Journal of Biogeography (J. Biogeogr.) (2014) 41, 1171–1182

ORIGINAL ARTICLE

Anchovies go north and west without losing diversity: post-glacial range expansions in a small pelagic fish

Gonçalo Silva, John B. Horne and Rita Castilho*

Centre of Marine Sciences, CCMAR/CIMAR Associate Laboratory, University of Algarve, Gambelas 8005-139 Faro, Portugal

*Correspondence: Centre of Marine Sciences, CCMAR/CIMAR Associate Laboratory, University of Algarve, Gambelas, 8005-139 Faro, Portugal.
E-mail: rcastil@ualg.pt

ABSTRACT

Aim As part of an emerging effort to understand the role played by climatic fluctuations in shaping the geographical distributions and abundances of marine organisms, we examined the genetic patterns of leading-edge populations in the European anchovy, *Engraulis encrasicolus*, and its American counterpart, the morphologically similar silver anchovy, *Engraulis eurystole*, in the North Atlantic Ocean.

Location Adults were collected from the western Atlantic, eastern Atlantic (from Norway to Ghana) and western Mediterranean.

Methods A 1045 bp fragment of the mtDNA cytochrome *b* gene was sequenced (n = 312) and nine microsatellite loci were genotyped (n = 462) for anchovies from 13 locations across the temperate North Atlantic. Populations were surveyed for diversity and differentiation with a range of summary statistics. Multivariate discriminant analysis of principal components was employed to detect the number of genetic clusters in the data and assign individuals to populations based on their microsatellite genotypes. Historical demographic inferences – mismatch distributions and Bayesian skyline plots – were used to observe population size changes relating to climatic oscillations.

Results Two mitochondrial clades were recovered, consistent with previous studies of *E. encrasicolus*, in which the frequency of each clade varied by latitude. Four genetic clusters corresponding loosely to large geographical regions were identified with microsatellite data. The north-western Atlantic *E. eurystole* was not reciprocally monophyletic for either mtDNA or microsatellite analyses and is probably conspecific with *E. encrasicolus*. Genetic diversity peaked in Iberian populations, but differences in genetic diversity were only statistically significant for the least diverse population, Tangier. The indications of demographic expansion were more pronounced in the southern clade and both mtDNA clades exhibited genetic diversity and expansion imprints that are likely to be older than climatic oscillations of the recent Pleistocene.

Main conclusions The highly mobile nature of anchovies has allowed them to track their optimal thermal physiological conditions during the extreme climate shifts of the Last Glacial Maximum and avoid wholesale population reductions and genetic bottlenecks. Both north-eastern and north-western Atlantic were probably rapidly recolonized after the Last Glacial Maximum by large numbers of anchovies, such that leading-edge populations retained the genetic diversity of parent populations.

Keywords

Anchovy, climatic fluctuation, dispersal dynamics, *Engraulis*, historical biogeography, leading-edge population, phylogeography, post-glacial expansion, trans-Atlantic migration.

© 2014 John Wiley & Sons Ltd
http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/jbi.12275

Journal of Biogeography

INTRODUCTION

A major goal of evolutionary biogeography is to understand the effects of climatic fluctuations on the geographical distribution and genetic diversity of species. Distribution ranges of marine temperate organisms in the Northern Hemisphere typically shift southwards during periods of climate cooling and northwards during climatic optima, in pursuit of optimal thermal conditions that change latitudinally. These changes in population ranges influence the genetic characteristics of populations, especially those at the periphery of the geographical range (McInerny *et al.*, 2009). By investigating genetic patterns in peripheral, leading- or rear-edge populations, we can better understand how natural populations have responded to changes in climate during their evolutionary past.

During the Last Glacial Maximum (LGM, *c.* 20 ka), sea surface temperatures in the North Atlantic fell by 10 °C on average (CLIMAP, 1981) and ice-sheet margins descended to 45° N (Renssen & Vandenberghe, 2003) (Fig. 1a). Sea levels dropped 120 m and altered coastal contours, creating land-bridge barriers to marine organisms in some areas (Lambeck *et al.*, 2002). Throughout the LGM, high latitude coastal areas of the North Atlantic were frequently glaciated and unavailable as pathways for dispersal and population persistence. Although some species can tolerate negative or near 0 °C temperatures (e.g. *Clupea harengus*; McInerny *et al.*, 2009), the majority of temperate marine organisms would have been unable to cope with LGM conditions throughout most of the North Atlantic and would have experienced dramatic shifts in their distributions and severe population bottlenecks as a result of widespread local extinction in the northern parts of their ranges (Alheit *et al.*, 2012). Therefore, present-day distributions of temperate marine species in the northern areas of the North Atlantic largely represent recolonizations from refugial populations that have occurred since the LGM (Maggs *et al.*, 2008).

