.1.2 software, Excoffier and Lischer 2010). Therefore, for the over-mentioned datasets, two values for the number of sequence-types were provided, one excluding gaps that were present in the subset, and one including all gaps.

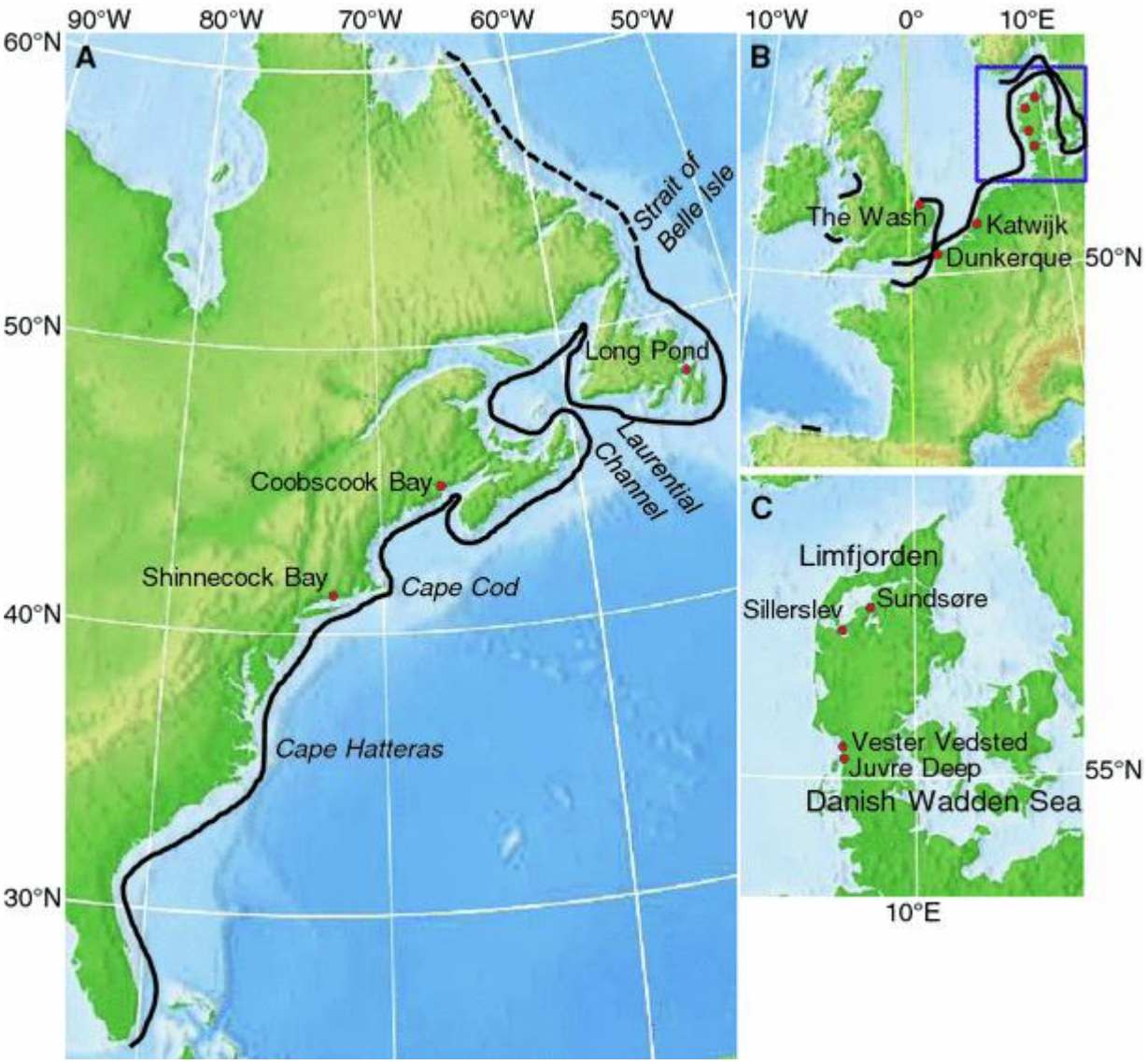

Fig. 1. Maps showing the approximate current distribution of *E. directus* according to Kenchington et al. (1998), Cosel (2009), Dansey (2011), and Arias and Anadón (2012). *Red dots* correspond to sampling sites. In the *dashed lined* area, the occurrence of the species is unclear. Potential barriers to gene flow in the Western Atlantic Boreal Region are indicated. a Native range. b Introduced range. c Introduced range (detail).

In Arlequin 3.5.1.2, we looked for shared haplotypes/sequence-types among sampling sites. Using this software, we also performed analyses of molecular variance (AMOVAs, Excoffier et al. 1992), in which sampling sites were grouped in different ways, and running 1,000 permutations. The Tamura—Nei distance was employed to build up the matrix. Danish sites were clustered in pairs considering their geographic proximity: Sillerslev and Sundsøre (Limfjorden); Juvre Deep and Vester Vedsted (Danish Wadden Sea). The fixation index F_{ST} was also calculated for each pair of population comparisons, and its significance was tested by 10,000 permutations. Finally, exact tests of population differentiation were carried out. They test the hypothesis of random distribution of individuals between pairs of populations as described in Raymond and Rousset (1995) and Goudet et al. (1996). The number of steps in Markov chain was set to 1,000,000, and the number of dememorisation steps, to 100,000.

In the F_{ST} and test of population differentiation analyses, we considered only four groups of sequences (Europe, Shinnecock Bay, Cobscook Bay, and Long Pond) following Fitzpatrick et al. (2012). Since these methods require populations to have reached a mutation-drift equilibrium, the authors pointed out that such

population genetic methods must not be applied to study population structure within a species' introduced range. They claimed that the time since the initial introduction and subsequent expansion of range of an introduced species is too recent for the effects of mutation and drift to be at equilibrium, particularly when effective population sizes are large (as in marine bivalves). Only regarding these comparisons and the neutrality tests (see below), we assume that European sequences represent the genetic variation of their source population. This would only be the case when no genetic bottlenecks occur, the source population is only one, and the introduction event is only one as well.

In order to investigate whether there was a native population more likely to be the source of European individuals, we input our COI-ITS1-ITS2 dataset in BAPS 5.3 (Corander et al. 2006, 2008) and ran a trained clustering analysis (Cheng et al. 2011) with a maximum number of clusters ranging between $K = 3$ and $K = 9$. The ANT sequences were not used because they were not available for all individuals. The analysis was repeated five times with invariable results. Individuals from Long Pond were excluded from this analysis because they did not share haplotypes/sequence-types with European individuals (see below).

Several neutrality tests were calculated with COI and ANT datasets. The ITS1 and ITS2 sequences were excluded since the occurrence of intragenomic divergence within this region in *E. directus* (Vierna et al. 2010) could violate some test assumptions and produce biased results. Tajima's *D* (Tajima 1989) and Fu's *Fs* (Fu 1997) were calculated in Arlequin 3.5.1.2, applying a number of simulated samples of 1,000, and Fu and Li's *D* and *F* (Fu and Li 1993), and Fay and Wu's *H* (Fay and Wu 2000), in DnaSP 5.10.01, using the total number of mutations, and *E. minor* as outgroup. The McDonald-Kreitman test (McDonald and Kreitman 1991) was also performed in DnaSP 5.10.01. In this case, the Canadian sequences were used as outgroup for each European partition, and the European sequences (all together) were used as outgroup for the Canadian sequences. The Hudson–Kreitman–Aguadé test (Hudson et al. 1987) was implemented for the COI + ANT datasets, taking *E. minor* sequences for interspecific comparisons, and using the 'direct mode', in DnaSP 5.10.01. This mode allows to compare loci that differ in their effective population sizes.

Mismatch distributions of the pairwise number of differences (Rogers and Harpending 1992) were obtained, and the goodness of fit of the observed and expected curves (under the 'sudden expansion' model) was assessed by the sum of squared deviations (SSD) and the raggedness statistic (Harpending 1994). The Θ_0, Θ_1, and τ parameters were calculated in Arlequin 3.5.1.2 with 100 bootstrap replicates, and the output values were introduced in DnaSP 5.10.01, where the histograms for each partition were obtained under a 'population growth-decline' model.

Mismatch analyses were also used to roughly estimate the time elapsed since expansion of the partitions that did not show significant SSD/raggedness statistic tests (i.e. those that may have undergone an expansion event). We used the equation $\tau = 2ut$ (Rogers and Harpending 1992), where $u = 2\ \mu l$, being μ the number of mutations per nucleotide site per generation, and l, the sequence length. Generation time was assumed to be 5 years, though *E. directus* reaches sexual maturity after 1 year (see Mühlenhardt-Siegel et al. 1983). Mutation rates for COI (0.14 and 0.52 % divergence per nucleotide site per million years) were obtained from Luttikhuizen et al. (2003) who studied another heterodont species. We are not aware of any reported estimate of mutation rates for the ANT region.

COI and ANT networks were calculated under the maximum parsimony criterion in TCS 1.21 (Clement et al. 2000) using the haplotype/sequence-type dataset, and applying a connexion limit high enough to permit the outgroup connexion. Due to their higher complexity, ITS1, ITS2, and ITS1 + ITS2 networks were built up from the sequence-type datasets using the neighbournet algorithm (Bryant and Moulton 2004) and uncorrected p-distances in SplitsTree4 (Huson and Bryant 2006).

