Post-glacial colonization of northern coastal habitat by bottlenose dolphins: A marine leading-edge expansion?
Nykänen, Milaja; Kaschner, Kristin; Dabin, Willy; Brownlow, Davison, Nicholas J; Deaville, Rob; Garilao, Christina; Kesner-Reyes, Kathleen; Gilbert, M. Thomas P.; Penrose, Rod; Islas-Villanueva, Valentina; Wales, Nathan; Ingram, Simon; Rogan, Emer; Louis, Marie; Foote, Andrew

Journal of Heredity

DOI: 10.1093/jhered/esz039
Published: 01/09/2019

Peer reviewed version

Dyfniad o’r fersiwn a gyhoeddwyd / Citation for published version (APA):

Hawliau Cyffredinol / General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Post-glacial colonization of northern coastal habitat by bottlenose dolphins: A marine leading-edge expansion?

Milaja Nykänen1*, Kristin Kaschner2, Willy Dabin3,4, Andrew Brownlow5, Nicholas J. Davison6, Rob Deaville6, Cristina Garilao7, Kathleen Kesner-Reyes8, M. Thomas P. Gilbert9, Rod Penrose10, Valentina Islas-Villanueva11, Nathan Wales9, Simon N. Ingram12, Emer Rogan1, Marie Louis3,13†, and Andrew D. Foote1,9†

†Joint senior authors

1 School of Biological, Earth and Environmental Sciences, University College Cork, Distillery Fields, North Mall, Cork, Ireland
2 Department of Biometry and Environmental System Analysis, Faculty of Environment and Natural Resources, University of Freiburg, Tennenbacher Straße 4, 79106 Freiburg, Germany
3 Centre d’Études Biologiques de Chizé, UMR 7372 CNRS-Université de La Rochelle, France
4 Observatoire Pelagis, UMS 3462, Université de La Rochelle, CNRS, 5 allées de l’Océan, La Rochelle, France
5 Scottish Marine Animal Stranding Scheme, SRUC Veterinary Services, Drummondhill, Stratherrick Road, Inverness, IV2 4JZ, UK
6 UK Cetacean Strandings Investigation Programme, The Wellcome Building, Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RY, UK
7 GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany
8 Quantitative Aquatics (Q-quatics), IRRI Khush Hall, College, Laguna 4031, Philippines
9 Section for Evolutionary Genomics, Department of Biology, University of Copenhagen, Øster Farimagsgade 5A, DK-1353 Copenhagen, Denmark
10 Marine Environmental Monitoring, Penwalk, Llechryd, Cardigan, Ceredigion, Wales, SA43 2PS, UK
11 CONACYT-Universidad del Mar, Cd. Universitaria s/n, San Pedro Pochutla 70902, Oaxaca, Mexico
12 Marine Vertebrate Research Group, School of Biological and Marine Sciences, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK
13 Scottish Ocean Institute, East Sands, St Andrews KY16 8LB, UK

Running title: Leading-edge expansion by bottlenose dolphins

*Correspondence to be sent to: Milaja Nykänen, School of Biological, Earth and Environmental Sciences, Distillery Fields, North Mall, University College Cork, Cork, Ireland, E-mail: m.nykanen@ucc.ie, Tel: +353 (0)21 4904617

Abstract

Oscillations in the Earth's temperature and the subsequent retreating and advancing of ice-sheets around the polar regions are thought to have played an important role in shaping the distribution and genetic structuring of contemporary high-latitude populations. After the Last Glacial Maximum (LGM), retreating of the ice-sheets would have enabled early colonizers to
rapidly occupy suitable niches to the exclusion of other conspecifics, thereby reducing genetic diversity at the leading-edge. Bottlenose dolphins (genus *Tursiops*) form distinct coastal and pelagic ecotypes, with finer scale genetic structuring observed within each ecotype. We reconstruct the post-glacial colonization of the Northeast Atlantic (NEA) by bottlenose dolphins using habitat modelling and phylogenetics. The AquaMaps model hindcasted suitable habitat for the LGM in the Atlantic lower latitude waters and parts of the Mediterranean Sea. The time-calibrated phylogeny, constructed with 86 complete mitochondrial genomes including 30 generated for this study and created using a multi-species coalescent model, suggests that the expansion to the available coastal habitat in the NEA happened via founder events starting ~15,000 years ago (95% highest posterior density interval: 4,900–26,400). The founders of the two distinct coastal NEA populations comprised as few as two maternal lineages that originated from the pelagic population. The low effective population size and genetic diversity estimated for the shared ancestral coastal population subsequent to divergence from the pelagic source population are consistent with leading-edge expansion. These findings highlight the legacy of the Late Pleistocene glacial cycles on the genetic structuring and diversity of contemporary populations.

**Keywords:** Last Glacial Maximum (LGM), multi-species coalescent, time-dependency, phylogenetics, genetic diversity, habitat modelling
Introduction

During the Late Quaternary period (1 Ma to present) the Earth’s climate was governed by a series of glacial and interglacial events and temperature fluctuations that occurred at approximately 100,000-year intervals (Shackleton, 2000). These glacial cycles are thought to have played an important role in shaping the current distribution and genetic structuring of species and populations. In the Northern Hemisphere, this is more pronounced at high latitudes where the presence of ice sheets and Arctic temperatures during cold stadial periods restricted the available habitat to warmer refugia for many temperate-adapted species (Hewitt, 2000). The distribution of temperate-adapted species at high latitudes is thought to have been characterized by cyclical range contractions and expansions throughout the Pleistocene (Hewitt, 2000), and present-day populations represent lineages descended from relict populations that survived in glacial refugia (Hofreiter & Barnes, 2010). However, the capacity to adapt to environmental change varies among species and populations through their dispersal ability, genetic diversity and generation time (Stewart et al., 2010; Montgomery et al., 2014; Younger et al., 2016) and has likely affected their present-day distribution.

