Meta-analysis of northeast Atlantic marine taxa shows contrasting phylogeographic patterns following post-LGM expansions

Tom L. Jenkins¹, Rita Castilho² and Jamie R. Stevens¹

¹ Department of Biosciences, University of Exeter, Exeter, United Kingdom
² Center for Marine Sciences, Campus de Gambelas, University of Algarve, Faro, Portugal

ABSTRACT

Background. Comparative phylogeography enables the study of historical and evolutionary processes that have contributed to shaping patterns of contemporary genetic diversity across co-distributed species. In this study, we explored genetic structure and historical demography in a range of coastal marine species across the northeast Atlantic to assess whether there are commonalities in phylogeographic patterns across taxa and to evaluate whether the timings of population expansions were linked to the Last Glacial Maximum (LGM).

Methods. A literature search was conducted using Web of Science. Search terms were chosen to maximise the inclusion of articles reporting on population structure and phylogeography from the northeast Atlantic; titles and abstracts were screened to identify suitable articles within the scope of this study. Given the proven utility of mtDNA in comparative phylogeography and the availability of these data in the public domain, a meta-analysis was conducted using published mtDNA gene sequences. A standardised methodology was implemented to ensure that the genealogy and demographic history of all mtDNA datasets were reanalysed in a consistent and directly comparable manner.

Results. Mitochondrial DNA datasets were built for 21 species. The meta-analysis revealed significant population differentiation in 16 species and four main types of haplotype network were found, with haplotypes in some species unique to specific geographical locations. A signal of rapid expansion was detected in 16 species, whereas five species showed evidence of a stable population size. Corrected mutation rates indicated that the majority of expansions were estimated to have occurred after the earliest estimate for the LGM (~26.5 Kyr), while few expansions were estimated to have pre-dated the LGM.

Conclusion. This study suggests that post-LGM expansion appeared to be common in a range of marine taxa, supporting the concept of rapid expansions after the LGM as the ice sheets started to retreat. However, despite the commonality of expansion patterns in many of these taxa, phylogeographic patterns appear to differ in the species included in this study. This suggests that species-specific evolutionary processes, as well as historical events, have likely influenced the distribution of genetic diversity of marine taxa in the northeast Atlantic.
INTRODUCTION

Comparative phylogeographic studies present opportunities to explore how historical events may have helped shape patterns of genetic structure amongst co-distributed species (Avise et al., 1987; Avise, 2009; Hickerson et al., 2010). Patterns of concordant phylogeographical structure across multiple taxa are particularly informative because, while some patterns of spatial genetic structure may be caused by species-specific evolutionary processes, patterns common across multiple taxa may suggest similar evolutionary histories, such as common barriers to gene flow (Avise, 2009; Hickerson et al., 2010). These findings can be important for conservation because of the potential to modify management actions in the light of the differing phylogeography of multiple species across the same geographical area (Pelc, Warner & Gaines, 2009; Toonen et al., 2011; Heyden et al., 2014; Liggins et al., 2016). In marine biology, such comparative studies have made important contributions to our understanding of how historical events, such as the Pleistocene glaciations, have helped shape the spatial patterns of contemporary genetic diversity of marine taxa (Patarnello, Volckaert & Castilho, 2007; Maggs et al., 2008; Marko et al., 2010; Ni et al., 2014).

The Pleistocene epoch was characterised by recurrent glaciations and intensive fluctuations in climate that periodically influenced the spatial distributions of plants and animals (Hewitt, 1999; Hofreiter & Stewart, 2009). The most recent glacial period began approximately 115 Ka and nearly all ice sheets were at their maximum (Last Glacial Maximum, LGM) between 26.5–19 Ka (Clark et al., 2009). The advances of the Northern Hemisphere ice sheets led to significant changes in temperature and sea levels (Lambeck & Chappell, 2001). This must have had profound implications for habitat availability and the population persistence of coastal species—large parts of species’ ranges would have been reduced, while other species may have survived in glacial refugia (Maggs et al., 2008; Provan & Bennett, 2008). As the ice retreated and the sea level rose, a number of individuals from refugial populations may have dispersed and recolonised areas unavailable during the glaciation (Hewitt, 2000). Changes in latitudinal ranges and population sizes can have distinct effects on the genetic architecture of a species due to the competing processes of mutation, drift and selection; moreover, the deep molecular divergence reported in taxa associated with several known European refugia suggests repeated expansion and contraction of conspecific populations were common throughout the Pleistocene (Hewitt, 2004).

In the northeast Atlantic, the ice sheets extended as far south as Britain and Ireland, leaving an ice-free zone in mid-southern England, with possibly a small area in southwest Ireland free of ice (Chiverrell & Thomas, 2010). However, the predicted extent of ice coverage across southern Ireland and the Celtic Sea differs among studies (e.g., Taberlet et al., 1998; Hughes et al., 2016). The advance of the ice sheets led to a drastic drop in sea levels in the English Channel, resulting in the complete emersion of the channel between
England and France, except for a palaeo-river that extended across the continental margin (Ménot et al., 2006). This suggests that extant coastal communities inhabiting these areas are likely recolonisers originating from glacial refugia. It has been suggested that Hurd Deep, a trench in the English Channel (Fig. 1), might have persisted as a marine lake during the LGM, thereby acting as a potential glacial refugium (Provan, Wattier & Maggs, 2005; Hoarau et al., 2007). Other areas further south, including Brittany (Coyer et al., 2003) and the Iberian Peninsula (Hoarau et al., 2007; Neiva et al., 2012) (Fig. 1), have also been postulated to act as refugia during the LGM. This was supported by high levels of genetic diversity found at these areas in the species studied, a key signature indicative of glacial refugia (Provan & Bennett, 2008).