Anchovies (genus *Engraulis*) are found in a wide range of temperatures (2–30 °C) and salinities (5–41 psu) (Whitehead *et al.*, 1988) and these physical parameter ranges define the ecophysiological tolerance and limitations of this species to geographical shifts. The European anchovy, *Engraulis encrasicolus* (Linnaeus, 1758), inhabits the eastern Atlantic, the Mediterranean Sea and the Black Sea (Whitehead *et al.*, 1988), while its less abundant congener, the silver anchovy, *Engraulis eurystole* (Swain & Meek, 1884), occupies the north and central western Atlantic. Morphological differences between these two species are minor, yet the geographical isolation between both sides of the Atlantic led previous authors to maintain these taxa as different species until 'careful revisionary work' is undertaken (Whitehead *et al.*, 1988).

Even though anchovies have a large dispersal potential, genetic studies show moderate levels of genetic structure among populations in the Mediterranean Sea (e.g. Magoulas *et al.*, 2006) and in the eastern North Atlantic (Petitgas *et al.*, 2012). The geographical distribution of genetic diversity

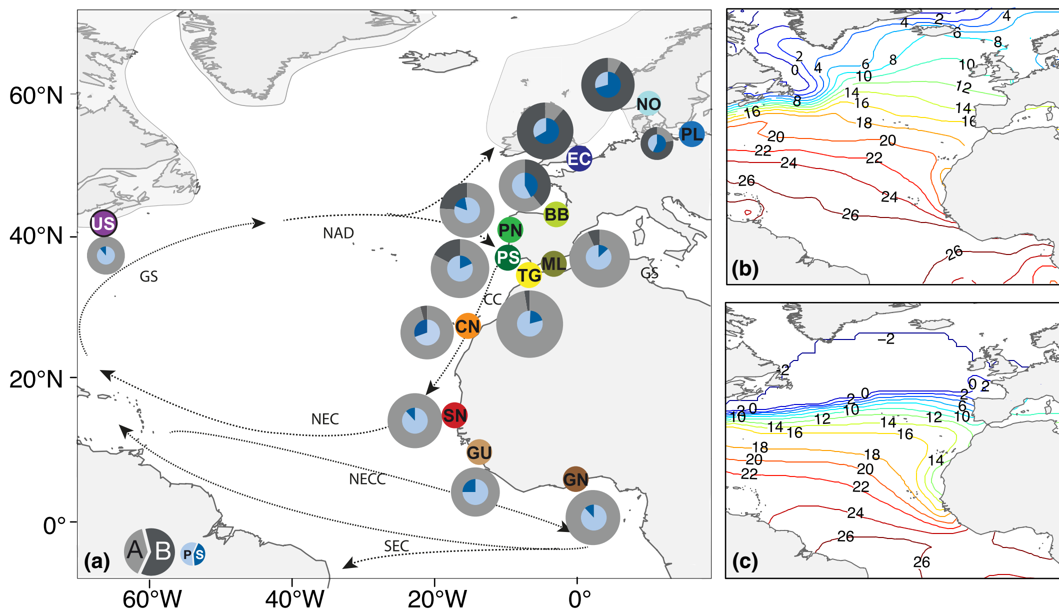

Figure 1 (a) Present-day Atlantic Ocean and western Mediterranean Sea showing *Engraulis encrasicolus* and *E. eurystole* sample locations. Light grey shading represents the extent of terrestrial ice sheets at the Last Glacial Maximum (LGM; 18–20 ka). Sample locations examined for mtDNA cytochrome *b* (cyt *b*) are indicated by coloured circles. Grey circles represent mtDNA cyt *b* clades A (light grey) and B (dark grey) proportions. Blue circles represent mtDNA cyt *b* private (light blue) and shared (dark blue) haplotype proportions. Black dashed lines represent main oceanic currents, and arrows show directionality (GS, Gulf Stream; NAD, North Atlantic Drift; CC, Canaries Current; NEC, North Equatorial Current; NECC, North Equatorial Counter Current; SEC, South Equatorial Current). Sample abbreviations are defined in Table 1. (b) Climatological present-day January mean sea surface temperature fields (SST, °C). (c) Climatological mean sea surface temperature fields (SST, °C) simulation results for the LGM based on seasonal estimates of CLIMAP (1981).