Phylogenies inferred under maximum likelihood (ML), bayesian (BA), and maximum parsimony (MP) criteria were obtained for the COI + ANT + ITS1 + ITS2 concatenated dataset (which comprised sequences

from 69 individuals and 1,710 nucleotides). RAxML-7.2.8 (Stamatakis 2006, 2008) was run from the CIPRES Science Gateway (Miller et al. 2010). This software is capable to assign and estimate separate model parameters for individual genes of multi-gene alignments (Stamatakis 2006) and implements the general time reversible (GTR) substitution model for all partitions. Therefore, we partitioned our data by genes. A GTRCATI model of nucleotide substitution was implemented, and 1000 automatic bootstraps were performed, followed by a search for the best-scoring ML tree. Gaps were considered as missing data. An additional ML analysis was performed in PhyML (Guindon et al. 2010) that was run through the ATGC Bioinformatics Platform (http://www.atgc-montpellier.fr/), implementing a GTR model of nucleotide substitution, 1,000 non-parametric bootstraps, and considering gaps as missing data. All other parameters were set to their default values, and in this case, data were not partitioned by genes. BA was carried out using the software MrBayes v. 3.0B4 (Huelsenbeck and Ronquist 2001) through the CIPRES Science Gateway. Models of evolution for each partition were obtained from MrModelTest v2.3 (Johan Nylander, http://www.abc.se/~nylander/). This software selected the HKY + I + G model for the COI partition, the K80 model for ANT, and the GTR + I + G model for either ITS. Gaps were treated as missing data. The analysis was performed with 15,000,000 generations initiated with a random starting tree, sampling every 1,000 generations and allowing the program to estimate the likelihood parameters required. Stationarity was assessed using the web-based software AWTY (Nylander et al. 2008). Results collected prior to stationarity were discarded as burn-in. A MP bootstrap consensus tree was retrieved from Paup4.0b10 (Swofford 2002) using the heuristic search method. Parameters were set as follows: gaps were treated as a 'fifth state' (this applies to the ITS1 and ITS2 sequences), multistate taxa were interpreted as uncertainty (this applies to the ANT sequences), starting trees were obtained via stepwise addition, the number of trees held at each step during stepwise addition was one, and the branch-swapping algorithm selected was TBR. The robustness of the obtained topology was assessed after running 1,000 non-parametric bootstraps. All phylogenetic trees were edited in FigTree 1.2.2. (Andrew Rambaut, http://tree.bio.ed.ac.uk/software/figtree/).

Finally, to compare our results to other reported cases, between-groups mean K2P distances were calculated using the MEGA 5.03 package. In the case of ITS1 and ITS2, gaps were considered in pairwise comparisons. Standard errors were obtained after running 1,000 permutations.

Morphometric analyses

Razor shells from the three native sites were additionally studied in terms of shell morphometrics. Two individuals from Cobscook Bay and two other from Long Pond that were studied genetically could not be studied morphologically because their valves were missing.

Using a vernier calliper, we measured the length of each right valve (considering the longest axis), and the width at the posterior adductor scar. The length of the posterior adductor scar (named here 'distance a'), and the distance between its posterior end and the beginning of the pallial sinus (along the dorsal line) ('distance b') were also recorded, since they seemed to differ among sites. Furthermore, the shape of muscle scars is a main taxonomic character in *Ensis* (see Cosel 2009). Distances measured are indicated in Fig. 2. We also weighed each pair of valves (whenever possible, since some of the shells were incomplete) and obtained the mean value for each valve.

To analyse morphometric data, we used the statistical programme SPSS vs 19. Prior to running ANCOVA, regression residuals were tested for homoscedasticity.

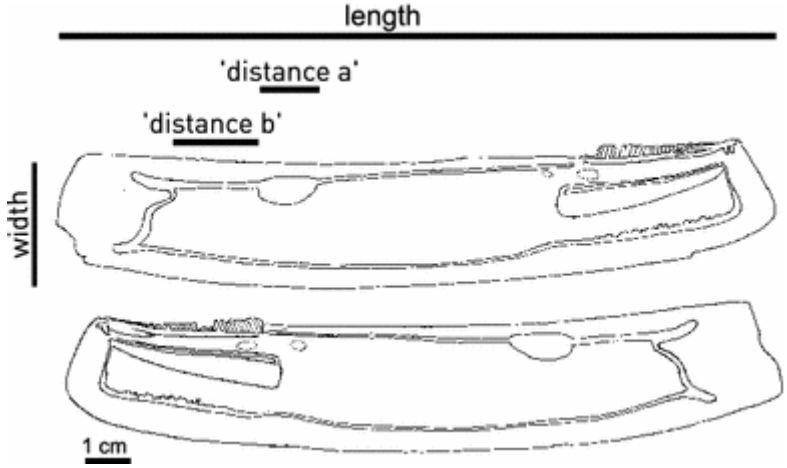

Fig. 2. Internal view of the valves of the holotype of *Ensis terranovensis* n.sp., showing the muscle scars and the distances considered in the morphometric analyses. '*Distance a*', length of the posterior adductor scar; '*distance b*', distance along the dorsal line between the posterior end of the posterior adductor scar and the beginning of the pallial sinus.

Results

Genetic variation

COI, ITS1, and ITS2 sequences were obtained for 77 European and 71 North American individuals. In addition, we obtained ANT sequences for a subsample of 45 European and 24 North American individuals. These sequences were converted into 90 European and 48 North American phased sequences.

After having deleted the primer-annealing regions and some low-quality terminal regions of the electropherograms, we obtained sets of sequences with the following lengths: COI, 454 bp; ITS1, 484–514 bp; ITS2, 295–301 bp; and ANT, 406 bp. All alignments were straightforward, despite some minor corrections performed in the ITS1 and ITS2 datasets. All sequences can be accessed at the DDBJ/EMBL/GenBank databases (accession numbers HE661632–HE662148).

Polymorphism values for each marker are indicated in Table 2 by site and geographic area. COI and ITS1 were more polymorphic than ITS2 and ANT. Considering all sequences (except the outgroup), COI displayed the highest proportion of segregating sites ($s = 0.165$), and the highest nucleotide diversity ($\Pi = 0.025 \pm 0.002$) and average number of nucleotide differences ($k = 11.303$). The values for the ITS1 region were higher in terms of the number of sequence-types obtained ($h = 73$; $h = 111$, depending on gaps, as explained above) and in terms of haplotype diversity ($H = 0.963 \pm 0.008$).

Strikingly, European sequences were, in general, more polymorphic than those from native sites (Table 2). Only the H value for Cobscook Bay ITS2 and for Shinnecock Bay ITS1 and the Π value for Long Pond COI sequences were higher in native populations. In the particular case of ITS1 and ITS2, European, Shinnecock Bay, and Cobscook Bay sequences were more polymorphic than those from Long Pond, but this was probably due to the occurrence of paralog ITS sequences in the genomes of these razor shells, which were not present in Long Pond individuals (see 'Evolutionary relationships among variants').

Table 2. Polymorphism in sampling sites or geographic areas.