Studies of single-species phylogeography across the northeast Atlantic are common; yet, because of the differences in molecular methodologies and analytical approaches, it can be difficult to compare results reliably. By applying a consistent methodology across all studies, this standardises the analysis (Harrison, 2011), enabling patterns of phylogeography to be explored and compared within and across taxa. Two comparative meta-analyses in the Atlantic Ocean have been published to-date: the first explored the feasibility of distinguishing genetic signatures of periglacial refugia from southern refugia
in eight benthic marine species (Maggs et al., 2008), and the second looked for concordance among phylogeographical breaks around the southeast coast of the United States of America (Pelc, Warner & Gaines, 2009). Systematic meta-analyses across diverse taxa in other seas and oceans have proved useful for exploring broad patterns of phylogeography (e.g., Patarnello, Volckaert & Castilho, 2007; Kelly & Palumbi, 2010; Marko et al., 2010; Ni et al., 2014); for example, one study of rocky-shore taxa from the northeastern Pacific found that 36% of species showed evidence of population expansions associated with the LGM, while 50% exhibited demographic patterns consistent with stable effective population sizes (Marko et al., 2010). However, such a study for marine taxa across the northeast Atlantic has yet to be undertaken.

In this study, we reanalyse available mitochondrial (mt)DNA data to compare the phylogeography of coastal benthic and demersal organisms across the northeast Atlantic (Fig. 1), an area characterised by complex oceanography and historical biogeographical events, such as the Pleistocene glaciations. Specifically, our aims were: (i) to identify commonalities (or otherwise) in contemporary genetic structure; (ii) to re-examine historical demography to test for signatures of population expansions; and (iii) to estimate the timings of any expansions detected. We discuss our findings in the context of the Pleistocene glaciations, asking in particular whether the LGM affected the phylogeography of marine taxa concordantly or discordantly.

MATERIAL AND METHODS

Literature search
To compare the phylogeography of benthic and demersal organisms across the northeast Atlantic, we undertook a meta-analysis of molecular phylogeographic studies. A literature search was conducted by TLJ and JRS using Web of Science (Thomson Reuters) in February 2015. Search terms were chosen to maximise the inclusion of articles reporting on population structure and phylogeography from the northeast Atlantic. The following sets of Boolean search terms were submitted to the Advanced Search Tool: (1) gene flow OR population structure OR genetic diversity OR phylogeograph*; (2) marine OR intertidal OR subtidal OR estuar*; and (3) Atlantic. Titles and abstracts were screened by TLJ and JRS to identify suitable articles within the scope of this study and only articles that matched the following criteria were retained: (a) organisms were fully marine or estuarine throughout their life history (diadromous species were excluded); (b) studies of temporal changes, hybridisation or introgression from closely related species were omitted; (c) the study included at least three sampling sites from within the northeast Atlantic (Fig. 1—sites outside of this area were not considered); (d) datasets contained a minimum of five individuals per site and a total sample size of at least 50; and (e) the study included latitude and longitude of the sampling sites or a detailed description or map which provided sufficient detail to determine the geographical location of sample origins. The studies were reviewed independently by TLJ and JRS and there were no disputes regarding inclusion or rejection that needed adjudication. Given the proven utility of mtDNA in comparative phylogeography (e.g., Patarnello, Volckaert & Castilho, 2007; Ni et al., 2014)
and the availability of these data in the public domain, a meta-analysis was conducted using published mtDNA gene sequences.

**Data reanalysis**

A standardised methodology was implemented to ensure that all mtDNA datasets were reanalysed in a consistent and directly comparable manner. Data analyses in the original studies were far from consistent, particularly with respect to the analysis of haplotype networks and historical demography. The majority of studies reported information about population structure, however, in several instances the studies included additional samples outside of the northeast Atlantic in their analysis. Therefore, standardised tests of population structure were undertaken *de novo* for each species. Sites that were genetically homogeneous (as described by the original authors) and which were spatially close or situated in the same geographical region were combined in some datasets. This ensured that phylogeography within and across seas was examined in this meta-analysis. Population differentiation was examined using global values of Jost’s $D$ \cite{Jost2008} and $F_{ST}$ \cite{WeirCockerham1984} using the `fastDivPart` function from the R package `diveRsity` \cite{KeenanR Core Team2013} and significance was assessed using 10,000 permutation replicates.

To examine the genealogical relationships within species, haplotype networks were constructed using the `haploNet` function from the R package `pegas` \cite{Paradis2010}. Tajima’s $D$ \cite{Tajima1989}, Fu’s $F_S$ \cite{Fu1997} and Ramos-Onsins’ $R_2$ \cite{RamosOnsinsRozas2002} neutrality tests were performed in DnaSP v5.10 \cite{LibradoRozas2009} to determine whether each species carried a signal that deviated from neutrality (significance was assessed using 10,000 bootstrap replicates). Mismatch analyses (frequency of pairwise nucleotide-site differences between sequences) were carried out using the population growth-decline model in DnaSP to further examine the demographic history, and Harpending’s raggedness index ($r$) \cite{Harpending1994} was used to evaluate the fit of the observed distribution to the growth-decline model (10,000 bootstrap replicates). A non-significant index suggests that the observed data have a relatively good fit to the growth-decline model. In contrast, a significant index is indicative of a stable population which is typically thought to show a ‘ragged’, multi-modal mismatch \cite{Harpending1994}.