© 2014 John Wiley & Sons Ltd

Table 1 Sample locations, sample abbreviations, collection dates, sample sizes and summary statistics for a 1045 bp sequence fragment of the mtDNA cytochrome *b* and eight nuclear microsatellites of the anchovies *Engraulis encrasicolus* and *E. eurystole*.

				Mitochondrial cytochrome *b*						Microsatellites (8 loci)						
Location	Code	Long	Lat	n	n_h	n_P	n_P/n_h	h	π	n	A_{avg}	$A_r = 9$	$A_r = 18$	*Effnum*	H_O	H_E
Norway	NO	10.6	59.0	24	17	5	0.29	0.953	0.009	40	13.1	7.6	10.5	7.0	0.714	0.850
Poland	PL	16.5	54.6	9	7	3	0.43	0.917	0.014	9	8.4	8.4	–	6.1	0.736	0.864
English Channel	EC	0.1	50.8	27	18	6	0.33	0.963	0.010	45	13.0	6.8	9.6	6.6	0.729	0.837
Bay of Biscay	BB	−2.9	43.5	23	19	11	0.58	0.980	0.015	45	16.0	8.4	11.6	6.9	0.762	0.837
Portugal – North	PN	−8.8	40.7	25	24	20	0.83	0.997	0.013	45	17.3	8.1	10.9	8.4	0.825	0.882
Portugal – South	PS	−8.4	37.1	29	27	22	0.81	0.995	0.011	43	17.9	8.0	12.3	8.6	0.777	0.880
Málaga	ML	−4.3	36.6	31	31	27	0.87	1.000	0.007	46	16.6	9.5	11.3	8.3	0.758	0.880
Tangier	TG	−5.9	35.9	38	10	8	0.80	0.423	0.003	46	14.0	6.1	9.3	5.3	0.686	0.800
Canary Islands	CA	−15.0	28.3	24	23	16	0.70	0.996	0.007	42	17.8	8.9	11.8	9.9	0.792	0.888
Senegal	SN	−17.6	14.8	25	25	22	0.88	1.000	0.006	37	12.5	7.3	8.5	5.0	0.693	0.797
Guinea-Bissau	GU	−14.2	9.7	20	20	20	1.00	1.000	0.006	19	13.1	7.8	12.8	6.8	0.653	0.868
Ghana	GH	0.0	5.6	25	25	25	1.00	1.000	0.006	27	15.5	8.4	13.6	8.7	0.766	0.883
USA*	US	−66.1	41.5	12	9	8	0.89	0.909	0.004	18	13.5	9.5	13.5	8.0	0.793	0.894
Total				312	210	208	0.99	0.985	0.014	462	14.6	25.1	32.0	6.7	0.745	0.858

Long, longitude; Lat, latitude; n, number of individuals; n_h, number of haplotypes; n_P, number of private haplotypes; n_P/n_h, proportion of private haplotypes; h, haplotype diverstity; π, nucleotide diversity; *Aavg*, average number of alleles, *Ar*, allelic richness; *Effnum*, effective number of alleles; H_O, observed mean heterozygozity; H_E, expected mean heterozygozity; *putative *Engraulis eurystole*.

indicates that the potential for dispersal may be limited (Magoulas *et al.*, 2006) by such factors as retentive currents, complex shorelines, oceanic fronts or isolation by distance (Agostini & Bakun, 2002). In addition to the genetic structure imposed by partial isolation between locations, two deeply separated mitochondrial DNA (mtDNA) clades varying in relative frequency were identified in *E. encrasicolus*, which appear to reflect ancient isolations and recent secondary contact (Magoulas *et al.*, 2006).