	Limfjorden		Danish Wadden Sea		The Wash	Katwijk	Dunkerque	European seq.	Shinnecock Bay	Cobscook Bay	Long Pond	*N* (total)
	Sillerslev	Sundsøre	Juvre Deep	Vester Vedsted								
COI												
N	8	8	8	8	15	15	15	77	21	22	28	148
l	454	454	454	454	454	454	454	454	454	454	454	454
S	20	16	15	13	18	15	18	43	21	12	20	75
s	0.044	0.035	0.033	0.029	0.040	0.033	0.040	0.095	0.046	0.026	0.044	0.165
h	7	7	7	6	11	9	10	36	12	7	15	62
H	0.964 ± 0.077	0.964 ± 0.077	0.964 ± 0.077	0.893 ± 0.111	0.952 ± 0.04	0.914 ± 0.052	0.914 ± 0.056	0.948 ± 0.013	0.905 ± 0.048	0.792 ± 0.066	0.884 ± 0.045	0.949 ± 0.011
Π	0.012 ± 0.001	0.012 ± 0.002	0.011 ± 0.002	0.008 ± 0.002	0.010 ± 0.001	0.009 ± 0.001	0.010 ± 0.001	0.010 ± 0.000	0.008 ± 0.001	0.006 ± 0.001	0.013 ± 0.009	0.025 ± 0.002
k	5.643	5.536	5.036	3.536	4.343	4.19	4.438	4.678	3.571	2.736	5.738	11.303
ITS1												
N	8	8	8	8	15	15	15	77	21	22	28	148
l	495	498	492	494	494	495	495	487	495	493	484	530
S	18	26	20	22	21	14	18	45	20	21	7	62
s	0.036	0.052	0.041	0.045	0.043	0.028	0.036	0.092	0.040	0.043	0.014	0.117
h	8 (8)	8 (8)	8 (8)	8 (8)	12 (14)	13 (14)	14 (15)	48 (64)	15 (20)	19 (19)	7 (9)	73 (111)
H	1.000 ± 0.063	1.000 ± 0.063	1.000 ± 0.063	1.000 ± 0.063	0.943 ± 0.054	0.981 ± 0.031	0.990 ± 0.028	0.964 ± 0.012	0.952 ± 0.032	0.983 ± 0.021	0.791 ± 0.041	0.963 ± 0.008
Π	0.014 ± 0.002	0.017 ± 0.005	0.015 ± 0.002	0.016 ± 0.002	0.010 ± 0.002	0.011 ± 0.001	0.011 ± 0.001	0.011 ± 0.001	0.009 ± 0.001	0.011 ± 0.001	0.004 ± 0.001	0.014 ± 0.001
k	7.179	8.500	7.357	7.857	5.143	5.314	5.676	5.292	4.533	5.632	2.013	6.794
ITS2												
N	8	8	8	8	15	15	15	77	21	22	28	148
l	295	295	295	295	295	293	294	292	294	295	295	320
S	9	8	10	9	10	10	12	18	12	12	4	25
s	0.031	0.027	0.034	0.031	0.034	0.034	0.041	0.062	0.041	0.041	0.014	0.078
h	6 (6)	7 (7)	7 (7)	6 (6)	12 (12)	8 (10)	10 (11)	22 (29)	12 (16)	13 (14)	5 (5)	31 (50)
H	0.929 ± 0.084	0.964 ± 0.077	0.970 ± 0.077	0.929 ± 0.084	0.962 ± 0.040	0.895 ± 0.003	0.933 ± 0.002	0.929 ± 0.012	0.933 ± 0.031	0.909 ± 0.043	0.725 ± 0.042	0.918 ± 0.009
Π	0.014 ± 0.002	0.010 ± 0.003	0.016 ± 0.002	0.016 ± 0.002	0.014 ± 0.001	0.011 ± 0.002	0.015 ± 0.002	0.014 ± 0.001	0.013 ± 0.001	0.013 ± 0.001	0.004 ± 0.000	0.014 ± 0.001
K	4.143	2.893	4.857	4.607	4.21	3.238	4.343	4.018	3.943	3.991	1.336	4.081
ANT												
N		4	8	10	24	24	20	90	6	12	30	138
l		406	406	406	406	406	406	406	406	406	406	406
S		3	2	3	9	7	7	16	2	5	5	22
s		0.007	0.005	0.007	0.022	0.017	0.017	0.039	0.005	0.012	0.012	0.054
h		3	3	3	11	9	8	17	4	5	6	24
H		0.833 ± 0.222	0.607 ± 0.164	0.622 ± 0.138	0.848 ± 0.063	0.855 ± 0.051	0.805 ± 0.07	0.821 ± 0.029	0.8 ± 0.172	0.742 ± 0.116	0.634 ± 0.08	0.860 ± 0.020
Π		0.004 ± 0.001	0.002 ± 0.001	0.003 ± 0.001	0.004 ± 0.001	0.004 ± 0.000	0.004 ± 0.001	0.004 ± 0.000	0.003 ± 0.001	0.003 ± 0.001	0.003 ± 0.000	0.004 ± 0.001
k		1.500	0.679	1.067	1.558	1.518	1.558	1.432	1.067	1.348	1.067	1.630

European sequences are those from Sillerslev, Sundsøre, Juvre Deep, Vester Vedsted, The Wash, Katwijk, and Dunkerque. *N* number of sequences considered in each analysis, *l* length of the genomic region studied, *S* number of segregating sites, *s* number of segregating sites per nucleotide site, *h* number of haplotypes or sequence-types, *H* haplotype diversity, Π nucleotide diversity, *k* average number of nucleotide differences, *h* values in brackets are those in which gaps were considered (see 'Materials and methods'). *H* and *Π* values are expressed with their SD.

Distribution of variants

A graphic representation of the distribution of all variants is available (Online Resource 1). We sampled 62 COI haplotypes (out of 148 sequences), and 36 of them were found in the introduced range. The number of private European haplotypes was 30, and the remaining six were shared with Shinnecock Bay (four haplotypes), Cobscook Bay (four haplotypes) or both sites (two haplotypes). Shinnecock Bay displayed 12 COI haplotypes, eight of which were private; and Cobscook Bay, seven haplotypes (three of them were private). The number of shared variants between Shinnecock Bay and Cobscook Bay was only two. In Long Pond, we sampled 15 haplotypes, all of them, private.

In the case of ITS1, 111 sequence-types were sampled, 63 of them were private to Europe, and one was shared between Europe and Cobscook Bay. Shinnecock Bay displayed 20 (all private) variants; and Cobscook Bay, 18 privates and the one shared with Europe. There were no shared variants between Shinnecock Bay and Cobscook Bay. Long Pond displayed nine variants, all private.

The number of ITS2 sequence-types obtained was 50, 29 of them occurring in Europe. Among those, 20 variants were private to this region, eight were shared with Shinnecock Bay, five with Cobscook Bay, and four with both sites. In Shinnecock Bay, we sampled 16 ITS2 sequence-types (seven were private); and in Cobscook Bay, 14 sequence-types (eight, private). The number of shared variants between Shinnecock Bay and Cobscook Bay was five. Again, all Long Pond sequence-types (five) were private to this sampling site.

All sampling sites displayed many private sequence-types when the ITS1 + ITS2 dataset was considered (see Online Resource 1).

Finally, the ANT region was more conserved, and the number of sequence-types sampled was 24 (in this case, out of 138 sequences); 13 of them were private to European waters, and four sequence-types were shared among European and North American sampling sites. In Shinnecock Bay, we sampled four sequence-types (one private, three shared with Europe and with Cobscook Bay, one shared with Long Pond). In Cobscook Bay, we found five (one private, three shared with Shinnecock Bay, four with Europe, one with Long Pond). In Long Pond, we found six (five private, one shared with Europe, Shinnecock Bay, and Cobscook Bay).

Population differentiation

AMOVA results showed that, when only introduced sampling sites were considered (Limfjorden–Danish Wadden Sea–The Wash–Katwijk–Dunkerque), the percentage of genetic variation within populations was very high (97.51 % for COI, 101.50 % for ITS1, 97.18 % for ITS2, and 99.21 % for ANT).

Similarly, in the comparison European sites (altogether)—Cobscook Bay—Shinnecock Bay, the percentage of genetic variation within populations was 95.31 % for COI, 100.51 % for ITS1, 101.71 % for ITS2, and 99.91 % for ANT, again indicating a lack of structure.

There was one grouping that maximised the percentage of variation among groups. It was the one in which Long Pond was separated from all other sampling sites (European sequences + Shinnecock Bay + Cobscook Bay − Long Pond). Specifically, the values obtained were: COI, 83.64 %; ITS1, 63.91 %; ITS2, 60.03 %; and ANT, 39.57 %.

F_{ST} values were significant in the COI comparisons between European sequences and either Shinnecock Bay ($F_{ST} = 0.03$) or Cobscook Bay ($F_{ST} = 0.03$). These values suggest low differentiation according to mitochondrial DNA. Tests of population differentiation were significant in the comparisons between European sequences and either Shinnecock Bay or Cobscook Bay according to ITS1.

In the comparisons between Shinnecock Bay and Cobscook Bay, F_{ST} values were non-significant, but the COI test of population differentiation suggested some degree of differentiation between these native sites.

Finally, both types of tests and all four molecular markers yielded significant results in all comparisons between Long Pond and all other sites or areas (Table 3).

Table 3. Population differentiation.

	European seq.	Shinnecock Bay	Cobscook Bay	Long Pond
European seq.				
COI	–	ITS1	ITS1	COI, ITS1, ITS2, ANT
ITS1	–			
ITS2	–			
ANT	–			
Shinnecock Bay				
COI	***0.03***	–	COI	COI, ITS1, ITS2, ANT
ITS1	0.01	–		
ITS2	−0.02	–		
ANT	0.02	–		
Cobscook Bay				
COI	***0.03***	0.01	–	COI, ITS1, ITS2, ANT
ITS1	0.01	0.03	–	
ITS2	−0.02	−0.02	–	
ANT	−0.01	−0.07	–	
Long Pond				
COI	***0.83***	***0.83***	***0.84***	–
ITS1	***0.64***	***0.74***	***0.67***	–
ITS2	***0.48***	***0.55***	***0.56***	–
ANT	***0.39***	***0.45***	***0.42***	–

European sequences are those from Sillerslev, Sundsøre, Juvre Deep, Vester Vedsted, The Wash, Katwijk, and Dunkerque. F_{ST} values are below diagonal (values in bold italics are significant). Above diagonal, tests of population differentiation that resulted to be significant for each marker. Significance level, $\alpha = 0.05$.

Most likely source population of European individuals

According to BAPS 5.3 analyses (Online Resource 2), the number of groups in the optimal partition was two, that is, the software did not need additional source populations to explain the diversity of European individuals. BAPS 5.3 assigned a probability to each European individual of belonging to a particular cluster (in this case, Shinnecock Bay or Cobscook Bay). Among European individuals, 49 out of 77 (63.6 %) were more likely to belong to the Cobscook Bay cluster, and the remaining 28 (36.4 %) were more likely to belong to Shinnecock Bay. Interestingly, a high number of individuals were more likely to belong to the Cobscook Bay cluster in the Limfjorden (75 %), Danish Wadden Sea (75 %), and Katwijk (73 %) areas. On the contrary, The Wash and Dunkerque individuals were more balanced (53.3 % were more likely to belong to the Shinnecock Bay cluster, and the remaining 46.7 %, to Cobscook Bay).