Hewitt (1999, 2000) proposed that colonization of terrestrial species during the warm interglacial periods typically occurred via long-range dispersal events, or ‘leading-edge range expansions’. During these events a few individuals from the source population, carrying a subset of the genetic diversity, first colonized and then expanded their range to fill emerging geographic and ecological niches leading to the exclusion of secondary waves of colonizers and thereby reducing genetic variability at this leading-edge. Unlike most terrestrial mammal species, marine mammals have a relatively low cost of movement (Tucker, 1975; Williams et al., 1992; Williams, 1999) and few geographic barriers to dispersal, so the leading-edge colonization model may, or may not, be realistic for such highly mobile organisms that are able to move thousands of kilometers within a few weeks (Gabriele et al., 1996; Mate et al., 1997). High mobility certainly makes them capable of undertaking long-distance dispersal events, but at the same time, also increases the potential to retain ongoing gene flow and migration between the leading-edge and source populations. Compared to terrestrial species, little is known about the post-glacial colonization history of Europe’s marine fauna, but during recent years a growing number of studies have explored the dynamics of postglacial range expansion in marine taxa (e.g. de Bruyn et al., 2009; Fontaine et al., 2010, 2014;
Catchen et al., 2013; Foote et al., 2013). Genetic analyses of carbon-dated sub-fossils indicate that temperate climate adapted species such as gray whales *Eschrichtius robustus* (Lilljeborg, 1861) and North Atlantic right whales *Eubalaena glacialis* (Müller, 1776) replaced cold-adapted species such as bowhead whales *Balaena mysticetus* (Linneaus, 1758) in mid-latitude European waters during the Late Pleistocene-Early Holocene transition (Foote et al., 2013; Alter et al., 2015). However, the refugial distribution during the Last Glacial Maximum (LGM) of the source populations for these re-colonizing temperate marine species is largely unknown..

Here, a multi-species coalescent model was used together with information on species and population structure to reconstruct the post-glacial colonization in a highly mobile marine genus *Tursiops*. Specifically, the aim was to determine the timing of the colonization into the northeastern Atlantic coastal habitats and to investigate whether the colonization occurred via leading-edge expansion. The genus *Tursiops* currently consists of two separate species recognized by the International Union for the Conservation of Nature, *T. truncatus* (Montagu, 1821; common bottlenose dolphin) and *T. aduncus* (Ehrenberg, 1833; Indo-Pacific bottlenose dolphin), although *T. australis* (Burrunan dolphin) inhabiting the coastal waters in southern Australia has been proposed as a separate species (Charlton-Robb et al., 2011). For simplicity, we refer to the latter as *T. australis* throughout this paper but acknowledge the unofficial status of this putative species. Species in the *Tursiops* genus are found in coastal inshore waters, continental shelf regions and open ocean environments (Wells & Scott, 2009) in tropical and temperate waters (Leatherwood & Reeves, 1989). Throughout much of their range, they exhibit a hierarchical population structure, with the greatest genetic divergence found between pelagic and coastal populations (Hersh & Duffield, 1990; Curry & Smith, 1998; Hoelzel et al., 1998; Natoli et al., 2005; Tezanos-Pinto et al., 2008; Louis et al., 2014b; a; Lowther-Thieleking et al., 2015; Nykänen et al., 2018). The broad-scale population structuring between coastal and pelagic ecotypes in the Northeast Atlantic (NEA) (Louis et al., 2014b) likely reflects colonization of emerging or available inshore habitats after the LGM from a pelagic source population followed by divergence in allopatry (Moura et al., 2013; Louis et al., 2014a). It has been suggested that the colonization of coastal habitats in the NEA was possibly achieved via a founder event by a small number of individuals (Louis et al., 2014a). However, the possible role of leading-edge colonization has not been fully investigated in this genus by quantifying historic population sizes. Therefore, to better understand the climatic, temporal and spatial context of the evolutionary processes that gave
rise to the present-day diversity and habitat use of bottlenose dolphins in the NEA, mitogenomic data were analyzed in combination with habitat mapping, using newly available tools that can robustly infer ancestral effective population sizes and therefore give insights to colonization processes.

Methods

Models of suitable habitat

Following the methods in Foote et al. (2013), Alter et al. (2015) and Morin et al. (2015), the AquaMaps modelling approach (Ready et al., 2010; Kaschner et al., 2011) was used to predict the distribution of suitable habitat for common bottlenose dolphins during the present, for the year 2100 (under the Institut Pierre Simon Laplace climate scenario SRES A2), and for the LGM (~20,000 ya). AquaMaps is a bioclimatic model that combines large sets of occurrence data (e.g. from visual observations) with available expert knowledge on species preference and tolerance to different environmental parameters and ultimately generates large-scale predictions of the probability of occurrence for different marine species. This way the preferred habitat of a species can be estimated based on a predefined set of environmental parameters including water depth, sea surface temperature, salinity, primary production, sea ice concentration and proximity to land, and projected into geographic space as relative probability of occurrence. For the purpose of this study, a slightly modified version of the original AquaMaps model (which is available on www.aquamaps.org) was used; specifically, primary production was excluded from the model, as there are no available data for this parameter for the Pliocene or Pleistocene. Current environmental conditions were assumed to be representative of the entire Holocene as the variability in conditions during this time period has been small compared to the differences between glacial cycles (Folland et al., 2002). AquaMaps has been previously used to hindcast suitable habitat predictions for bowhead whales, gray whales and killer whales, Orcinus orca (Linnaeus, 1758), during the LGM (Foote et al., 2013; Alter et al., 2015; Morin et al., 2015) by using mean annual environmental conditions during the LGM based on the GLAMAP project data set (Schäfer-Neth & Paul, 2003).
Sample collection

Tissue samples from twelve individuals from *T. truncatus* were obtained by biopsy sampling free-ranging animals (see Krützen *et al.*, 2002) in coastal waters of Ireland and France in 2005–2012. Additionally, skin samples were collected from 21 individuals that stranded along the coast of Ireland, France and the UK in 1991–2010 (Fig. 1). Samples were specifically selected for sequencing based on previous genetic assignment to distinct ecotypes/populations using microsatellites, supported by significant $F_{ST}$ values (Islas-Villanueva, 2010; Mirimin *et al.*, 2011; Louis *et al.*, 2014b) and the lack of evidence for recent migration between the populations (Louis *et al.*, 2014b). For consistency, we refer to individuals not assigned to a coastal population as belonging to a ‘pelagic’ ecotype. However, we acknowledge that little is known about the ecology of these individuals, and they may inhabit both neritic and oceanic waters. In addition to the 30 samples successfully sequenced for the study, 57 mitochondrial haplotypes from the genus *Tursiops* from other regions of the world, along with 17 delphinid sequences, were downloaded from GenBank (Appendix S1, Electronic Supplementary Material).

DNA extraction, amplification and sequencing

DNA was extracted from 33 tissue samples using the Qiagen DNeasy (Qiagen DNeasy, Valencia, CA, USA) kit following the manufacturer’s guidelines. DNA yield was quantified using a Qubit (Life Technologies) and ranged between 10 and 300 ng/μL for all samples. The DNA samples were then sheared to fragments of ~150-200 bp using a Diagenode Bioruptor NGS run with 20 cycles of 30 seconds on, and 30 seconds off.