The equation $t = \tau/(2\mu k)$ was used to estimate the timing of a population expansion ($t$), where $\tau$ is the date of the expansion measured in units of mutational time (Tau –estimated using DnaSP), $\mu$ is the mutation rate per site per year and $k$ is the sequence length. In addition, Bayesian Skyline Plots (BSPs) were run using BEAST2 v2.5.0 \cite{Drummond2005,Bouckaert2014}. BEAST2 uses a Markov chain Monte Carlo (MCMC) sampling procedure to estimate effective population size ($N_e$) through time based on the temporal distribution of coalescences in gene genealogies. For each dataset, the substitution model was selected using bModelTest \cite{BaridoSottani2018}, which uses reversible jump MCMC that allows the Markov chain to jump between states representing different possible substitution models. A strict clock and a coalescent Bayesian Skyline prior was implemented. Each run consisted of 100 million steps with a burn-in of one million and
parameters were sampled every 10,000 steps. Chain convergence and BSPs were analysed with Tracer v1.7.1 (Rambaut et al., 2018).

Recent studies have shown that the use of mutation rates derived from ancient calibration dates or from phylogenetic analyses may not be appropriate for studies at the population level (Ho et al., 2008; Ho et al., 2011). In this study, therefore, mutation rates were chosen based on the most recent calibration date available for the closest taxonomic relative (Table S1). In published studies where a mutation rate was not specified, the genetic distance provided by the study was divided by the date of the calibration event (in Myr) to obtain a % mutation rate per Myr. For cases where only calibration dates older than 5 Myr were available for the species and gene of interest, a three-fold correction in mutation rate was applied to the original rate to control for the potential time-dependency of molecular rates. This adjustment was implemented because rates have been found to vary by three to six-fold for several marine species when calibration dates younger than 5 Myr vs. older dates have been tested (Crandall et al., 2012; Laakkonen, Strelkov & Väinölä, 2015). A range of mutation rates based on the rates reported by previous studies were used to calculate a minimum, maximum and average time estimate since a population expansion.

RESULTS
Literature search
The initial search using Boolean terms identified 1,120 articles, which was reduced to 56 articles after the titles and abstracts were examined and the search criteria were applied (Fig. S1). The final database for the meta-analysis consisted of mtDNA gene sequence data from 21 studies (Table 1); some studies from the previous step were not included due to the use of RFLPs in mtDNA or because some mtDNA datasets were not publicly available. The final database spanned several taxonomic groups, with fishes, molluscs and crustaceans accounting for the majority of species (81%). The most common mitochondrial gene across all studies was cytochrome oxidase I (COI), followed by cytochrome b (Cyt b), the control region (CR) and the intergenic spacer region (IGS). COI was the most commonly used gene for invertebrate studies, IGS for macroalgae, and studies of fish used either the CR or the Cyt b gene.

Genetic structure
Sixteen species showed significant global Jost’s D and F_{ST} values, indicative of population differentiation (Table 2), while the remaining five species showed little evidence of population differentiation. Across the 21 datasets, four different types of haplotype network were putatively identified based on the structure of the networks (Fig. 2) (all haplotype networks are presented in Fig. S2):