Neutrality tests and past changes in population size

Results of all tests performed are recorded in Table 4. Only the Fu's *Fs* test was significant in one data partition (European sequences) after applying the Bonferroni correction. It was significantly negative in the COI (*Fs* = −22.784) and ANT (*Fs* = −10.415) datasets, suggesting recent population expansion or genetic hitchhiking. Since both markers showed congruent results, and considering that it is not feasible that selection had been acting both over the nuclear and the mitochondrial regions under study (see Beaumont 2007), the past population expansion hypothesis is more likely.

The HKA test did not reveal significant deviations from neutrality of any of these two markers in any partition, meaning that the genomic regions chosen seem to have evolved neutrally.

Mismatch distributions of the pairwise number of differences were obtained for the European partition (the one in which we detected a population expansion event) (Table 5 and Online Resource 3). The ANT mismatch distribution observed curve did not perfectly fit the expected curve (the Harpending's RI was significant). However, in all other cases, tests of goodness of fit were not significant, and therefore, the observed curves fitted a population expansion model, as expected.

Estimates of time elapsed since expansion according to COI sequences were quite high, regardless of the mutation rate used. Expansion of the (hypothetical) source population of European individuals was, at the latest, 6 mya (Pleistocene), assuming all European individuals came from the same population and that there were no bottleneck effects during introduction or mass mortality events (Table 5). We should be cautious in relation to these results because the mutation rates employed could not be suitable for this species.

Table 4. Neutrality tests.

	European seq.	Shinnecock Bay	Cobscook Bay	Long Pond
COI				
N	77	21	22	28
Tajima's *D*	−1.506 (*P* = 0.032)	−1.464 (*P* = 0.056)	−0.592 (*P* = 0.308)	0.411 (*P* = 0.690)
Fu's *Fs*	***−22.784 (P = 0.000)***	−3.977 (*P* = 0.037)	−0.060 (*P* = 0.528)	−3.002 (*P* = 0.118)
Fu and Li's *D*	−1.949 (0.1 > P>0.05)	−0.470 (*P* > 0.1)	0.126 (*P* > 0.1)	−1.140 (*P* > 0.1)
Fu and Li's *F*	−2.235 (0.1 > P>0.05)	−1.007 (*P* > 0.1)	−0.325 (*P* > 0.1)	−1.073 (*P* > 0.1)
Fay and Wu's *H*	−18.064	−11.048	−8.537	−2.365
MK (Fisher's exact test)	*P* = 0.354	nct	nct	*P* = 0.354
ANT				
N	90	6	12	30
Tajima's *D*	−1.549 (*P* = 0.042)	1.032 (*P* = 0.853)	−0.684 (*P* = 0.287)	−0.422 (*P* = 0.413)
Fu's *Fs*	***−10.415 (P = 0.000)***	−1.685 (*P* = 0.021)	−1.159 (*P* = 0.143)	−1.454 (*P* = 0.201)
Fu and Li's *D*	−1.135 (*P* > 0.1)	0.883 (*P* > 0.1)	−1.553 (*P* > 0.1)	0.244 (*P* > 0.1)
Fu and Li's *F*	−1.712 (*P* > 0.1)	1.005 (*P* > 0.1)	−1.783 (*P* > 0.1)	0.051 (*P* > 0.1)
Fay and Wu's *H*	−1.113	0.267	0.697	−1.177
MK (Fisher's exact test)	nct	nct	nct	nct
COI+ANT				
HKA test (direct mode)	*P* = 0.895	*P* = 0.418	*P* = 0.778	*P* = 0.697

European sequences are those from Sillerslev, Sundsøre, Juvre Deep, Vester Vedsted, The Wash, Katwijk, and Dunkerque. Significance of some of the tests was evaluated by means of the *P* value (*P*). Significant values after the Bonferroni correction (α = 0.0125) are bold italic. *MK* McDonald-Kreitman test, *nct* the contingency table could not be computed by the software because there were no fixed differences between data sets.

Table 5. Mismatch distribution parameters and estimates of time elapsed since expansion in the European partition.

	COI	ANT
N	77	90
k	4.678	1.432
Observed variance	4.340	0.964
Θ_0	0.000	0.002
Θ_1	83.438	99999
τ	5.318	1.574
SSD	0.008 (*P* = 0.06)	0.012 (*P* = 0.07)
Harpending's RI	0.021 (*P* = 0.18)	***0.100 (P = 0.01)***
TESE (years)		
μ (%)	**0.14**	20917244
μ (%)	**0.52**	5631566

European sequences are those from Sillerslev, Sundsøre, Juvre Deep, Vester Vedsted, The Wash, Katwijk, and Dunkerque. Significance of the SSD and Harpending's RI tests were evaluated by means of the *P* value (*P*). Values in bold italics and bold are significant (α = 0.05). *k* average number of nucleotide differences, *TESE* time elapsed since expansion, *SSD* sum of squared deviations, *Harpending's RI* Harpending's raggedness index.

Evolutionary relationships among variants

The network obtained from the COI haplotype dataset (Fig. 3a) revealed two main haplogroups; one including all variants from Europe, Shinnecock Bay, and Cobscook Bay; and another one, separated by 28 mutational steps, which included all Long Pond haplotypes. Within the first haplogroup, no clear subgroups were recognised in terms of geography, that is, haplotypes from US and European sites were intermixed in the network. The ANT network (Fig. 3b) supported the separation of US/Europe and Long Pond individuals though not as clearly. In this case, all Canadian variants except one clustered apart all others. The remaining sequence-type corresponded to the one sampled in all sites except Vester Vedsted, which was the most frequent variant of the dataset. Networks performed with the ITS1, ITS2, and ITS1 + ITS2 variants (Fig. 3c, e) showed a similar picture. Long Pond sequence-types appeared clustered apart from all others in all three networks, and no subdivision by sampling site or geographic area was found among the remaining sequences (from Europe or US). Remarkably, in the case of the ITS1 + ITS2 concatenated dataset network (Fig. 3d), it was particularly evident that variants from the US and Europe were clearly distributed into two subgroups. This revealed the existence of two different ITS1-ITS2 (paralog) regions in the genomes of these animals, which were not present in the genomes of Long Pond individuals.

In the phylogenetic trees obtained from the COI + ANT + ITS1 + ITS2 concatenated dataset (Fig. 4), two reciprocally monophyletic groups were recognised, regardless of the phylogenetic method employed, and with the highest support. One group included all European sequences, those from Shinnecock Bay, and those from Cobscook Bay; and the other one, comprised all Long Pond sequences. K2P distances between those groups were: COI, 0.063 ± 0.012; ITS1, 0.028 ± 0.007; ITS2, 0.019 ± 0.006; and ANT, 0.005 ± 0.003.