To generate mitochondrial genome sequences we employed a simple shotgun sequencing method (see Tilak *et al.*, 2015). Illumina sequencing libraries were built on the sheared DNA extracts using NEBNext (Ipswich, MA, USA) DNA Sample Prep Master Mix Set 1 following Meyer and Kircher (2010). Libraries were subsequently index amplified for 15 cycles using Phusion High-Fidelity Master Mix (Finnzymes) in 50-μL reactions following the manufacturer guidelines. The libraries were then purified using a MinElute PCR purification kit (Qiagen, Hilden, Germany). Concentrations of amplified libraries were initially checked using a Qubit (Life Technologies) and fragment size distribution was visualized on agarose gel, before a 1/10 diluted aliquot was run on Agilent Bioanalyser 2100 (Palo Alto, CA, USA) to determine molarity and concentration and facilitate equimolar pooling of index amplified libraries. The index amplified libraries were then sequenced in sub-partitions of single
channels on an Illumina HiSeq 2500 ultra-high-throughput sequencing platform using single
read (SR) 100-bp chemistry.

Conversion of Illumina's *.bcl files to fastq, and demultiplexing were performed using
Illumina’s CASAVA (version 1.8.2) software allowing for no mismatch in the 6-nucleotide
indices used for barcoding. Sequencing reads within the generated fastq files were processed
with ADAPTER-REMOVAL (Lindgreen 2012) to trim residual adapter sequence
contamination and to remove adapter dimer sequences as well as low-quality stretches at 3’
ends (i.e. consecutive stretches of N’s and of bases with a quality score of 2 or lower). Reads
that were ≤30 bp following trimming were discarded. Filtered reads with average phred-
scores ≥30 were then mapped to a reference T. truncatus mitogenome sequence
(KF570351.1) using BWA version 0.5.9 (Li & Durbin, 2009), requiring a mapping quality of
Q ≥30 for all bases. The reference mitochondrial sequence was modified as per Morin et al.
(2015) to improve assembly coverage at the ‘ends’ of the linearized mitogenome by adding
40-bp from each end to the opposite end (so that reads could map across the artificial break
point of the linearized sequence). The influence of potential NUMTs was excluded given the
absence of stop codons or frame shifts in the aligned protein encoding genes, and the lack of
regions of excessive coverage due to the mapping of both mt and nuDNA. Clonal reads were
collapsed using the rmdup program of the SAMTOOLS (version 0.1.18) suite (Li et al.,
2009). Consensus mitogenome sequences were then reconstructed using bam files, which
were assembled and visualised in GENEIOUS (Biomatters Ltd.), allowing indels and unique
variants to be visually verified in the BAM files, and only including nucleotide positions with
a read depth of ≥3× coverage; ambiguous bases were assigned as Ns.

Time-calibrated phylogeny of delphinids and the genus Tursiops

There were insufficient data for the reconstruction of complete mitochondrial genomes for
three sequenced samples, which were therefore excluded from further analyses. The
remaining 30 whole mitochondrial T. truncatus genomes were used in the subsequent
analyses. The inference of time-calibrated phylogenies followed a two-step methodology
following Morin et al. (2015) with the purpose of the first step (delphinid phylogeny) to
estimate a calibration point (divergence of the two most evolutionarily distant Tursiops spp.
haplotypes) for the second step (Tursiops spp. phylogeny).

The 30 whole mitochondrial T. truncatus genomes sequenced for this study were first aligned
with 56 sequences from the genus Tursiops and a rough-toothed dolphin (Steno bredanensis)
sequence downloaded from GenBank (see Appendix S1 for sequence accession numbers) that was used as an outgroup using the MUSCLE algorithm in MEGA 6 (Tamura et al., 2007). A topology tree in MRBAYES (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003) was then built with these sequences after initial model selection for the best substitution scheme in JMODELTEST2 (Posada, 2008; Darriba et al., 2012), and the resulting consensus tree was inspected to find the two most divergent Tursiops haplotypes (SABD1, representing a haplotype of T. australis obtained from South Australia, and SA99 representing a haplotype of T. aduncus collected from South Africa (see Moura et al., 2013). The thirteen protein coding genes of the mitochondrial genome from these two most divergent Tursiops sequences were then aligned with the additional 18 delphinid sequences downloaded from GenBank (Appendix S1). We used the package ClockstaR (Duchêne et al., 2014) in R 3.3.3 (R Core Team, 2016) to investigate whether the different mitochondrial genes had different mutation rates or whether a single rate could be applied in further time-calibrated models.

The ‘greedy’ search in PARTITIONFINDER (v1.1.0) (Lanfear et al., 2012) was used to find the best partitioning scheme for each gene, and a time-calibrated phylogenetic tree was built in BEAST2 (Drummond et al., 2012) with three different data partitions for nucleotide substitution models (Appendix S2). We used a single clock model based on the ClockstaR results and estimated a single tree due to the non-recombining properties of mtDNA to find the time to most recent common ancestor (TMRCA) between the two most divergent Tursiops haplotypes. The TMRCA of Delphinidae (mean = 10.08 Myr, SD = 1.413, (McGowen et al., 2009) was used to calibrate the root of the tree and a Calibrated Yule prior was used for the branching rate. Both uncorrelated log-normal relaxed clock (Drummond et al., 2006) and strict clock models were tested with 40,000,000 Markov Chain Monte Carlo (MCMC) steps, 25% pre-burn-in, and a sampling frequency of 4,000. Each model was run twice and the convergence of chains and the Effective Sample Size (ESS) values relating to the model parameters were checked in TRACER (Rambaut et al., 2014). After verifying convergence, LOGCOMBINER and TREEANNOTATOR (Drummond et al., 2012) were used to combine and summarize the trees, respectively. Model selection between the two clock models was done by inspecting the uclStdev parameter in the relaxed clock model; a standard deviation close to zero (<0.1) in this parameter indicates negligible variation in the substitution rates across branches and a better fit of the strict clock model (Drummond & Bouckaert, 2015).
We tested for potential saturation (multiple substitutions) of the third codon positions by plotting the uncorrected pairwise genetic distances and model-corrected (Tamura-Nei substitution model, Tamura & Nei, 1993) genetic distances in R. To investigate the potential effect of incomplete purifying selection, we compared the node TMRCA estimates and clock rates derived using all codon positions (as described above) to a time-calibrated phylogeny created using only the third codon positions, as most mutations that occur in third codon positions are silent and therefore code for the same amino acid (Lagerkvist, 1978). This makes them less likely to be directly targeted by purifying selection (but does not account for purifying selection upon a linked site), thereby maintaining a more constant substitution rate over evolutionary time. The approach of using only the third codon positions has been recently used in a study resolving killer whale (Orcinus orca) mitochondrial lineages (Morin et al., 2015).