Our study is part of an emerging effort to understand the role played by climatic fluctuations in shaping the geographical distributions and genetic diversity of marine organisms in the North Atlantic (e.g. Maggs *et al.*, 2008, and references therein). We provide a phylogeographical study based on a large fragment of the mitochondrial cytochrome *b* gene (cyt *b*) and nine nuclear microsatellites. The emphasis of the work is: (1) on the levels of genetic diversity of recently established northern populations compared to persistent more southern populations; and (2) on the genetic diversity and differentiation between *E. encrasicolus* and *E. eurystole*, in the eastern and western North Atlantic, respectively. We build on previous studies (Magoulas *et al.*, 2006; Bouchenak-Khelladi *et al.*, 2008; Sanz *et al.*, 2008; Borrell *et al.*, 2012; Zarraonaindia *et al.*, 2012) to provide insights on the timeframe, the source and the mode of colonization of marginal populations. We predict that recently colonized northern populations of the European anchovy are comparable in genetic diversity to populations further south, given that anchovies are highly mobile pelagic fish able to rapidly track suitable habitats in large numbers. Recolonization involving large numbers of migrants does not result in decreased genetic diversity in the new populations, as is expected when populations are originally founded by only a few colonizers (McInerny *et al.*, 2009). For the same reason, we do not predict finding signals of sudden demographic population growth within the post-LGM timeframe, contrary to expectations of newly founded populations from low dispersal species.

MATERIALS AND METHODS

Sample collection, DNA extraction and PCR amplification

Samples of *E. encrasicolus* and *E. eurystole* were collected at a total of 13 sites in the eastern and western North Atlantic Ocean (Table 1, Fig. 1a) and at one site in the western Mediterranean, where the *E. encrasicolus* population was previously described as being genetically close to Atlantic populations (Magoulas *et al.*, 2006). Fish were purchased at small coastal fish markets, as artisanal fisherman do not venture far, or were collected on scientific cruises (see Acknowledgements). A small portion of white muscle or fin was preserved in 96% ethanol and stored at −20 °C. Total genomic DNA was extracted by a saline method (Sambrook & Russell, 2001). A 1045 bp fragment from the mitochondrial cyt *b* and nine nuclear microsatellite loci were amplified by polymerase chain reaction (PCR). Sequences and fragment lengths (using the GeneScan-500 LIZ standard; Life Technologies Europe BV, Porto, Portugal) were obtained using an ABI 3130XL automated sequencer (Applied Biosystems, Foster City, CA, USA) see Appendix S1 in the Supporting Information for primers, amplification details and fragment characteristics). Microsatellite raw allele sizes were manually scored in STRand 2.4.59 (Toonen & Hughes, 2001).

Genetic analysis

Cyt *b* sequences were aligned using ClustalX 2.0.3 with default settings, implemented in Geneious 5.4 (Drummond

© 2014 John Wiley & Sons Ltd

et al., 2011) and checked manually. We used Collapse 1.2 (Posada, 2004) to reduce sequences to haplotypes. For cyt *b*, number of individuals (n), frequency (f), number of haplotypes (n), number of private haplotypes (n_P), and haplotype (h) and nucleotide diversities (π) were calculated in Arlequin 3.5.1.2 (Excoffier & Lischer, 2010). To compare haplotype diversities, the Salicru method (Salicru *et al.*, 1993) was used for both overall diversity and pairwise locations.

In previous analyses, the European anchovy displayed a high proportion of null alleles for microsatellite loci (e.g. Zarraonaindia *et al.*, 2009). Therefore, we used micro-checker 2.2.3 (van Oosterhout *et al.*, 2004) and freena (Chapuis & Estoup, 2007) to calculate the frequency of null alleles at different loci, and F_{ST} values (Weir, 1996) were re-calculated after correcting for the presence of null alleles.

Summary statistics, number of individuals (n), mean number of alleles (n_a), observed heterozygosity (H_O) and expected heterozygosity (H_E) were calculated for each location and for each locus with genodive, whereas mean allelic richness (A_r) was calculated with fstat 2.9.3.2 (Goudet, 1995).