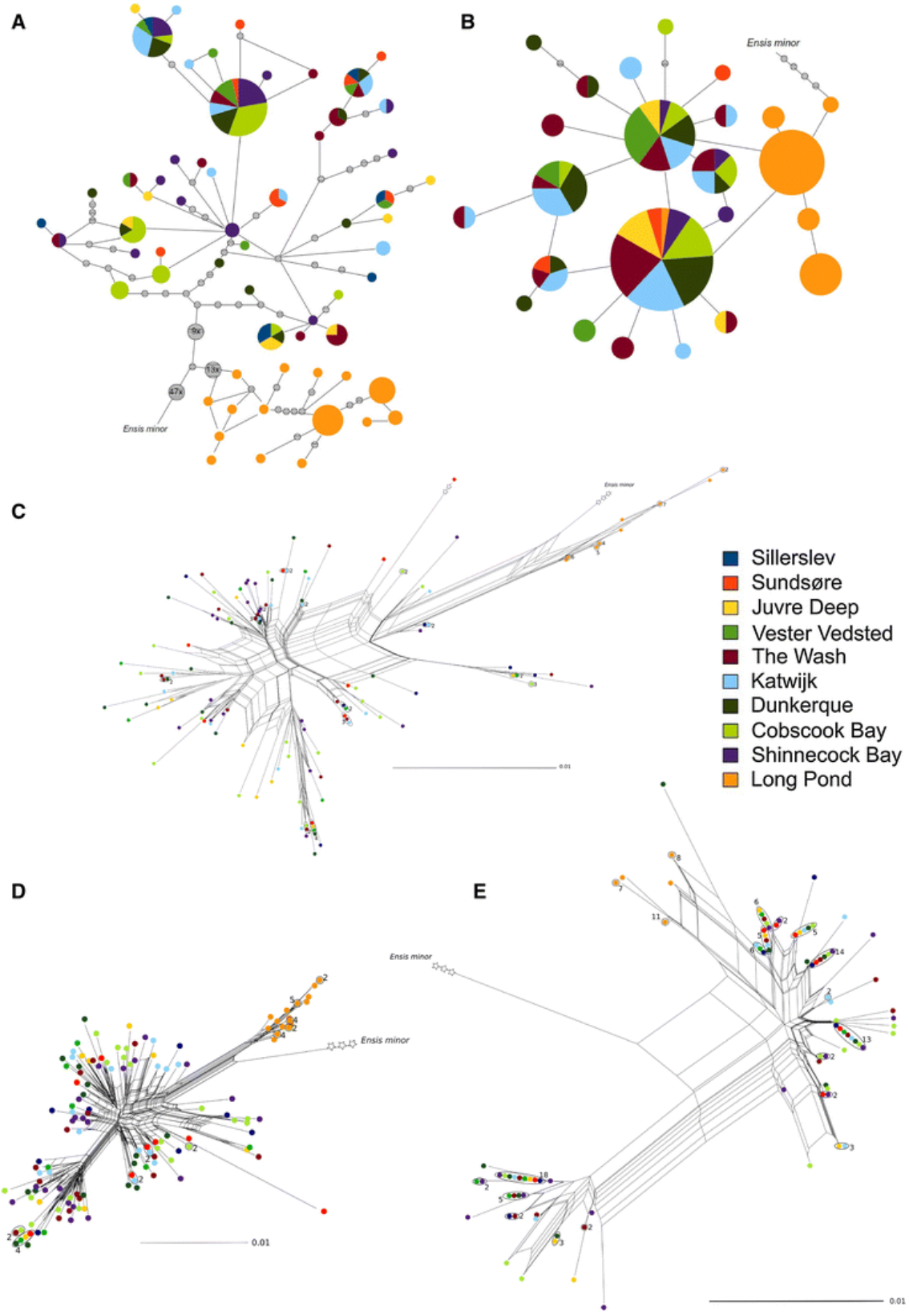

Fig. 3. Networks showing the phylogenetic relationships of obtained variants. a COI network. b ANT network. c ITS1 network. d ITS1 + ITS2 network. E ITS2 network. Each colour represents a sampling site, according to the legend. a, b Parsimony networks; the size of *circles* is proportional to variants frequency; *lined circles* are non-sampled variants inferred by the software; each line between circles represents a mutational step. c–e Distance networks; sequence-types with frequency = 1 are represented by *single dots*; sequence-types with higher frequencies are represented by more than one dot surrounded by ovals or circumferences out of which the absolute frequency is indicated by a *number*; lines between sequence-types are proportional to genetic distance; *lines* that have been shortened to fit in the figure were marked with a *star*.

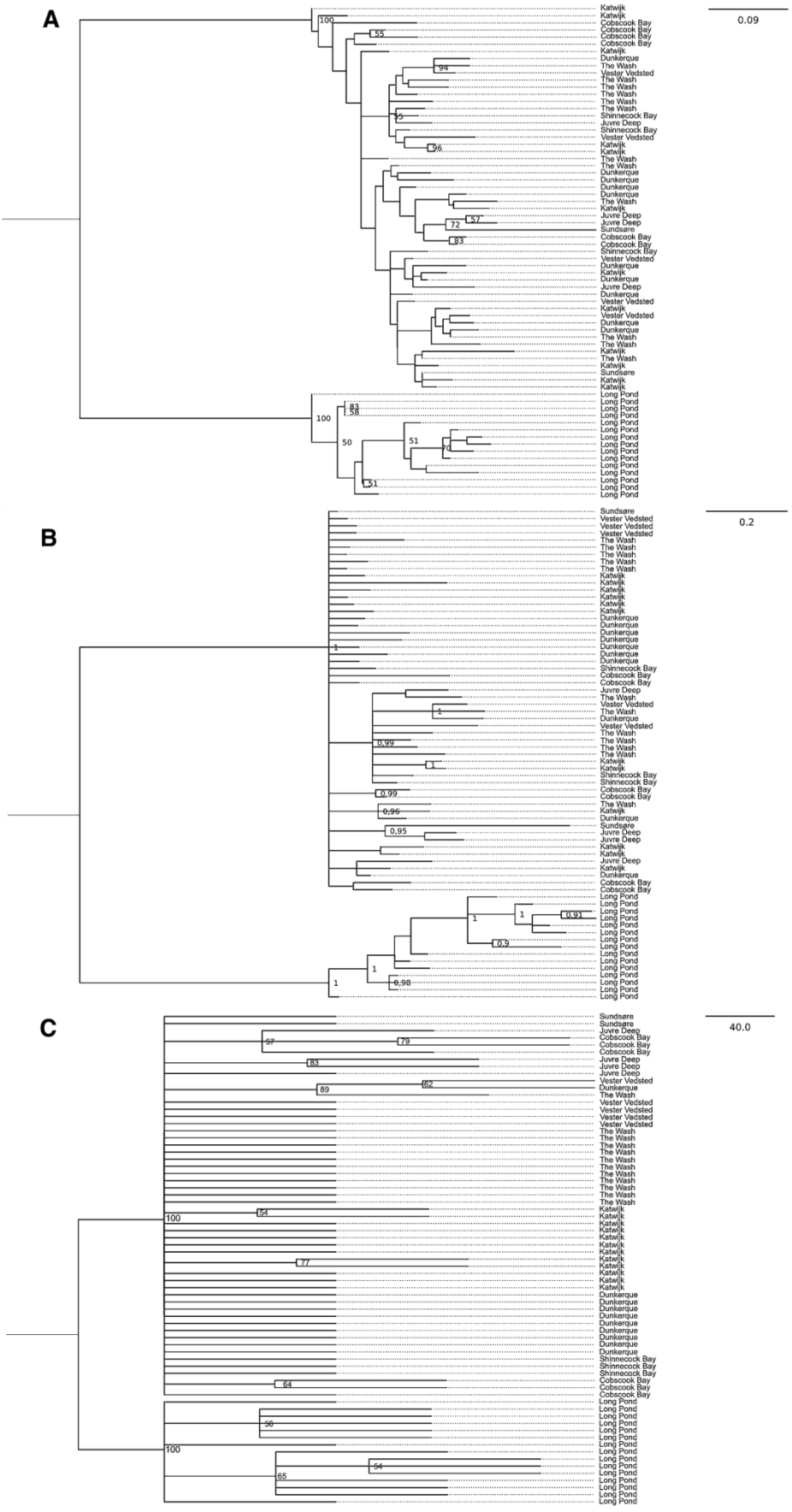

Fig. 4. Trees showing the phylogenetic relationships of sequences of the COI + ANT + ITS1 + ITS2 concatenated dataset. a Best-scoring maximum-likelihood tree (tree likelihood, −4,461.9); bootstrap values ≥50 indicated at the nodes. b Bayesian tree; posterior probability values ≥0.95 indicated at the nodes. c Maximum parsimony bootstrap 50 % majority-rule consensus tree; bootstrap values ≥50 indicated at the nodes.

Shell morphometrics

All measurements obtained and the ratios calculated from them were recorded in Online Resource 4. Remarkably, 'distance b' values were greater for *Ensis* specimens from Long Pond than from Shinnecock Bay and Cobscook Bay. There was a significant site effect ($F_{2,\ 62} = 161.854$, p value = 0, ANCOVA; homoscedasticity among residuals according to Levenes's test and equal regression slopes). For example, for a shell of 135 mm, the mean 'distance b' value (95 % confidence interval in brackets) was 16.36 mm (15.65–17.08) in Long Pond, 7.47 mm (6.67–8.27) in Shinnecock Bay, and 9.01 mm (8.13–9.89) in Cobscook Bay. In Figs. 5and 6, the relationships between 'distance b' values, and length or width were recorded per sampling site. To make an easy identification, a Long Pond individual showed a shell width (measured at the posterior adductor scar) less than twice 'distance b', whereas a shell from Shinnecock Bay or Cobscook Bay showed a width more than twice 'distance b'.

Taxonomy

Considering the genetic and morphometric results obtained in the present work, we propose to include the individuals from Long Pond within the new taxon *Ensis terranovensis* Vierna and Martínez-Lage sp.n. (Fig. 2, Online Resource 5).

Type material: MNCN-15.07/15001 (holotype; Fig. 2, Online Resource 5); MNCN-15.07/15002 (paratypes); MNHN-IM-2009-16706 to 16709 (paratypes); ZMUC-BIV-394 (paratypes); CMNML-096168 (paratypes). Type locality: Long Pond, Conception Bay, Newfoundland (Canada), 47° 30′ 54″ N, 52° 58′ 30″W. Individuals collected by scuba divers P. Sargent and R. O'Donell in 2007.

Etymology

The word *terranovensis* means 'from Newfoundland', referring to the area where the species was found.

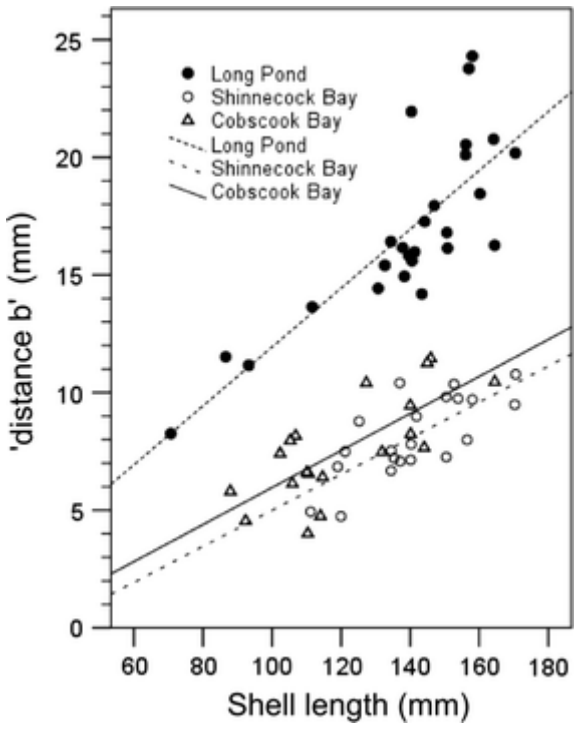

Fig. 5. The relationship of the distance along the dorsal line between the posterior end of the posterior adductor scar and the beginning of the pallial sinus ('*distance b*') and shell length of Ensis specimens. The coefficient of determination (r^2) for the linear relationship is 0.66, 0.54, and 0.56 for Long Pond, Shinnecock Bay and Cobscook Bay, respectively.