(i) A ‘Star’ network (Fig. 2A), in which a single, widespread haplotype is typically positioned at the centre of the network and is thought to be the ancestral haplotype. Additional haplotypes are linked to this dominant haplotype by a single (or a few) mutational step(s), suggesting these haplotypes are the product of recent mutation events. Eight species showed this type of relationship (*Celleporella hyalina*, *Conger conger*, *Nassarius nitidus*, *Nassarius reticulatus*, *Palinurus elephas*, *Pelvetia canaliculata*, *Pomatoschistus*).
<table>
<thead>
<tr>
<th>Taxon species</th>
<th>MtdNA gene</th>
<th>No. sites; N</th>
<th>Sampling site distribution</th>
<th>Larval development</th>
<th>No. of lineages</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crustacean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcinus maenas</td>
<td>COI</td>
<td>13; 200</td>
<td>SW Spain to Norway</td>
<td>PLD, long</td>
<td>1</td>
<td>Roman &amp; Palumbi (2004)</td>
</tr>
<tr>
<td>Maja brachydactyla</td>
<td>COI</td>
<td>13; 291</td>
<td>SW Spain to W Ireland</td>
<td>PLD, 2–3 wk</td>
<td>1</td>
<td>Sotelo et al. (2008)</td>
</tr>
<tr>
<td>Neomysis integer</td>
<td>COI</td>
<td>9; 379</td>
<td>SW Spain to E Scotland</td>
<td>No PLD, brooder</td>
<td>1</td>
<td>Remerie et al. (2009)</td>
</tr>
<tr>
<td>Palinurus elephas</td>
<td>COI</td>
<td>6; 119</td>
<td>S Portugal to W Scotland</td>
<td>PLD, up to 1 yr</td>
<td>1</td>
<td>Palero et al. (2008)</td>
</tr>
<tr>
<td><strong>Fish</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conger conger</td>
<td>CR</td>
<td>4; 232</td>
<td>Azores to Ireland</td>
<td>Leptocephalus, up to 2 yr</td>
<td>1</td>
<td>Correia et al. (2012)</td>
</tr>
<tr>
<td>Dicentrarchus labrax</td>
<td>CR</td>
<td>9; 93</td>
<td>Bay of Biscay to Norway</td>
<td>PLD, 8–12 wk</td>
<td>1</td>
<td>Coscia &amp; Mariani (2011)</td>
</tr>
<tr>
<td>Labrus bergylta</td>
<td>CR</td>
<td>7; 279</td>
<td>W Ireland to Norway</td>
<td>PLD, 37–49 d</td>
<td>1</td>
<td>D’Arcy, Mirimin &amp; FitzGerald (2013)</td>
</tr>
<tr>
<td>Pomatoschistus microps</td>
<td>Cyt b</td>
<td>10; 232</td>
<td>Bay of Biscay to Norway</td>
<td>PLD, 6–9 wk</td>
<td>1</td>
<td>Gysels et al. (2004)</td>
</tr>
<tr>
<td>Pomatoschistus minutus</td>
<td>Cyt b</td>
<td>8; 165</td>
<td>S Portugal to Norway</td>
<td>PLD, unknown</td>
<td>1</td>
<td>Larmuseau et al. (2009)</td>
</tr>
<tr>
<td>Raja clavata</td>
<td>Cyt b</td>
<td>9; 315</td>
<td>Azores to North Sea</td>
<td>No PLD, oviparous</td>
<td>1</td>
<td>Chevrot et al. (2006)</td>
</tr>
<tr>
<td>Solea solea</td>
<td>Cyt b</td>
<td>10; 645</td>
<td>Bay of Biscay to Skagerrak</td>
<td>PLD, up to 3 wk</td>
<td>1</td>
<td>Caveliers et al. (2012)</td>
</tr>
<tr>
<td>Symphodus melops</td>
<td>CR</td>
<td>10; 263</td>
<td>S Portugal to Skagerrak</td>
<td>PLD, 14–25 d</td>
<td>1</td>
<td>Robalo et al. (2012)</td>
</tr>
<tr>
<td><strong>Macroalgae</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelvetia canaliculata</td>
<td>IGS</td>
<td>15; 429</td>
<td>Portugal to Norway</td>
<td>No PLD, external fertilisation</td>
<td>1</td>
<td>Neiva et al. (2014)</td>
</tr>
<tr>
<td><strong>Mollusc</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerastoderma edule</td>
<td>COI</td>
<td>12; 300</td>
<td>Portugal to Norway</td>
<td>PLD, up to 4 wk</td>
<td>1</td>
<td>Krakau et al. (2012)</td>
</tr>
<tr>
<td>Macoma balthica</td>
<td>COI</td>
<td>15; 399</td>
<td>Bay of Biscay to North Sea</td>
<td>PLD, 2–3 wk</td>
<td>2</td>
<td>Becquet et al. (2012)</td>
</tr>
<tr>
<td>Modiolus modiolus</td>
<td>COI</td>
<td>4; 73</td>
<td>Irish Sea to Norway</td>
<td>PLD, up to 24 wk</td>
<td>2</td>
<td>Halanych et al. (2013)</td>
</tr>
<tr>
<td>Nassarius nitidus</td>
<td>COI</td>
<td>3; 62</td>
<td>NW Spain to Sweden</td>
<td>PLD, 4–8 wk</td>
<td>1</td>
<td>Couceiro et al. (2012)</td>
</tr>
<tr>
<td>Nassarius reticulatus</td>
<td>COI</td>
<td>6; 156</td>
<td>S Portugal to UK</td>
<td>PLD, 4–8 wk</td>
<td>1</td>
<td>Couceiro et al. (2007)</td>
</tr>
<tr>
<td><strong>Polychaete</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Owenia fusiformis</td>
<td>COI</td>
<td>11; 283</td>
<td>Portugal to North Sea</td>
<td>PLD, up to 28 d</td>
<td>3</td>
<td>Jolly et al. (2005)</td>
</tr>
<tr>
<td>Pectinaria koreni</td>
<td>COI</td>
<td>10; 289</td>
<td>Portugal to North Sea</td>
<td>PLD, up to 15 d</td>
<td>2</td>
<td>Jolly et al. (2006)</td>
</tr>
<tr>
<td><strong>Bryozoan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Celleporella hyalina</td>
<td>COI</td>
<td>9; 63</td>
<td>NW Spain to Iceland</td>
<td>PLD, 1–4 h</td>
<td>1</td>
<td>Gómez et al. (2007)</td>
</tr>
</tbody>
</table>