Genetic structure and population differentiation

To examine the relationship between mitochondrial haplotypes, a minimum spanning network was constructed with Arlequin 3.5.1.2 (Excoffier & Lischer, 2010) and visualized with HapStar (Teacher & Griffiths, 2011). Population pairwise genetic differentiation was estimated with G_{st_est} (Hedrick & Goodnight, 2005) and Jost's D_{est} value (Jost, 2008) following Pennings *et al.* (2011) for mtDNA, and using the R package diversity (Keenan *et al.*, 2013) for microsatellites.

Spatial analysis of shared alleles (SAShA) was used to detect subtle geographical structuring of mtDNA haplotypic and microsatellite allele co-occurrences (Kelly *et al.*, 2010). This spatial analysis is done by comparing the observed distance distribution (ODD) between occurrences of each allele, with a null expected distance distribution (EDD) generated from the data. To test whether the geographical pattern of genetic differentiation is caused by isolation by distance (IBD) we ran Mantel tests for pairwise matrices between geographical distances (kilometres) of the shortest marine path among locations measured in Google Earth and genetic differentiation [$D_{est}/(1 - D_{est})$]. Mantel tests (1000 randomizations) were performed using mantel.xla 1.2.4 (Briers, 2003). To determine genetic structuring and individual assignments based on the autosomal microsatellite data set, we used discriminate analysis of principal components (DAPC) a multivariate ordination method (Jombart *et al.*, 2008) implemented in the adegenet package (Jombart, 2008) of R 2.15.3 (R Development Core Team, 2009; http://www.r-project.org). This method does not assume Hardy–Weinberg equilibrium or linkage disequilibrium and is more appropriate for situations where such assumptions are not met, as is often the case with anchovies (Zarraonaindia *et al.*, 2009), than conventional approaches such as structure (Pritchard *et al.*, 2000). DAPC yields similar results to structure (van der Meer *et al.*, 2012; Molfetti *et al.*, 2013) predicting genetic clusters based on the results of principal components analysis. A user-specified number of principal components is retained to act as predictors, in this case, representing 90% of the cumulative variance. The optimal number of populations was identified as the one for which the Bayesian information criterion (BIC) showed the lowest value and after which BIC increased or decreased by the least amount. We followed Jombart's (2013) recommendation to perform a cross-validation of the robustness of cluster assignments, by splitting the data in two parts. Twenty-five per cent of the individuals from each sample were chosen as training data and the remaining 75% were hold-out data.

Historical demography

Two neutrality tests were used to assess population expansion, Fu's F_S (Fu, 1997) and R_2 (Ramos-Onsins & Rozas, 2002). Mismatch distributions, frequencies of pairwise differences between haplotypes, were estimated for each clade. Significance of F_S and R_2 was evaluated by comparing the observed value with a null distribution generated by 10,000 coalescent simulations, using the empirical population sample size and observed number of segregating sites implemented in DnaSP 5.10 (Librado & Rozas, 2009). Sum of squares deviations (SSD) and raggedness statistics (*rg*; Harpending, 1994) significances were obtained based on 10,000 permutations. Lower R_2 and *rg* values are expected for a population growth scenario (Harpending, 1994; Ramos-Onsins & Rozas, 2002). These analyses were performed in Arlequin 3.5.1.2 (Excoffier & Lischer, 2010). DnaSP (Librado & Rozas, 2009) was used to obtain observed and expected distributions under the constant population model and the growth population model. Time and magnitude of inferred population expansion can be determined by mismatch analysis (Rogers & Harpending, 1992) by calculating θ_0, θ_1 and τ, where $\theta_0 = 2N_0\mu$ (N_0 = population size before expansion); $\theta_1 = 2N_1\mu$ (N_1 = population size after expansion); and $\tau = 2ut$ (u = is the mutation rate over the whole sequence; t = time since population growth expansion). Coalescence analysis requires an estimate of generation time and mutation rate. The generation time, defined as the average age of reproductively mature individuals in the population, is one year in *E. encrasicolus* (Parada *et al.*, 2003). The only specific divergence rate available for cyt *b* of Engraulidae is 1.9% Myr^{-1} (Grant *et al.*, 2010), roughly equivalent to a substitution rate of 1% Myr^{-1}, well within what has been usually accepted in many species of bony fishes (Bowen *et al.*, 2001; and references therein). Given a mean generation time of 1 year, the substitution rate per site per generation in cyt *b* is 1×10^{-8}. However, this rate is merely a heuristic and absolute molecular dating was not inferred using this rate. The Bayesian skyline plot (BSP) analysis of population size history, using beast 1.7.5 software (Drummond *et al.*, 2012) was applied to cyt *b* sequences. Genealogies were combined from 10 runs of 1×10^8 steps with a burn-in of 1×10^7 (for clade A) and 10 runs of 10^7 steps with a burn-in of 1×10^6 steps (for clade B). Runs were performed