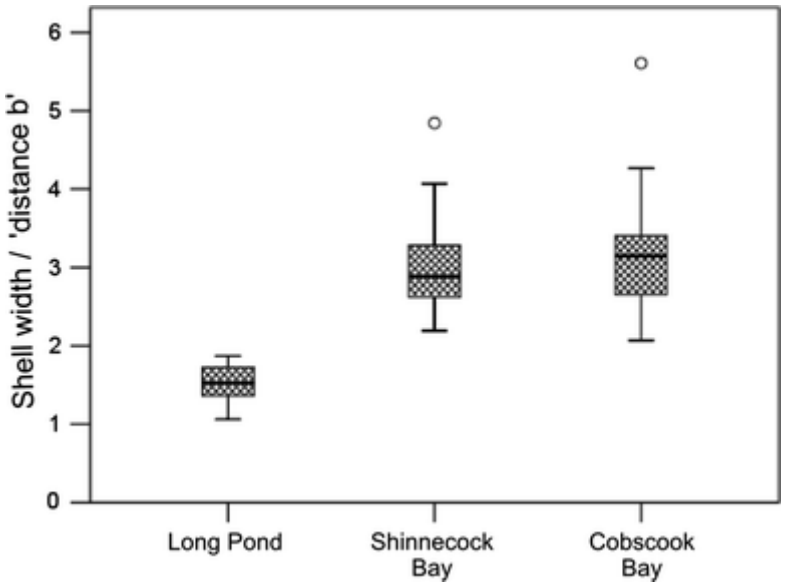

Fig. 6. Boxplot of the shell width—'*distance b*' ratio of *Ensis* specimens from Long Pond, Shinnecock Bay, and Cobscook Bay. '*Distance b*' is the distance along the dorsal line between the posterior end of the posterior adductor scar and the beginning of the pallial sinus.

Morphological description

In studied individuals (n = 26), shell length between 70.7 and 170.4 mm (mean 139.2 mm; SD 24.3 mm); shell width at the posterior adductor scar between 14.5 and 30.5 mm (mean 24.9 mm; SD 3.9 mm). Shell variable in curvature, in two individuals (whose lengths were 93.2 and 170.4 mm), straight or almost straight; in all others, curved to very curved. Valve margins parallel or almost parallel; some, slightly tapering posteriorly. Length/width ratio ranging between 4.8 and 7.2 (mean 5.6; SD 0.5). Anterior margin rounded, posterior margin truncated. Dorsal margin concave in some individuals, and almost or completely straight in others; ventral margin weakly to conspicuously convex.

Shell thick and strong (increasing with age); interior part, white coloured and slightly translucent. Posterior end of anterior adductor scar and posterior end of hinge, located at the same cross section, or more often, posterior end of anterior adductor scar, slightly surpassing posterior end of hinge towards posterior margin. Anterior adductor scar at the anterior end, narrow, broadening posteriorly; its posterior end, usually rounded, but sometimes slightly truncated. Shape of posterior adductor scar, irregularly oval. The ratio 'distance b'/'distance a', ranging between 1.1 and 2.3 (mean 1.6; SD 0.3), therefore, distance between the posterior end of the posterior adductor scar and the beginning of the pallial sinus ('distance b'), always higher than the length of the posterior adductor scar itself ('distance a'). Dorsal pallial line, much closer to dorsal margin than ventral pallial line to ventral margin. Dorsal and ventral pallial lines not completely parallel to corresponding margins, quite irregular in some individuals. Pallial sinus formed by two concave areas separated by a convex one; dorsal part of pallial sinus more concave than ventral part; in general, pallial sinus shape, similar to an irregular W.

Exterior part of the shell with irregular growth lines; conspicuous and rough in older individuals. Younger individuals, olive green, dark brown, and grey coloured; older individuals, dark brown, whitish, and in some cases, reddish.

Many individuals displaying a bulge on the most anterior hinge tooth of the left valve.

Biotope and distribution

Razor shells were collected along a channel to a harbour (Long Pond), in Conception Bay, Newfoundland, Canada, from a depth of 6–11 m. The substrate consisted of a combination of hard to loose packed mud, sand, and ground sea shells. Animals were burrowed in this material 15–30 cm. In shallower waters (6 m), substrate turned to mostly fine sand. Since the sampling of additional sites within Newfoundland was out of the scope of this work, the distributional range of the species, apart from its occurrence in Conception Bay, remains unknown.

Remarks

In terms of genetics, our results indicate *E. terranovensis* and *E. directus* are two significantly different lineages regardless of the phylogenetic method employed, according to four different nuclear and mitochondrial markers. The sequences obtained here will be useful to complement morphological identifications of newly sampled individuals. In addition, the occurrence of two paralog ITS groups in *E. directus* and not in *E. terranovensis* is a further evidence supporting strong lineage divergence.

In terms of shell morphometrics, individuals from *E. terranovensis* are easily discernible from the other *Ensis* spp., including its sister taxon *E. directus*, after a careful analysis of the valves. *E. terranovensis* is characterised by having thicker and stronger valves comparing to *E. directus*, and individuals studied appeared to have undergone slower growth. The position of the posterior adductor scar is a clear and statistically significant difference between taxa; *E. terranovensis* individuals showed a shell width less than twice the distance between pallial sinus and posterior adductor scar, whereas a *E. directus* shell from Shinnecock Bay or Cobscook Bay showed a width more than twice that distance.

It should be pointed out that one individual from Graye–sur–Mer, Calvados, France (therefore, identified as *E. directus*) was reported to have a similar distance between pallial sinus and posterior adductor scar to *E. terranovensis* (Cosel 2009, Fig. 1f). No more data about this specimen are available, and its origin remains unclear.

Discussion

Genetic variation of *E. directus* in native and introduced sites

The introduction of *E. directus* in Europe is well documented. Several ecological studies focusing on this species have been conducted in the last decades (Swennen et al. 1985; Beukema and Dekker 1995; Armonies and Reise 1999; Cadée 2000; Palmer 2004; Krakau et al. 2006; Cardoso et al. 2009; Freudendahl et al. 2010; Tulp et al. 2010; Dannheim and Rumohr 2011; Dekker and Beukema 2012), but so far there have been no reports about the genetic variation of native and introduced populations except our report on ITS sequence variation of *Ensis* species (Vierna et al. 2010). After having recognised that the *Ensis* specimens from Long Pond belong to a new species, the conclusion about the applicability of ITS to differentiate among individuals from different geographic areas cannot be maintained.

One of our goals was to obtain preliminary information on the origin of European individuals. However, to identify the geographic source of introduced populations, the native range of the species must be thoroughly sampled and potential source populations must be sufficiently differentiated (Fitzpatrick et al. 2012). This is an important issue since the detection of source populations will depend not only on sampling intensity but also on the differentiation of these populations along the native range of the species. Indeed, our results suggest that both Shinnecock Bay and Cobscook Bay could be the origin of European individuals, since all of them could be assigned with high probability to one of these US sites. The site of Long Pond was

completely discarded as potential source population. Nonetheless, if data from a third (southern) potential source population become available, these results could be re-interpreted. Despite several attempts, we failed to obtain samples from the southern native range of the species that otherwise might have contributed to a more complete description of the genetic variation in the native range. Armonies and Reise (1999) discussed the possibility of linking the mass mortality events with the origin of European individuals. According to them, these individuals may have originated from an American population at the southern limit of its distributional range that may not be adapted to cold winter conditions supposed to cause mass mortality in Europe. However, it is unknown either to what extent native populations are subjected to mass mortality, and the role that factors other than storms may play in the mass mortality events, both in the native and in the introduced ranges. The higher differentiation of the European population to native ones compared to differentiation between Shinnecock Bay and Cobscook Bay might support that the source of European individuals is at the southern limit of the native range.

In their review on the genetic variation of various taxa in native and introduced populations, Dlugosch and Parker (2008) concluded that genetic variation is usually lower in introduced areas. However, according to Holland (2000), if the introduction involves a large (more than 1,000), genetically diverse assortment of individuals, we might expect to see little or no reduction in heterozygosity and allelic diversity relative to the gene pool of the source population. Here, after having sampled the main parts of the species' native and introduced ranges, we show that the European population as a whole, far from displaying low levels of genetic variation, is more variable than native sites in the US. In the same way, several sampling sites within the European range are more variable than native sites. For instance, Dunkerque is almost twice as variable as Cobscook Bay in terms of COI nucleotide diversity. The sites from the Limfjorden area are even more variable.

Taking into account the intensity of transatlantic shipping, the hypothesis of multiple-introduction events seems rather likely. The higher probability of Limfjorden, Danish Wadden Sea, and Katwijk to be assigned to Cobscook Bay seems to support this hypothesis. Multiple introductions appear to be common and contribute to increase variation (Dlugosch and Parker 2008). In this sense, the higher levels of variation detected in Europe agree with a multiple-introduction scenario. On the contrary, the time-line for the appearance of *E. directus* along the European coast suggests that colonisation happened gradually (for colonisation pattern see Cosel 2009). The very recent findings of isolated populations in western Britain (Paul Dansey personal communication; Dansey 2011) and northern Spain (Arias and Anadón 2012) could be a result of dispersion from European populations but it might also be a result of new introductions from native sites.