**Notes.**
MtdNA, mitochondrial DNA; No. of sites, number of sampling sites; N, total number of sequenced individuals; PLD, pelagic larval duration.
<table>
<thead>
<tr>
<th>Species</th>
<th>Population differentiation</th>
<th>Demography</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Jost’s D</td>
<td>F&lt;sub&gt;ST&lt;/sub&gt;</td>
<td>Tajima’s D</td>
</tr>
<tr>
<td>Crustacean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Carcinus maenas</em></td>
<td>0.584</td>
<td>0.157</td>
<td>−1.73</td>
</tr>
<tr>
<td><em>Maja brachydactyla</em></td>
<td>0.298</td>
<td>0.045</td>
<td>−1.86</td>
</tr>
<tr>
<td><em>Neomysis integer</em></td>
<td>0.956</td>
<td>0.554</td>
<td>0.14</td>
</tr>
<tr>
<td><em>Palinurus elephas</em></td>
<td>0.023</td>
<td>0.000</td>
<td>−2.31</td>
</tr>
<tr>
<td>Fish</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Conger conger</em></td>
<td>0.124</td>
<td>0.000</td>
<td>−2.58</td>
</tr>
<tr>
<td><em>Dicentrarchus labrax</em></td>
<td>0.540</td>
<td>0.031</td>
<td>−1.88</td>
</tr>
<tr>
<td><em>Labrus bergylta</em></td>
<td>0.672</td>
<td>0.135</td>
<td>−0.53</td>
</tr>
<tr>
<td><em>Pomatoschistus microps</em></td>
<td>0.391</td>
<td>0.385</td>
<td>−1.39</td>
</tr>
<tr>
<td><em>Pomatoschistus minutus</em></td>
<td>0.652</td>
<td>0.100</td>
<td>−1.96</td>
</tr>
<tr>
<td><em>Raja clavata</em></td>
<td>0.375</td>
<td>0.330</td>
<td>−0.09</td>
</tr>
<tr>
<td><em>Solea solea</em></td>
<td>0.049</td>
<td>0.002</td>
<td>−2.02</td>
</tr>
<tr>
<td><em>Symphodus melops</em></td>
<td>0.578</td>
<td>0.349</td>
<td>−1.70</td>
</tr>
<tr>
<td>Macroalgae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pelvetia canaliculata</em></td>
<td>0.689</td>
<td>0.482</td>
<td>−1.53</td>
</tr>
<tr>
<td>Mollusc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Cerastoderma edule</em></td>
<td>0.662</td>
<td>0.304</td>
<td>−2.24</td>
</tr>
<tr>
<td><em>Macoma balthica</em></td>
<td>0.702</td>
<td>0.470</td>
<td>−0.80</td>
</tr>
<tr>
<td>lineage 1</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>lineage 2</td>
<td>–</td>
<td>–</td>
<td>−0.99</td>
</tr>
<tr>
<td><em>Modiolus modiolus</em>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.083</td>
<td>&lt;0.001</td>
<td>−1.79</td>
</tr>
<tr>
<td><em>Nassarius nitidus</em></td>
<td>0.222</td>
<td>0.302</td>
<td>−1.49</td>
</tr>
<tr>
<td><em>Nassarius reticulatus</em></td>
<td>0.047</td>
<td>0.000</td>
<td>−2.51</td>
</tr>
<tr>
<td>Polychaete</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Owenia fusiformis</em></td>
<td>0.788</td>
<td>0.055</td>
<td>–</td>
</tr>
<tr>
<td>lineage 1</td>
<td>–</td>
<td>–</td>
<td>−2.34</td>
</tr>
<tr>
<td>lineage 2</td>
<td>–</td>
<td>–</td>
<td>−2.06</td>
</tr>
<tr>
<td>lineage 3</td>
<td>–</td>
<td>–</td>
<td>−1.26</td>
</tr>
<tr>
<td><em>Pectinaria koreni</em></td>
<td>0.596</td>
<td>0.112</td>
<td>–</td>
</tr>
<tr>
<td>lineage 1</td>
<td>–</td>
<td>–</td>
<td>−1.99</td>
</tr>
<tr>
<td>lineage 2</td>
<td>–</td>
<td>–</td>
<td>−2.63</td>
</tr>
<tr>
<td>Bryozoan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Celleporella hyalina</em></td>
<td>0.513</td>
<td>0.488</td>
<td>−1.35</td>
</tr>
</tbody>
</table>

Notes.
*<0.05.
**<0.01.
***<0.001.
F<sub>s</sub>, Fu’s F<sub>s</sub>; R<sub>2</sub>, Ramos-Onsins’ R<sub>2</sub>; r, Harpending’s raggedness index.
<sup>a</sup>Only statistics for lineage 1 are shown.
Figure 2  Haplotype networks showing four different network structures. Haplotype networks showing (A) ‘star’ *(Palinurus elephas)*, (B) ‘complex star’ *(Carcinus maenas)*, (C) ‘reciprocally monophyletic’ *(Macoma balthica)* and (D) ‘complex mutational’ *(Dicentrarchus labrax)* structures. Each circle represents a unique haplotype and the sizes of the circles are proportional to the haplotype frequencies for each network but are not comparable across studies. Each line represents one mutation step and two or more steps are indicated by bars or numbers. Colours inside the circles correspond to sites which have individuals represented in that particular haplotype. Species illustrations by Guy Freeman.
and *Raja clavata*). In one case, the dominant haplotype had far fewer connections than a low-frequency haplotype in the network, making it difficult to distinguish the centre of the network with confidence (*Pomatoschistus microps*);

(ii) A ‘Complex star’ network (Fig. 2B), in which there are multiple high-frequency haplotypes and connections. Six species showed this type of relationship (*Carcinus maenus, Cerastoderma edule, Maja brachydactyla, Pomatoschistus minutus, Solea solea, Symphodus melops*);

(iii) A ‘Reciprocally monophyletic’ network (Fig. 2C), in which more than one lineage is apparent and each lineage is linked by a long branch associated with numerous mutations. Four species showed this type of relationship (*Macoma balthica, Modiolus modiolus, Owenia fusiformis* and *Pectinaria koreni*);

(iv) A ‘Complex mutational’ network (Fig. 2D), in which some branches were separated by a very large number of mutations, while other branches had contrarily one or two mutations. Three species showed this type of relationship (*Dicentrarchus labrax, Labrus bergylta* and *Neomysis integer*). In most cases, a dominant haplotype was present and was presumed to be the ancestral form. However, *Neomysis integer* presented an unusual network in which a distinct ancestral haplotype was not apparent and the centre of the haplotype network was not readily distinguishable.