The best nucleotide partitioning scheme for the coalescent analysis of the genus Tursiops was determined for the data set, consisting of thirteen protein coding gene regions of the 30 mitochondrial genomes sequenced for this study supplemented by 56 Tursiops haplotypes downloaded from GenBank (Appendix S1, Electronic Supplementary Material), using PARTITIONFINDER (v1.1.0) (Lanfear et al., 2012) (Appendix S2, Electronic Supplementary Material). These samples/haplotypes from T. truncatus, T. aduncus and T. australis had been determined to have originated from coastal or pelagic regions in different parts of the world in previous studies and had been delineated into separate populations or ecotypes (Fig. 1; Appendix S1, Electronic Supplementary Material). This information on population structure was used in this study together with mitochondrial gene regions to create a time-calibrated phylogeny for the genus Tursiops and to estimate the divergence times for the different populations of T. truncatus in the North Atlantic.

As the multi-species coalescent (MSC) model has been found not only to delimit species but to identify genetic structuring between putative populations (Sukumaran & Knowles, 2017), this approach was chosen to construct the time-calibrated phylogeny using StarBEAST2 (Ogilvie et al., 2017), an extension of BEAST2 (Drummond et al., 2012). This method is thought to provide accurate inferences of species tree topologies, divergence times, and substitution rates even in the presence of incomplete lineage sorting, i.e., when individual gene trees do not correspond well with the species tree (Ogilvie et al., 2017). For this reason, information on population structure obtained from previous studies on nuclear microsatellite markers were included in the MSC model. In order to estimate historical effective population
287 sizes \((N_e)\) for the populations sequenced for this study, the MSC model was run with the
288 population model ‘Linear with Constant Root Populations’ assuming that population sizes
289 change over time but keeping the size of the root (the ancestor of all species/populations)
290 constant. The scaled population sizes for pelagic Northeast Atlantic, Coastal North and
291 Coastal South produced by the StarBEAST2 models were adjusted by the average generation
292 time of 21 years estimated for \(T.\ truncatus\) and \(T.\ aduncus\) (Taylor et al., 2007) to convert the
293 estimates to population sizes. As with the delphinid phylogeny, to investigate the possible
294 effect of incomplete purifying selection, StarBEAST2 was run with two different models; a
295 model including all codon positions and a second model including only the third codon
296 positions. Both models were run with an uncorrelated log-normal relaxed clock (Drummond
297 et al., 2006) and a strict clock and with 200,000,000 MCMC steps, 10% burn-in and a
298 sampling frequency of 20,000. Based on the results of ClockstaR analysis (Duchêne et al.,
299 2014) on delphinids, a single clock model was used for all the data partitions. We applied a
300 constraint on the root age obtained in the delphinid phylogeny (split of the two most
301 divergent \(Tursiops\) haplotypes; 3.662Myr (95% highest posterior density interval, HPDI:
302 2.567–4.725) with the all codon model, 3.692 Myr (95% HPDI: 2.567–4.153) with the third
303 codon model) to calibrate the trees. The convergence of chains and model performance was
304 inspected in TRACER (Rambaut et al., 2014) after running each model twice, and
305 LOGCOMBINER and TREEANNOTATOR (Drummond et al., 2012) were used to combine the
306 log- and tree-files and summarize the trees, respectively. The resulting summary trees were
307 plotted using phytools (Revell, 2012) in R.
308 Using DnaSP 5.0 (Librado & Rozas, 2009), we calculated the number of segregating sites \((k)\)
309 and four different indices of historical population demographic changes (bottlenecks or
310 expansions); Tajima’s D (Tajima, 1989), Fu and Li’s D* and F* (Fu & Li, 1993) and Ramos-
311 Onsins & Rozas’ R2 (Ramos-Onsins & Rozas, 2002), for the pelagic, the Coastal North and
312 the Coastal South populations. The examination of deviation from the neutral theory model at
313 equilibrium between mutation and genetic drift by these indices was based on 10,000
314 coalescent simulations. Expectations of these statistics are ~zero in a constant-size
315 population, whereas significant negative or positive values can indicate a sudden expansion
316 or a recent population bottleneck, respectively. In addition, differences in nucleotide and
317 haplotype diversity \((\pi_{nuc} \text{ and } \pi_{hap}, \text{ respectively})\) were compared between pelagic and the
318 Coastal North and Coastal South populations with the “genetic_diversity_diffs” function
319 (Alexander et al., 2016) implemented in R and using 10,000 re-sampling permutations, in
order to test whether the observed differences in diversity between the populations were
greater than expected by chance. A haplotype network of the whole mtDNA sequences
(16,390 bp) and including all the North Atlantic *T. Tursiops* samples was created in popART
(Leigh & Bryant, 2015) using the median-joining network (Bandelt *et al.*, 1999).

### Results

#### Models of suitable habitat

According to the AquaMaps model, the core suitable habitat for common bottlenose dolphins
during the LGM ranged from ~42°N (approximately the latitude of the state of Connecticut)
in the Northwest Atlantic southwards along the coast to lower latitudes (Fig. 2, Appendix
S3). In the Northeast Atlantic, their coastal range was restricted to approximately 40°N,
corresponding to the south of the Iberian Peninsula, and to an area around the Alboran Sea in
the Mediterranean. In addition, suitable habitat was hindcasted to cover an area in the pelagic
North Atlantic around the Mid-Atlantic ridge and around the islands on both sides of the
Atlantic. Note that AquaMaps hindcasted suitable habitat also in the Black Sea, however, this
remained closed and thus inaccessible for the dolphins until after 10 kya (Kerey *et al.*, 2004).