Apart from the already mentioned difficulties in determining the precise source population(s) and assessing the number of introductions, there is another factor that should be considered to understand the observed genetic variation both at native and introduced sites. The genetic variation of potential source populations that we have measured might not be the same as it was when introduction took place. In fact, temporal fluctuations in genetic variation have been reported for other *Ensis* (Varela et al. 2009, 2011). Even though it would have been ideal to have analysed individuals sampled in the native populations around 1978 (when the first migrants arrived into European waters), such samples are, to our best knowledge, unavailable. In the same way, fluctuations of genetic variation in the introduced range are also possible. In fact, Hedgecock and Pudovkin (2011) stressed the importance of studying temporal stability of genetic variation when conducting marine population genetic studies, but this is rarely done due to considerable time and economic limitations.

Therefore, we can conclude that (1) individuals from European sites can potentially be assigned to either Shinnecock Bay and Cobscook Bay but not to Long Pond. (2) Multiple introductions seem likely, but our data cannot prove this. And (3) potential bottlenecks during introduction(s) and mass mortality events seem not to have affected genetic variation in the introduced range.

Within the native range, the northern site (Cobscook Bay) is much less variable than Shinnecock Bay according to mitochondrial DNA, but not according to nuclear DNA, a fact that is quite intriguing. It is usual that northern populations in the Northern Hemisphere display less neutral genetic variation due to the effect of Quaternary glaciations (even though in marine animals, there are several reports of populations surviving in northern periglacial refugia, see Maggs et al. 2008; Krakau et al. 2012). However, nuclear and mitochondrial loci should be concordant. Therefore, a mitochondrial selective sweep in Cobscook Bay might be the cause of the reduced genetic variation. But neutrality tests did not detect deviations from neutrality so this question remains unanswered.

Our results suggest that F_{ST} is more conservative than tests of population differentiation to detect genetic differences among sites. Both Shinnecock Bay and Cobscook Bay show little but significant differentiation from Europe (considered as a representative of the source population) according to COI in the F_{ST} analyses, and to ITS1 in the test of population differentiation. Since the only significant comparison obtained between Shinnecock Bay and Cobscook Bay sites was the COI test of population differentiation, they seem to be less differentiated to each other than to Europe.

An absence of population structure over hundreds of kilometres of coast is not unexpected in populations of marine bivalves (e.g. Strasser and Barber 2008; Baker et al. 2008; Arnaud-Haond et al. 2008) since they are often characterised by frequent gene flow among sites. Nonetheless, there are also several examples of structure (e.g. Arnaud-Haond et al. 2008; Xiao et al. 2010; Mao et al. 2011). Dispersal and therefore gene flow is facilitated by external fertilisation and a planktonic larval stage but could also be reduced by marine currents and other physical barriers. The Western Atlantic Boreal Region extends from the Strait of Belle Isle to Cape Hatteras (Briggs and Bowen 2012). In this region, some main barriers to gene flow have been described, namely the Laurentian Channel, Cape Cod, and Cape Hatteras itself (see Fig. 1). The weak but significant population differentiation that we detected between Shinnecock Bay and Cobscook Bay may support Cape Cod as a barrier to gene flow.

The four sequence-based markers selected in this work showed different degrees of conservation, but all of them were informative both at the population and species levels. Indeed, the degree of conservation is expected to vary among different genomic regions. For example, in phylogenetics, more conserved genes are usually employed to resolve internal nodes in the phylogeny, whereas those more variable are, in general, able to resolve the external nodes. Here, we have shown that COI and ITS1 regions were more polymorphic than ITS2 and ANT. Population genetic results based on these four markers were, in general, concordant. Even though ITS2 and ANT sequences did not show extreme sequence differences among *E. directus* and *E. terranovensis* individuals (as COI and ITS1 did), they still separated both lineages, with the exception of the shared ANT sequence-type, a phenomenon known as incomplete lineage sorting. Because of this phenomenon, some haplotypes can remain identical in two isolated gene pools, a situation that mainly occurs when divergence is recent (April et al. 2011).

A new *Ensis* species

Taxonomy and systematics of *Ensis* razor shells have traditionally been based on shell morphology such as continuous shell characters (e.g. valve shape, length and shape of muscle scars, shell colour) (see Cosel 2009 and references therein), which are often overlapping among species. This absence of autapomorphies makes *Ensis* spp. a good candidate for combined studies of genes and morphometrics.

Though several species concepts have been proposed since the 1940s there is a working definition that considers species as separately evolving lineages. Kevin de Queiroz (2007) has proposed a unified species concept that treats this property as the only necessary property of species. Here, we demonstrate that *directus* and *terranovensis* are separately evolving lineages. There is a high genetic divergence between individuals belonging to *E. terranovensis* (the Long Pond population) and individuals belonging to *E.*

directus (European, Shinnecock Bay, and Cobscook Bay sites) according to both nuclear and mitochondrial DNA that can only be explained by a two-lineages scenario. This divergence was confirmed by the morphometric analysis carried out.

In a recent paper, Kong et al. (2012) studied COI and ITS1 sequence data of the marine bivalves *Macrodiscus* spp. They described new species based on these two molecular markers and morphometrics, supporting the suitability of combining morphometrics and DNA to clarify taxonomy. K2P distances between *E. directus* and *E. terranovensis* were intermediate compared to the ones obtained for *Macrodiscus* spp.

The Laurentian Channel (Fig. 1) seems important in the speciation process of *E. directus* and *E. terranovensis* as it is a main barrier to gene flow between Coobscook Bay and Long Pond. This channel (>400 m depth) is a geographic barrier separating the Scotian Shelf and Newfoundland Shelf marine ecosystems (Sargent et al. 2008 and references therein) and seems to be a good candidate for vicariance.

Baker et al. (2008) detected significant (weak) population differentiation of the bivalve *Mercenaria mercenaria* between a population from Prince Edward Island (Gulf of St. Lawrence, Canada) and New York (US), and Hare and Weinberg (2005) found significant and strong genetic differentiation between a Îles de la Madeleine (Gulf of St. Lawrence) population of the bivalve *Spisula solidissima solidissima* and other Atlantic US populations. In the same way, Kenchington et al. (2006) detected significant genetic differentiation between Canadian populations of *Placopecten megallanicus* (at both sides of the Laurentian Channel), and Atlantic US populations. These examples support a genetic isolation of Gulf of St. Lawrence populations in bivalve species that might have facilitated speciation in *Ensis*. Nonetheless, it is unknown whether the two species co-occur somewhere in the NW Atlantic region. So, further studies on the present distribution of *E. directus* in its northernmost range and that of *E. terranovensis* are needed to unveil the evolutionary history of both taxa.

Conclusion

We found genetic variation at mitochondrial and nuclear markers in *E. directus* in its introduced (European) range to be higher than in native (North American) populations. This contrasts with our initial expectation of a strong effect of random genetic drift due to both potential bottlenecks during introduction(s) and to mass mortalities observed several times in the introduced range. The observed patterns of genetic variation could be due to temporal fluctuations of genetic variation; to the fact that potential bottlenecks and mass mortalities might not have affected genetic variation in the introduced range; and to the occurrence of multiple introductions. The hypothesis of multiple introductions seems likely since it is supported by trained clustering analysis and by the intensity of transatlantic shipping. However, it contrasts with the observed gradual colonisation of European coastal waters. Population genetic analyses enabled us to identify a very divergent population from Newfoundland (Canada). Based on genetic and morphometric evidence, the examined specimens from this population belong to a new *Ensis* species that we described and named *E. terranovensis*.

Acknowledgments

We are very grateful to Rudo von Cosel for his support and his drawing of *E. terranovensis* holotype, and to the following colleagues who helped us in some way or another during the execution of this work: André Martel, Anja Schulze, Barbara Buge, David Palmer, Diego Fonataneto, Ferruccio Maltagliati, Horacio Naveira, Jean-Marc Gagnon, Jean-Marie Dewarumez, Jeroen Goud, Jukka Corander, Manuel Pimentel, Mark

Graham, Marta Vila, Neus Marí, Nicolas Puillandre, Ole S. Tendal, Paul Dansey, Philip Sargent, Rafael Araújo, Ray J. Thompson, Robert O'Donnell, Stephen T. Tettelbach, Tim Sheehan, and Tom Schioette. We would also like to thank two anonymous reviewers for critical and helpful comments. JV has been supported by a 'María Barbeito' fellowship and a travel grant, both from the Consellería de Economía e Industria, Xunta de Galicia (Spain) and the European Social Fund.