**Historical demography**

Historical demography was inferred for each species based on the observed mismatch distribution, neutrality tests and the raggedness index (Table 2). Four main types of mismatch distributions were observed: unimodal, skewed unimodal, multimodal and bimodal (Fig. 3) (all mismatch distributions are presented in Fig. S3). Unimodal is associated with a sudden population expansion (e.g., *Maja brachydactyla*; Fig. 3A), and skewed unimodal is generally associated with a recent expansion or bottleneck (e.g., *Nassarius reticulatus*; Fig. 3B). Multimodal (e.g., *Labrus bergylta*; Fig. 3C) and bimodal (e.g., *Macoma balthica*; Fig. 3D) are usually associated with constant population size. However, previous research has suggested that bimodal peaks may indicate the presence of two distinct lineages (e.g., *Alvarado-Bremer et al., 2005*), which would potentially violate the assumptions of coalescent theory if analysed as one ‘genetic’ population. In this case, the first peak would represent intra-clade pairwise differences, whereas the second peak would likely represent more ancient inter-clade pairwise differences (Fig. 3D). For each instance of bimodality, the haplotype network was inspected for evidence of two or more lineages. The networks indicated that more than one distinct lineage was evident for all bimodal mismatches (*Macoma balthica, Modiolus modiolus, Owenia fusiformis* and *Pectinaria koreni*) and, therefore, mismatch analysis and neutrality tests were carried out on each lineage separately. These analyses were not conducted for lineage 2 of *Modiolus modiolus* due to the small number of individuals (*N* = 3) comprising this lineage.

Neutrality statistics for testing the drift–mutation equilibrium (Tajima’s *D*, *F*<sub>S</sub> and *R*<sub>2</sub>) were found to be contrasting between species (Table 2). These tests tended to be significant for species that showed a star-shaped network and for which the mismatch graph was unimodal or skewed unimodal. This supported evidence that a signal of rapid population
expansion was detected; however, a selective sweep can also produce the same genetic signal. Harpending’s $r$ suggested that two datasets departed from a model of demographic expansion (Table 2), but inspection of the mismatch graphs and neutrality tests indicated there was strong evidence to support a rapid population expansion (or selective sweep) in both datasets. No signatures of rapid population expansion were detected in five species (Celleporella hyalina, Macoma balthica, Nassarius nitidus, Neomysis integer and Raja clavata), suggesting a stable constant population size.

For the remaining 19 datasets (16 species, 19 including lineages), a historic population expansion was assumed and the timing of the expansion was estimated (Fig. 4). All expansions were found to take place during the Pleistocene or the Holocene epoch. Estimated timings for 17 datasets were after or overlapped the earliest estimate for the LGM ($\sim$26.5 Ka). Expansion estimates for one fish (Labrus bergylta) and one lineage of the polychaete Owenia fusiformis pre-dated the LGM but were still positioned during the last glacial period. Bayesian Skyline Plots (Fig. 5) were generally consistent with the results from the mismatch analyses. Among the 17 datasets for which from the mismatch analyses expansion times were estimated to have occurred after the LGM, a rise in $N_e$ post-LGM was apparent in 15 of these datasets, but the strength of the increase varied across datasets.
Figure 4 Estimated dates of expansion for species or lineages (L) in which the demographic expansion hypothesis was not rejected. A minimum and maximum time since expansion is plotted as horizontal bars for some datasets, estimated from a minimum and maximum mutation rate (Table S1). The beginning of the last glacial period (dotted line) and the estimated time-frame of the Last Glacial Maximum (grey shaded area) are displayed. Species are organised by taxa: crustaceans, *Carcinus maenas* – *Palinurus elephas*; fish, *Conger conger* – *Symphodus melops*; macroalgae, *Pelvetia canaliculata*; molluscs, *Cerastoderma edule* – *Nassarius reticulatus*; polychaetes, *Owenia fusiformis* – *Pectinaria koreni*. Species illustrations by Guy Freeman.

In comparison to the mismatch analysis, the BSP for *L. bergylta* (Fig. 5F) and *O. fusiformis* lineage 2 (Fig. 5P) indicated a population expansion after the earliest estimate for the LGM as opposed to pre-dating the LGM. In addition, although the mismatch analyses inferred a post-LGM expansion for *M. modiolus* lineage 1 (Fig. 5M) and *O. fusiformis* lineage 3 (Fig. 5Q), BSPs generally suggested $N_e$ was constant after the LGM.