#### DNA sequencing

A total of 292×10⁶ sequencing reads were generated from the samples; >1 million reads per
individual for 28 of the 33 individuals included on the sequencing lane. Following QC
filtering, removal of duplicate reads, mapping and only including nucleotide positions with a
read depth of ≥3× coverage, complete mitochondrial genomes were assembled for 30
individuals with a mean sequence coverage ranging from 10× to >100×. Control region
sequences were compared with those generated for the same individuals by Sanger
sequencing in previously published studies (Islas-Villanueva, 2010; Mirimin *et al.*, 2011;
Louis *et al.*, 2014b), in some cases the same individual had been previously sequenced
independently by all three studies. In all cases our high-throughput sequencing generated data
agreed 100% with the Sanger sequence data. Twenty new haplotypes from the samples
sequenced for this study were identified (Appendix S1), which are deposited in GenBank
(KT601188-KT601207).
Time-calibrated delphinid phylogeny

The approximately linear relationship between the uncorrected and model-corrected genetic distances indicated no evidence of saturation of the third codon position of the protein coding genes (Appendix S4). Low standard deviation (<0.1) of the clock rate parameter (ucldStdev) in both delphinid models, the all codon model and the third codon model, run with an uncorrelated log-normal relaxed clock indicated low variation in substitution rates between lineages; therefore, the models run with a strict clock were considered to be a better fit to the data. Average substitution rate for the strict clock all codon model was estimated as 6.663 × 10^{-3} substitutions/site/Myr (95% HPDI: 4.872 × 10^{-3} – 8.865 × 10^{-3}), and for the third codon model as 0.017 substitutions/site/Myr (95% HPDI: 0.012 – 0.022). Both models produced similar topology and node age estimates for the clade including the *Tursiops* and *Stenella* species, *Sousa chinensis*, *Steno bredanensis* and *Delphinus capensis* (Appendix S5). However, *Lagenorhynchus albirostris* and *Cephalorhynchus heavisidii* were placed in different clades in the trees depending on the model used and the models also estimated the divergence of killer whales (*Orcinus orca*) differently. The placement of these species has been uncertain in previous studies (LeDuc *et al*., 1999; Caballero *et al*., 2008; McGowen *et al*., 2009; Steeman *et al*., 2009; Vilstrup *et al*., 2011) and is thought to reflect rapid radiation in the Delphinidae family (Vilstrup *et al*., 2011). We compared the time to most recent common ancestor (TMRCA) estimates for the common nodes derived using the two models and found them to be significantly correlated (Pearson’s $r^2 = 0.99$, $P < 0.001$, Appendix S6) with a slope of 1.01, implying that incomplete purifying selection was not an issue in the delphinid phylogeny. The TMRCA of the two most divergent *Tursiops* samples was estimated as 3.662 Myr (95% HPDI: 2.567 – 4.725) with the all codon model and as 3.692 Myr (95% HPDI: 2.567 – 4.772) with the third codon only model. The Effective Sample Size (ESS) values for the different parameters were all >500 in the third codon only model (most of them >2000), whereas in the all codon model they were >200 (most of them >2000), except for one parameter for the gamma shape and one parameter for the proportion of invariant sites, which were both >150 (Appendix S7).

Time-calibrated *Tursiops* phylogeny

Very low standard deviation (<0.01 in both all codon and third codon only models) in the substitution rates estimated with the uncorrelated log-normal relaxed clock indicated negligible variation in substitution rates between lineages; therefore, in StarBEAST2, a strict
380 clock was used to estimate the time-calibrated phylogeny of the genus *Tursiops*. The ESS
381 values for the parameters in both models were all >300 (most of them >6000), indicating no
382 sign of autocorrelation between samples and a good convergence of chains (Appendix S8,
383 Electronic Supplementary Material). Both models resulted in identical species/population
384 topologies and similar node age estimates and posterior probabilities (Appendix S9,
385 Electronic Supplementary Material). Similar to the Delphinid phylogeny, the node ages were
386 significantly correlated with a slope of 1.01 (Pearson’s $r^2 > 0.99$, $P < 0.001$, Appendix S10),
387 indicating no evidence of incomplete purifying selection affecting the *Tursiops* tree.
388
389 The summary consensus species tree, inferred using the third codon only model and showing
390 only the NEA and Mediterranean and Black Sea *T. truncatus* populations (Fig. 3), indicates
391 that the branch including haplotypes from the Mediterranean and Black Seas split from the
392 NEA branch ~35,900 years ago (95% HPDI: 11,200–60,500). The coastal NEA populations,
393 on the other hand, diverged from the pelagic population around 14,800 years ago (95%
394 HPDI: 4,900–26,400) with a further more recent divergence of the northernmost coastal
395 population (Coastal North) occurring ~7,200 years ago (95% HPDI: 0–14,200). However, the
396 node posterior probabilities of the pelagic and coastal NEA populations are quite low (Fig. 3),
397 likely due to incomplete lineage sorting. Indeed, incomplete lineage sorting, especially of the
398 coastal North Atlantic *T. truncatus* sequences, is evident in the consensus gene tree consisting
399 of all samples from the genus *Tursiops* (green tips in the tree, Appendix S12, Electronic
400 Supplementary Material), and is accompanied by older node coalescence times compared to
401 the time-calibrated species tree (Fig. 3). The consensus gene tree also indicates that the
402 coastal bottlenose dolphin population currently occupying the northern parts of the NEA (i.e.
403 Coastal North) originates from at least two separate mitochondrial lineages.
404
405 The summary consensus species tree depicting the effective population sizes ($N_e$) through
406 Late Pleistocene–Holocene (Fig. 3, Appendix S11) indicates that, unlike the pelagic
407 population, which is estimated to have retained a relatively stable effective size throughout
408 the Late Pleistocene and Holocene, all of the present-day coastal populations, and especially
409 the Black Sea and the Coastal North populations, are inferred to have undergone a genetic
410 bottleneck around 7,000–8,000 years ago followed by expansion towards the present day
411 (Fig. 3, Appendix S11). The significantly lower nucleotide ($\pi_{nuc}$) and haplotype diversities
412 ($\pi_{hap}$) in the Coastal North population compared to the pelagic population, and the
413 significantly lower $\pi_{nuc}$ in the Coastal South population compared to the Pelagic East Atlantic
414 (all $P < 0.05$, based on 10,000 permutations) are further indications of ancestral population
bottlenecks in the coastal populations, as are the reduced number of segregating sites ($k$) (see Table 1). Conversely, the non-significant values of Tajima’s $D$, Fu and Li’s $D^*$ and $F^*$ and Ramos-Onsins & Rozas’ R2, albeit the first three indicators being negative for the Coastal South population (Table 1), fail to provide support for our inference of post-glacial demographic expansion from a small founder population and indicate that there is no significant excess of low frequency variants in the Coastal North and Coastal South populations. This could be an artefact of lack of power to detect significance, or it could be due to a violation of the assumption of panmixia inherent in these statistics. For example, the Coastal North population consists of demes that are interconnected by limited and varying gene flow (Louis et al., 2014b; Nykänen et al., 2018). This could increase the number of rare variants, but it could also inflate intermediate variants; for example, variants that are at high frequency in one or two demes would be identified as intermediate frequency variants in our sample set. The possible unaccounted population structuring, in turn, would inflate the effective population sizes, which could help explain the unrealistically high recent time $N_e$ values estimated for the coastal populations. Furthermore, the fact that the divergence of the coastal populations from the pelagic population was estimated to have occurred already ~15,000 years ago, could mean that the rare variants expected to arise in a population after a bottleneck or founder event have had sufficient time to drift to extinction or intermediate frequency.