References

April J, Mayden RL, Hanner RH, Bernatchez L (2011) Genetic calibration of species diversity among North America's freshwater fishes. PNAS 108:10602–10607

Arias A, Anadón N (2012) First record of *Mercenaria mercenaria* (Bivalvia: Veneridae) and *Ensis directus* (Bivalvia: Pharidae) on Bay of Biscay, Iberian Peninsula. J Shellfish Res 31:57–60

Armonies W (2001) What an introduced species can tell us about the spatial extension of benthic populations. Mar Ecol Prog Ser 209:289–294

Armonies W, Reise K (1999) On the population development of the introduced razor clam *Ensis americanus* near the island pf Sylt (North Sea). Helgolan Mar Res 52:291–300

Arnaud-Haond S, Vonau V, Rouxel C, Bonhomme F, Prou J, Goyard E, Boudry P (2008) Genetic structure at different spatial scales in the pearl oyster (*Pinctada margaritifera cumingii*) in French Polynesian lagoons: beware of sampling strategy and genetic patchiness. Mar Biol 155:147–157

Audzijonyte A, Vrijenhoek RC (2010) Three nuclear genes for phylogenetic, SNP and population genetic studies of molluscs and other invertebrates. Mol Ecol Resour 10:200–204

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

Beaumont MA (2007) Conservation genetics. In: Balding DJ, Bishop M, Cannings C (eds) Handbook of statistical genetics, 3rd edn. Wiley, Chichester, pp 1021–1066

Beukema JJ, Dekker R (1995) Dynamics and growth of a recent invader into European coastal waters: the American razor clam, *Ensis directus*. J Mar Biol Assess UK 75:351–362

Briggs JC, Bowen BW (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. J Biogeogr 39:12–30

Bryant D, Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. Mol Biol Evol 21:255–265

Cadée GC (2000) Herring gulls feeding on a recent invader in the Wadden Sea, *Ensis directus*. In: Harper EM, Taylor JD, Crame JA (eds), *The evolutionary biology of the Bivalvia*. Geol Soc Lond Special Publ 177:459–464

Cardoso JFMF, Witt JIJ, van der Veer HW (2009) Reproductive investment of the American razor clam *Ensis americanus* in the Dutch Wadden Sea. J Sea Res 62:295–298

Cheng L, Connor TR, Aanensen DM, Spratt BG, Corander J (2011) Bayesian semi-supervised classification of bacterial samples using MLST databases. BMC Bioinformatics 12:302. doi:10.1186/1471-2105-12-302

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1660

Corander J, Marttinen P, Mäntyniemi S (2006) Bayesian identification of stock mixtures from molecular marker data. Fis B-NOAA 104:550–558

Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9:539

Cosel von R (2009) The razor shells of the eastern Atlantic, part 2. Pharidae II: the genus *Ensis* Schumacher, 1817 (Bivalvia, Solenoidea). Basteria 73:1–48

Dannheim J, Rumohr H (2011) The fate of an immigrant: *Ensis directus* in the eastern German Bight. Helgol Mar Res. doi:10.1007/s10152-011-0271-2

Dansey P (2011) *Ensis directus* (Conrad 1843) (Bivalvia: Solenoidea) found in Liverpool Bay (Sea area S24). J Conchol 40:679

de Queiroz K (2007) Species concepts and species delimitation. Syst Biol 56:879–886

Dekker R, Beukema JJ (2012) Long-term dynamics and productivity of a successful invader: The first three decades of the bivalve *Ensis directus* in the western Wadden Sea. J Sea Res http://dx.doi.org/10.1016/j.seares.2012.04.004

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol Ecol 17:431–449

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2011) Geneious v5.4. Available from http://www.geneious.com

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res 10:564–567

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

Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. Genetics 155:1405–1413

Fitzpatrick BM, Fordyce JA, Niemiller ML, Reynolds RG (2012) What can DNA tell us about biological invasions? Biol Invasions 14:245–253

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

Freudendahl ASL, Nielsen MM, Jensen T, Jensen KT (2010) The introduced clam *Ensis americanus* in the Wadden Sea: field experiment on impact of bird predation and tidal level on survival and growth. Helgoland Mar Res 64:93–100

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

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics 133:693–709

Giegerich R, Meyer F, Schleiermacher C (1996) GeneFisher—software support for the detection of postulated genes. Proc Int Conf Intell Syst Mol Biol 4:68–77

Goudet J, Raymond M, de Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. Genetics 144:1933–1940

Griffiths-Jones S (2005) RALEE—RNA ALignment editor in Emacs. Bioinformatics 21:257–259

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321

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

Hare MP, Weinberg JR (2005) Phylogeography of surfclams, *Spisula solidissima*, in the western North Atlantic based on mitochondrial and nuclear DNA sequences. Mar Biol 146:707–716

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum Biol 66:591–600

Heath DD, Rawson PD, Hilbish TJ (1995) PCR-based nuclear markers identify alien blue mussel (*Mytilus* spp.) genotypes on the west coast of Canada. Can J Fish Aquat Sci 52:2621–2627

Hedgecock D, Pudovkin AI (2011) Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. Bull Mar Sci 87:971–1002

Holland BS (2000) Genetics of marine bioinvasions. Hydrobiologia 420:63–71

Hudson R, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. Genetics 116:153–159

Huelsenbeck JP, Ronquist F (2001) MRBAYES: bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267

Kenchington E, Duggan R, Riddell T (1998) Early life history characteristics of the razor clam (*Ensis directus*) and the moonsnails (*Euspira* spp.) with applications to fisheries and aquaculture. Can Tech Rep Fish Aquat Sci 2223:1–32

Kenchington EL, Patwary MU, Zouros E, Bird CJ (2006) Genetic differentiation in relation to marine landscape in a broadcast-spawning bivalve mollusc (*Placopecten magellanicus*). Mol Ecol 15:1781–1796

Kong L, Matsukuma A, Hayashi I, Takada Y, Li Q (2012) Taxonomy of *Macridiscus* species (Bivalvia: Veneridae) from the western Pacific: insight based on molecular evidence, with description of a new species. J Moll Stud 78:1–11

Krakau M, Thieltges DW, Reise K (2006) Native parasites adopt introduced bivalves of the North Sea. Biol Invasions 8:919–925

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

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, López R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

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. Mol Ecol 12:2215–2229

Maggs CA, Castilho R, Foltz D, Henzler C, Jolly MT, Kelly J, Olsen J, Perez KE, Stam W, Väinölä R, Viard F, Wares J (2008) Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. Ecology 89:S108–S122

Maine Sea Grant (2012) Accessed at http://www.seagrant.umaine.edu/resources-for-shellfish-growers/species/razor-clam. On 01 Feb 2012

Mao Y, Gao T, Yanagimoto T, Xiao Y (2011) Molecular phylogeography of *Ruditapes philippinarum* in the northwestern Pacific Ocean based on COI gene. J Exp Mar Biol Ecol 407:171–181

Marine Stewardship Council (2012) Accessed at http://www.msc.org/track-a-fishery/in-assessment/north-east-atlantic/dfa-dutch-north-sea-ensis. On 01 Feb 2012

McDonald J, Kreitman M (1991) Adaptative protein evolution at *adh* locus in *Drosophila*. Nature 351:652–654

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the gateway computing environments workshop (GCE), 14 November 2010, New Orleans, LA pp 1–8

Mühlenhardt-Siegel U, Dörjes J, von Cosel R (1983) Die amerikanische Schwertmuschel *Ensis directus* (Conrad) in der Deutschen Bucht: 2. Populationsdynamik. Senckenb Marit 15:93–110

Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24:581–583

Palmer DW (2004) Growth of the razor clam *Ensis directus*, an alien species in the Wash on the east coast of England. J Mar Biol Assess UK 84:1075–1076

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49:1280–1283

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569

Sargent PS, Methven DA, Hooper RG, McKenzie CH (2008) A range extension of the Atlantic silverside, *Menidia menidia*, to coastal waters of southwestern Newfoundland. Can Field Nat 122:338–344

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690

Stamatakis A, Hoover P, Rougemont J (2008) A fast bootstrapping algorithm for the RAxML web-servers. Syst Biol 57:758–771

Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet 76:449–462

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989

Strasser CA, Barber PH (2008) Limited genetic variation and structure in softshell clams (*Mya arenaria*) across their native and introduced range. Conser Genet 10:803–814

Swennen C, Leopold MF, Stock M (1985) Notes on growth and behaviour of the American razor clam *Ensis directus* in the Wadden Sea and the predation on it by birds. Helgolan Mar Res 39:255–261

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version. Sinauer Associates, Sunderland, MA, USA

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Tulp I, Craeymeersch J, Leopold M, van Damme C, Fey F, Verdaat H (2010) The role of the invasive bivalve *Ensis directus* as food source for fish and birds in the Dutch coastal zone. Estuar Coast Shelf Sci 90:116–128

Varela MA, Martínez-Lage A, González-Tizón AM (2009) Temporal genetic variation of microsatellite markers in the razor clam *Ensis arcuatus* (Bivalvia: Pharidae). J Mar Biol Assess UK 89:1703–1707

Varela MA, Martínez-Lage A, González-Tizón AM (2011) Genetic heterogeneity in natural beds of the razor clam *Ensis siliqua* revealed by microsatellites. J Mar Biol Assess UK 92:171–177

Vierna J, Martínez-Lage A, González-Tizón AM (2010) Analysis of ITS1 and ITS2 sequences in *Ensis* razor shells: suitability as molecular markers at the population and species levels, and evolution of these ribosomal DNA spacers. Genome 53:23–34

von Cosel R, Dörjes J, Mühlenhardt-Siegel U (1982) Die amerikanische Schwertmuschel *Ensis directus* (Conrad) in der Deutschen Bucht. I. Zoogegraphis und Taxonomie im Vergleich mit den einheimischen Schwertmuschel-Arten. Senckenbergiana marit 14:147–173

Xiao J, Cordes JF, Wang H, Guo X, Reece KS (2010) Population genetics of *Crassostrea ariakensis* in Asia inferred from microsatellite markers. Mar Biol 157:1767–1781