**DISCUSSION**

The results of this study show a range of contemporary genetic patterns across the coastal marine taxa analysed in the northeast Atlantic. In general, genealogical patterns were
not uniform within taxonomic groups, though common patterns were observed in both polychaete species, which implies that historical events may have affected these polychaete species similarly. Most species (76%) showed evidence of population structuring, suggestive of restricted contemporary or historical gene flow between the sites studied. Of the species that exhibited no population differentiation, all five species have a pelagic larval phase, with a pelagic larval duration (PLD) ranging from up to three weeks (S. solea) to a year or more (P. elephas and C. conger) (Table 1). However, most of the species that demonstrated significant population differentiation also had a pelagic larval phase, ranging from a

Figure 5  Bayesian Skyline Plots for species or lineages (L) in which the demographic expansion hypothesis was not rejected. Solid black lines show the medium effective population size over time ($N_e = \) effective population size and $T =$ generation time); dashed black lines represent the 95% confidence intervals. The estimated time-frame of the Last Glacial Maximum is denoted by the area shaded dark grey. Species are organised by taxa: crustaceans, Carcinus maenas (A), Maja brachydactyla (B), Palinurus elephas (C); fish, Conger conger (D), Dicentrarchus labrax (E), Labrus bergylta (F), Pomatoschistus microps (G), P. minutus (H), Solea solea (I), Symphodus melops (J); macroalgae, Pelvetia canaliculata (K); molluscs, Cerastoderma edule (L), Modiolus modiolus Lineage 1 (M), Nassarius reticulatus (N); polychaetes, Owenia fusiformis Lineage 1 (O), O. fusiformis Lineage 2 (P), O. fusiformis Lineage 3 (Q), Pectinaria koreni Lineage 1 (R), P. koreni Lineage 2 (S).
relatively short PLD of 1–4 h (C. hyalina) to a relatively long PLD of 8–12 weeks (D. labrax) (Table 1). Although speculative, taken altogether, this may suggest that larval development and PLD could be important factors in maintaining gene flow in some, but not all, of these species; however, more evidence is needed to confirm this. Indeed, whether a general correlation exists between PLD and genetic differentiation measures remains unclear because some studies have reported poor correlations between the two (Weersing & Toonen, 2009; Kelly & Palumbi, 2010; Riginos et al., 2011), while other studies have reported the opposite (Siegel et al., 2003; Selkoe & Toonen, 2011) suggesting that PLD and genetic metrics can indeed reflect scales of dispersal if the sampling design is robust (Selkoe & Toonen, 2011). As a result, speculative relationships between PLD and genetic differentiation should be interpreted with caution.

In some of the species studied, certain geographical areas were dominated by a particular haplotype that was rarely or not present in other areas across the sampled range. For example, the green crab Carcinus maenas showed highly significant differentiation and distinctive haplotypes in the Faroe Islands and Iceland, a pattern detected by the original authors who subsequently concluded that a deep-water barrier to dispersal in green crabs was the driver of this pattern (Roman & Palumbi, 2004). A similar pattern was also observed for two species around western Ireland in the northeast Atlantic. In Celleporella hyalina and Macoma balthica, distinct haplotypes composed a population around western Ireland; however, unique haplotypes were not apparent in other species analysed in this study with similar sampling coverage (e.g., Labrus bergylta, Palinurus elephas and Pelvetia canaliculata). A discrepancy in genetic structure between species at this spatial scale has also been observed between two temperate octocoral species (Eunicella verrucosa and Alcyonium digitatum) using microsatellite markers, whereby northwest Ireland samples were found to be genetically isolated from other northeast Atlantic samples in E. verrucosa, but not in A. digitatum (Holland, Jenkins & Stevens, 2017). This suggests that historical or contemporary gene flow between areas in the northeast Atlantic and western Ireland is likely possible, but in some cases the spatial patterns of genetic structure could be influenced by other processes such as strong selection pressures, species-specific life history traits, demographic fluctuations, or range expansions occurring at different times in different species (Hellberg, 2009).

Demographic history
Demographic history was variable across species in the northeast Atlantic, as evidenced by both the diverse structuring of the haplotype networks and the observed mismatch distributions within species. The presence of one or more lineages and the complexity of mutational patterns in several networks suggested some species have undergone pronounced changes in their demography and genealogy. Connections with large mutation steps separating some haplotypes are indicative of deep phylogenetic splits in the genealogies and suggests the persistence of old populations in these species. Accumulating new mutations is a relatively slow process and, therefore, sufficient time since coalescence must have elapsed to facilitate these large sequence divergences (Avise, 2009).
In the northeast Atlantic, the LGM has often been viewed as a possible explanation for discrepancies in genealogies and for rapid population expansions via recolonisation as glaciers started to retreat from their maximum positions (Hewitt, 2004). In this study, we detected rapid expansions in many different taxa, of which the majority were estimated to occur after the LGM. This supports evidence for post-LGM expansions, possibly from periglacial refugia (Maggs et al., 2008) or via recolonisation of areas previously affected by the Northern Hemisphere ice sheets. These results are in contrast to the northeast Pacific where regional persistence during the LGM appeared to be common in rocky-shore organisms (Marko et al., 2010). The conclusions of several previous studies reanalysed in this meta-analysis also detected rapid expansions (e.g., Jolly et al., 2006; Sotelo et al., 2008; Larmuseau et al., 2009); however, the authors of these studies estimated the dates of these expansions to have occurred pre-LGM. This discrepancy could be due to the differences in mutation rates, whereby the original authors typically used rates derived from ancient calibrations, while in this study we attempted to use more recent calibration dates to correct for the potential time-dependency of molecular rates (Ho et al., 2011).