The haplotype network shows the two source lineages of the haplotypes shared among different NEA coastal populations (Fig. 4). One of the haplotypes is shared between six individuals from two different populations, five from the Coastal North and one from the Coastal South, and another haplotype is private to the Coastal North population and is shared by five individuals. Thus, 71% of the Coastal North individuals sampled in this study shared just these two haplotypes; the remainder have haplotypes recently derived from these ancestral lineages. This is consistent with the demographic expansion of the Coastal North population from just two founding lineages and a probable ancestral origin in the pelagic population. Similarly, the Coastal South population may have originated from only one or two pelagic founding lineages; however, sequencing more samples from this population will be required to determine this with certainty. Nevertheless, it is likely that the pelagic population gave rise to coastal lineages with current populations located in the NEA as well as in the Mediterranean and Black seas.
Discussion

Our analyses indicate that bottlenose dolphins have colonized the northernmost range of their distribution through a leading-edge expansion, in which the early colonists have expanded to fill the available coastal niches, to the exclusion of others. Due to its strict maternal inheritance the mitochondrial genome tracks only the matrilineal population history, a very stochastic coalescent process, and generally has less power to detect ancestral demographic changes than estimates based on analyses of multiple nuclear markers (Ho & Shapiro, 2011). For this reason, a multi-species coalescent (MSC) model that utilized information on population structure from previous studies on biparentally inherited nuclear DNA markers (microsatellites) was used to infer divergence times of distinct species and populations of bottlenose dolphins. Incomplete lineage sorting and rapid radiation among delphinids (e.g. McGowen et al., 2009; Amaral et al., 2012; Moura et al., 2013; Louis et al., 2014b; Morin et al., 2015; Foote & Morin, 2016) can complicate the estimation of the timing of speciation events and population divergence in the genus Tursiops over phylogenetic time-scales (Moura et al., 2013). One advantage of using MSC model, such as StarBEAST2 (Ogilvie et al., 2017), compared to traditional methods which assume that evolutionary histories of gene trees match the evolutionary history of the species tree, is that the branch lengths and divergence times are not overestimated (Ogilvie et al., 2017). Another advantage of StarBEAST2 is the possibility to infer historical effective population sizes while accounting for population structuring, as this structuring combined with uneven sampling can lead to bias in Ne estimates inferred with traditionally used Bayesian Skyline methods (Heller et al., 2013) that assume panmixia. Additionally, time-dependency of increasing mutation rate towards the present caused by incomplete purifying selection – the time lag between the appearance of a mutation and its removal due to purifying selection - has the potential to further bias divergence time estimates (e.g. Ho et al., 2005; Ho et al., 2011). The use of sites under strong selective constraint extends terminal branches potentially leading to overestimation of older node ages (Ho et al., 2011; He et al., 2019). In this study, based on the comparison of time-calibrated phylogenies inferred using all codon positions and only the third codons, we found no evidence of incomplete purifying selection affecting the divergence time estimates in the delphinid family or in the phylogeny consisting of Tursiops spp. sequences and haplotypes. However, this may have been an artefact of the small number of common nodes (N=8) included in the comparison or the relatively young node ages of the branches consisting of T. truncatus samples, and the issue of time-dependency in Tursiops
phylogenies should be re-examined with more samples collected from various populations. Nevertheless, third codon positions are thought to be less constrained by purifying selection compared to first and second codon positions, and they are therefore expected to evolve in a more clock-like fashion, accumulating mutations at a more constant rate and thus minimizing the effect of time-dependence (Ho et al., 2011).

The average substitution rate for the third codon positions across all branches for the samples from the genus *Tursiops* was estimated as 0.017 substitutions/site/Myr. This estimate corresponds well with previously estimated rates of 0.024 for third codon positions for cetaceans (Ho & Lanfear, 2010) and 0.021 substitutions/site/Myr for bottlenose dolphin mitogenome-wide rates (Moura et al., 2013). Similarly, the estimated divergence date for the second youngest node for the genus *Tursiops* in this study, represented by the divergence between Mediterranean and Black Sea bottlenose dolphins (around 7,900 years ago), is concordant with that derived previously by Moura et al. (2013), and is consistent with the geological timeframe of the opening of the Bosphorus Strait as well as the climatic context linking to the LGM. In contrast, the average root age for the genus *Tursiops* in this study was estimated as 3.7 Myr, which is considerably older than 1.1 Myr estimated by Moura et al. (2013). Possible reasons for this incongruence may be a result of the different modelling strategies applied to different parts of the mitogenome (mtDNA protein coding regions were used in this study whilst Moura et al. (2013) used the whole mtDNA sequence). However, a more likely reason is the fact that these authors used a relatively narrow prior for their biogeographical calibration (Uniform prior of 0.003–0.01, supplementary table S5 in Moura et al. (2013) that overpowered the fossil calibrations with much wider priors, and this ultimately resulted in poor convergence of the older node age posterior estimates based on low (<100) ESS values (supplementary table S6 in Moura et al. (2013)).