Of course, we acknowledge that the signal of deviation from neutrality we detected may, in some cases, be the result of a selective sweep and not a rapid expansion. This signal could be distinguished by incorporating multi-locus data; nevertheless, given that a variety of species in this study showed similar genealogical patterns consistent with demographic expansion, it seems likely that most of them did indeed experience demographic changes associated with the end of the LGM, rather than selective sweeps. Moreover, distinctive haplotypes were found in several species networks (Pelvetia canaliculata, Pomatoschistus minutus, Owenia fusiformis and Pectinaria koreni) to the south of where the Eurasian ice sheet is proposed to have extended during the LGM (Fig. 1). This finding suggests populations of these species may have survived in southern glacial refugia; though, as pointed out by some of the original authors, deep sequence divergences in some species (e.g., O. fusiformis and P. korena) and the lack of a species-specific molecular clock calibration makes inferences about refugia challenging (Jolly et al., 2005; Jolly et al., 2006).

It is difficult to suggest an explanation for the two expansions estimated to have pre-dated the LGM (using mismatch analysis), but which fall within the last glacial period. This pattern of pre-LGM expansion has also been reported in a number of previous studies for a variety of marine taxa (e.g., Hoarau et al., 2007; Marko et al., 2010; Ni et al., 2014; Almada et al., 2017). One potential explanation for this pattern is that sea level during the last glacial cycle did not decrease uniformly towards the level observed at the LGM, but oscillated rapidly over a period of 60 Ka to 30 Ka (see Fig. 3A in Lambeck, Esat & Potter, 2002). Therefore, it may be possible that we are detecting the signature of a population expansion during one of these sudden increases in sea level during the last glacial period. Alternatively, as the BSP analysis inferred a post-LGM expansion for these two datasets, this could be a limitation associated with the mismatch analysis approach, which does not consider genealogy, and may, therefore, produce a less precise estimation. In addition, the sample of genetic diversity for this species may not be representative (Karl et al., 2012) or the genetic signal we detected may have been the result of a selective sweep and not a rapid expansion.
The use of single marker mtDNA genealogies and coalescence theory can introduce challenges associated with the interpretation of data and these limitations should be acknowledged (Karl et al., 2012). For example, the populations under study may have experienced multiple episodes of growth and decline; however, only the most recent expansion event can be detected using coalescence analysis and, in some cases, these events may not be sufficiently severe to be detected (Karl et al., 2012). In addition, coalescent histories can differ amongst loci because they can experience mutation and drift independently. Therefore, analysis of a single gene only gives insight into the coalescent history of that locus, which may not always be representative of population history. Analysis of multiple loci and genomics would help to alleviate these concerns, and would likely provide enhanced resolution for exploring the phylogeography of northeast Atlantic marine fauna.

Although population expansions were detected in a number of species in this study and also in the wider literature, populations of other marine species, including five from this study, have been found to remain stable throughout the LGM. As previously reported, not all coastal marine taxa appear prone to demographic changes during or after ice ages (Janko et al., 2007; Marko et al., 2010; Olsen et al., 2010). It is also important to acknowledge that earlier events in the Pleistocene and more ancient events that pre-date the Pleistocene may have helped shape the contemporary patterns of genealogical structure observed in this study.

CONCLUSION

The findings of this meta-analysis indicate that species in the northeast Atlantic do not show a uniform pattern of phylogeography, but rather a mixture of complex contemporary genealogical structure. Reanalysis of demographic histories indicated that a large proportion of the species included in this study have experienced post-LGM expansions, supporting the general expectation that rapid population expansions occurred after the LGM as the ice sheets started to retreat (Hewitt, 2000; Hewitt, 2004). This suggests that regional extirpation during the LGM appears to be a common biogeographic history for many northeast Atlantic marine taxa. However, improvements in mutation rate estimates, as well as the incorporation of multi-locus markers and genomics, would likely provide greater accuracy and resolution for overcoming the challenges associated with single mtDNA genealogies, and for improving our understanding of phylogeography in the northeast Atlantic Ocean.

ACKNOWLEDGEMENTS

We thank the many authors who kindly provided or contributed to the original datasets used in the meta-analysis. We also thank SYW Ho (Sydney), CN Roterman (Oxford) and KF Thompson (Exeter) for insightful comments on the manuscript, and GW Freeman (Exeter) for species illustrations.
Funding
This research was funded by a Natural Environment Research Council (UK) GW4+ DTP studentship (Grant No. NE/L002434/1), Natural England (Ref. PO 904130) and the University of Exeter. Rita Castilho was funded by FCT—Foundation for Science and Technology through project UID/Multi/04326/2013. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the authors:
Natural Environment Research Council (UK): NE/L002434/1.
Natural England: PO 904130.

Competing Interests
Rita Castilho is an Academic Editor for PeerJ.

Author Contributions
• Tom L. Jenkins conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
• Rita Castilho and Jamie R. Stevens conceived and designed the experiments, authored or reviewed drafts of the paper, approved the final draft.

Data Availability
The following information was supplied regarding data availability:
The raw data is available in the Supplemental File.

Supplemental Information
Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.5684#supplemental-information.

REFERENCES