Pleistocene climatic oscillations are thought to have played a major role in shaping species distribution and divergence and in promoting speciation (Avise et al., 1998). The divergence of the Northwest Atlantic coastal population from other Atlantic *T. truncatus* populations around 600,000 years ago coincides with the end of a cooler period during the mid-
Pleistocene when the SST started to gradually rise again up to ~10°C in the Northern Hemisphere (Clark et al., 2006). Consistent with previous studies (Natoli et al., 2005; Moura et al., 2013; Louis et al., 2014a), the phylogenetic tree estimated in this study for the genus *Tursiops* (Fig. 3) suggests that the colonization of coastal habitats in the NEA occurred from
a pelagic source population. According to our results, this happened via an ancestor that occupied lower latitude Atlantic waters. The AquaMaps model also indicated that bottlenose dolphins may have inhabited parts of the Mediterranean Sea during the LGM, as some areas were highlighted as suitable habitat (Fig. 2b). Even if the northern parts of the Mediterranean Sea (e.g. Tyrrhenian Sea) were too cold or otherwise unsuitable for the bottlenose dolphins during the LGM, a refugial population may have existed in the areas around the warmer Alboran Sea, with models of SST showing consistently warmer annual temperatures above 10°C for this area throughout the last ice age (Cacho et al., 2001). Similar refugia at lower latitudes have been documented, for example, with Atlantic salmon, *Salmo salar* (Consuegra et al., 2002; Finnegan et al., 2013) and European eel, *Anguilla anguilla* (Kettle et al., 2011). Suitable habitat for bottlenose dolphins during the LGM was also hindcasted with AquaMaps in the Black Sea. However, the land bridge between the Black Sea and the Mediterranean collapsed only ~10,000 years ago (Kerey et al., 2004). Moreover, an accumulation of species found in the Black Sea sediment that are indicative of marine, rather than brackish, conditions was dated to 7,000–6,400 yr (the Kalamitian transgression (Yanko-Hombach et al., 2007)). This makes the Black Sea an unlikely habitat for bottlenose dolphins until ~7,000 years ago at the earliest, and the divergence time between the eastern Mediterranean and Black Sea populations estimated as ca. 7,900 ya in this study fits quite well within this timeframe.

Temperature changes after the LGM have likely also played a role in the colonization of the northern parts of the NEA. The post-LGM divergence of the coastal populations in the NEA from the pelagic ~15,000 years ago coincides with the rising temperatures in the lower latitudes of the North Atlantic Ocean. The subsequent recent split of the Coastal North population from its southern counterpart currently occupying lower latitude European waters was estimated to have occurred around 8,000 years ago, and the timing of the split is supported by two *T. truncatus* subfossils found in the Dutch Southern Bight and radio-carbon aged to 7,390 and 8,135 yrs, respectively (Post, 2005). It thus appears that these cladogenetic events are correlated with periods of temperature changes, with warmer temperatures leading to an increase in sea-level in coastal areas and the subsequent release of available habitat. Further, it is also likely that suitable coastal habitat became available gradually in the NEA coastal areas after the LGM and the onset of the warmer Holocene leading to sequential colonization and divergence.
According to Hewitt (1999, 2000), the colonization of suitable niches in high latitudes would occur rapidly via long-range dispersal events. Being highly mobile, cetaceans are capable of undertaking long distance movements on relatively short time-scales (Forcada, 2009; O’Brien et al., 2009; Robinson et al., 2012). Coupled with this potential for long-range dispersal, the fact that bottlenose dolphins form small local resident populations in coastal European waters (e.g. Ingram & Rogan, 2002; Cheney et al., 2013; Louis et al., 2015), makes post-glacial leading-edge colonization of the northern NEA a likely dispersal scenario. The demographic contraction and subsequent expansion observed in the coastal populations of this study, as well as the significantly reduced nucleotide and haplotype diversities, suggest that bottlenose dolphins indeed expanded their range northwards to coastal areas via founder events during leading-edge expansion. A similar founder event has recently been suggested to give rise to the present-day resident-type killer whale communities in the Russian Pacific (Filatova et al., 2018), and rapid demographic expansions after the LGM have been documented across other marine taxa (e.g. Klimova et al., 2014; Jenkins et al., 2018).

Our results indicate that the colonization and expansion to fill the available coastal habitat around the British Isles happened via serial founder events after the LGM by a small leading-edge founder group, that comprised as few as two maternal lineages, from the source pelagic population. This finding is supported by previous publications of larger datasets of 136 mitochondrial control region sequences from coastal populations around the British Isles, which comprise the same two mitochondrial haplogroups (Mirimin et al., 2011; Louis et al., 2014b). The inferred process of post-glacial expansion into emerging habitat reflects the model inferred by Hoelzel et al., (2007) for colonization of the coastal regions of the North Pacific and the formation of structure in killer whales. This process may thus reflect a commonality in the demographic history of high latitude social odontocetes. More broadly, serial founder events by a small subset of the parent population, leading to sequential loss of genetic diversity, have been linked to the phylogeographic history in a number of species, including modern humans (Homo sapiens) during the migration out of Africa (DeGiorgio et al., 2009; Deshpande et al., 2009; Henn et al., 2012). Still today the greatest nuclear and haplotype diversity among modern humans is found in Africa (reviewed in Campbell & Tishkoff, 2008), and it is generally accepted that Africa was the source of all current modern human populations. Similarly, we identify a loss of genetic diversity in bottlenose dolphins at the northern edge of their coastal range relative to that found in the large pelagic population (e.g. Quérouil et al., 2007; Mirimin et al., 2011; Louis et al., 2014b).
In summary, our results suggest that the present day geographic and genetic structuring of bottlenose dolphins in the North Atlantic were shaped by the climatic cycles of the Late Quaternary, and the most northerly coastal population represents a leading-edge expansion by two founding mitochondrial lineages. We estimate these lineages originated from a diverse pelagic source population, and then rapidly spread throughout and retained the available coastal territory and associated resources around the British Isles. However, the responses to climate change are likely to vary even among mobile species depending on their tolerance to environmental conditions and the derived habitat preference (Foote et al., 2013; Sydeman et al., 2015). Although it seems likely that the amount of available suitable habitat for bottlenose dolphins will increase and expand northwards in the next hundred years with the ongoing warming of the oceans (Appendix S13, Electronic Supplementary Material), the complexity of ecosystems that bottlenose dolphins live in, combined with a range of social behaviors that they exhibit, makes predicting the consequences of the climate change to populations an extremely difficult task.

Funding

This work was supported by Thomas Crawford Hayes Scholarship (School of Biological, Earth and Environmental Sciences, University College Cork) to MN; National Parks and Wildlife Service, The Department of Culture, Heritage and the Gaeltacht, Ireland, to ER and SI; Fondation Fyssen to ML; Fondation Total; Agence de l’Eau Seine-Normandie; Fonds de Dotation pour la Biodiversité; Agence des Aires Marines Protégées; Direction Régionale de l’Environnement, de l’Aménagement et du Logement; Ministère de l’Ecologie, du Développement Durable et de l’Energie; Conseil Général de la Manche; and Communauté d’Agglomération de la Ville de La Rochelle.

Acknowledgements

The authors wish to thank all the people who were involved in biopsy sample collection, including Anneli Englund and Ailbhe Kavanagh (University College Cork), François Gally and volunteers of the Groupe d’Etudes des Cétacés du Cotentin. Our warmest gratitude also to everyone who helped in tissue sample collections; Harry Ross, Tony Patterson, Bob Reid and Mary Harman, as well as Amelia Viricel (University of La Rochelle), and volunteers from the French Stranding Network (Réseau National Echouages), Cécile Dars, Fabien Demaret, Ghislain Doremus and Olivier Van Canneyt (Observatoire Pelagis). Some of the samples
from stranded animals were obtained from the Irish Cetacean Genetic Tissue Bank, Natural History Museum, Dublin. UK stranded samples were provided by the UK Cetacean Strandings Investigation Programme, which is jointly funded by DEFRA and the Devolved Governments of Scotland and Wales. We acknowledge The Danish National High-Throughput DNA Sequencing Centre for sequencing the samples and, particularly, Andaine Seguin-Orlando, Lillian Petersen, Cecilie Demring Mortensen, Kim Magnussen and Ian Lissimore for technical support. We also thank Michael Krützen, Nick Matzke, Tom Reed and Vincent Ridoux for their input on the earlier version of the manuscript, and to two anonymous reviewers for their helpful comments. Biopsy samples in Ireland were collected under licenses from the National Parks and Wildlife Service (license numbers C111/2013 and C043/2014) and the Health Products Regulatory Authority (license number: AE19130/P014). Biopsies in France were collected under permit 9/115/DEROG from the French ministry.

618 Supporting information

619 Supporting information (Appendices S1-S13) have been uploaded as part of the Electronic Supplementary Material.

621 Data availability

622 The datasets supporting this article have been uploaded in dryad at doi:10.5061/dryad.XXXX.

623 References


concordant assessment from archaeozoology and palaeotemperature reconstructions. 


Tucker, V.A. 1975. The energetic cost of moving about: walking and running are extremely inefficient forms of locomotion. Much greater efficiency is achieved by birds, fish—and bicyclists. *Am. Sci.* **63**: 413–419.


### Table and figure captions

**Table 1.** Summary statistics of genetic diversity and the site frequency spectra used to infer recent population expansion/contraction in Northeast Atlantic bottlenose dolphin populations. Only the sequences generated for this study (N=30) have been included in the analyses.

**Figure 1.** (a) *Tursiops* spp. haplotypes retrieved from GenBank and the approximate sampling location of the species/populations they represent; SABD – *T. australis* (South Australia, N=7), SA – *T. aduncus* (South Africa, N=10), IPTA – *T. aduncus* (Southeast Australia, N=10), WNAC – *T. truncatus* (coastal Northwest Atlantic, N=9), EMED – *T. truncatus* (Eastern Mediterranean Sea, N=10), BSEA – *T. truncatus ponticus* (Black Sea, N=10). (b) the biopsy/stranding locations of the 30 new *T. truncatus* samples sequenced for this study with the populations represented by the different symbols; PEA – *T. truncatus*, pelagic (Northeast Atlantic, N=11), CS – *T. truncatus*, Coastal South (Northeast Atlantic, N=5), CN – *T. truncatus*, Coastal North (Northeast Atlantic, N=14). Note that some of the sample locations overlap in map (b). The area in map (b) is marked with the rectangle in map (a). See online version for full color.

**Figure 2.** AquaMaps suitable habitat map for *T. truncatus* for (a) the present day, and (b) the Last Glacial Maximum in the North Atlantic. White to red colors represent least to most suitable habitat, respectively, and land is shown in light gray. Dark green color in Fig. (b) represents areas with permanent ice sheet and light green color areas with >50% sea ice concentration. The predictions for the entire world are given in Appendix S3.
Figure 3. Time-calibrated phylogenetic tree depicting the effective population size ($N_e$) over time, shown only for the Northeast Atlantic *T. truncatus* populations and estimated with StarBEAST2 model ‘Linear with Constant Root Populations’, using only the third codon positions from 13 mitochondrial genes. The numbers in rectangles are node posterior probabilities, numbers below nodes median node ages, and the bars represent 95% HPDI in node TMRCA. The width of the line depicts $N_e$ with the values given in Appendix S11. The tips represent current populations/species and their sampling locations, see Fig. 1. The full *Tursiops* species trees are shown in Appendix S9, Electronic Supplementary Material.

Figure 4. Median-joining haplotype network of North Atlantic (including Mediterranean and Black Seas) *T. truncatus* based on 16,390 bp mtDNA sequence constructed in PopART. Haplotypes (circles) have been assigned a color based on population assignment in previous studies and their sampling locations (see Fig. 1). The size of the circle represents the frequency of occurrence of each haplotype and the number inside it the number of individuals. Mutational steps between haplotypes are indicated by hatch marks and small black circles are missing haplotypes. See online version for full color.
## Tables and figures

**Table 1.** Summary statistics of genetic diversity and the site frequency spectra used to infer recent population expansion/contraction in Northeast Atlantic bottlenose dolphin populations. Only the sequences generated for this study (N=30) have been included in the analyses.

<table>
<thead>
<tr>
<th>Population</th>
<th>k</th>
<th>$\pi_{nuc}$</th>
<th>$\pi_{hap}$</th>
<th>Tajima's $D$</th>
<th>Fu and Li's $D^*$</th>
<th>Fu and Li's $F^*$</th>
<th>Ramos-Onsins &amp; Rozas' $R_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pelagic (PEA, n=11)</td>
<td>247</td>
<td>0.0044†</td>
<td>1.000††</td>
<td>-0.722</td>
<td>-0.325</td>
<td>-0.552</td>
<td>0.111</td>
</tr>
<tr>
<td>Coastal North (CN, n=14)</td>
<td>66</td>
<td>0.0021</td>
<td>0.780</td>
<td>2.641</td>
<td>1.461</td>
<td>2.253</td>
<td>0.245</td>
</tr>
<tr>
<td>Coastal South (CS, n=5)</td>
<td>64</td>
<td>0.0016</td>
<td>0.900</td>
<td>-1.119</td>
<td>-0.321</td>
<td>-0.613</td>
<td>0.330</td>
</tr>
</tbody>
</table>

$k$ = number of segregating sites, i.e. polymorphisms

$\pi_{nuc}/\pi_{hap} = $ nucleotide/haplotype diversity

* an outgroup (*Steno bredanensis*) was used in the calculations of the indices

† = significant difference at $\alpha=0.05$ level was detected between the pelagic and both of the coastal populations

‡ = significant difference at $\alpha=0.05$ level was detected between the pelagic and the Coastal North population
Figure 2
Figure 3
Figure 4