© 2014 John Wiley & Sons Ltd

under the GTR+I+Gamma model, with a strict molecular clock and a stepwise skyline model with 20 piecewise intervals. Genealogies and model parameters were sampled every 10,000 iterations and operators were optimized automatically. Effective sample size (ESS) for each parameter exceeded 200. The trajectories were plotted by Tracer 1.5 (Rambaut & Drummond, 2007). Mismatch distributions and BSP trajectories were performed to evaluate expected signatures of demographic expansion between subsets of the data, judged to represent a coherent demographic history based on the geographical settings and the main genetic mtDNA structure.

RESULTS

Mitochondrial DNA

A total of 312 individuals were analysed (accession numbers JQ716609–JQ716731, JQ716748–JQ716756 and JX683020–JX683113). Sequences were polymorphic at 216 sites (119 parsimony-informative) defining 210 haplotypes, of which 97 were singletons and 184 (87.6%) were private (i.e. unique to a single locality). Haplotype diversity (h) was high, ranging from 0.909 to 0.980 in northern locations (Bay of Biscay to Norway and USA) and from 0.423 (Tangier) to 1.000 in southern locations (Table 1). Differences in haplotype diversity among the 13 locations were only significant when Tangier was included (χ^2 = 39.54, d.f. = 12, $P > 0.01$; range of z-values = 0.0–1.7; global $P = 0.00009$). All haplotype diversity pairwise comparisons with Tangier were also significant. When Tangier was removed from the dataset, global haplotype diversity values were not significantly different (χ^2 = 8.01, d.f. = 11, range of z-values = −5.63–5.65, $P = 0.71$), nor were any of the remaining pairwise comparisons. Nucleotide diversity (π) was low, ranging from 0.3% (Tangier) to 1.5% (Bay of Biscay) (Table 1). Locations north of Tangier displayed higher nucleotide diversities while Senegal, Guinea-Bissau and Ghana had lower diversity values (Table 1). Diversities in the sample from the Alboran Sea (h = 1.000, π = 0.007) were similar to those for north-eastern Atlantic samples. On the whole, diversity measures did not decrease towards the marginal northern locations (see Appendix S2).

Haplotypes grouped into two previously described clades (Magoulas *et al.*, 1996) with frequencies shifting clinally. Clade A haplotypes were more frequent in the southern samples while clade B predominated in samples north of the English Channel. North-eastern Atlantic locations exhibited a higher proportion of shared haplotypes, in contrast with more southern locations (Fig. 1a). Both clades were present in all locations, with the exception of West Africa (Senegal, Guinea-Bissau and Ghana) and USA, where clade B is absent (Fig. 1a). Strong clade-frequency shifts occurred between the English Channel and the Bay of Biscay. Net sequence divergence between clades was 1.87% (SE 0.37%).

In the haplotype network, clades A and B were separated by 14 mutations (Fig. 2). The two clades exhibited different haplotype patterns: clade A was characterized by multiple star-like radiations with relatively shallow genetic divergences; clade B lacked distinct star patterns and exhibited many unsampled or extinct haplotypes. USA haplotypes were separated by two to five mutations from the eastern Atlantic haplotypes, with one haplotype shared with the Canary Islands. Furthermore, western Atlantic haplotypes were not reciprocally monophyletic with respect to European lineages.

SAShA analyses rejected population panmixia (ODD = 785 km, EDD = 3295 km, $P < 0.001$), because the mean geographical dispersal of alleles was smaller than expected. Pairwise differentiation was much greater when measured using D_{est} than G_{st_est} (Fig. 3). Additionally, adjacent locations had consistently lower D_{est} values. G_{st_est} values were mostly between 0 and 0.05 with the notable exception of Tangier and USA. IBD was significant, when considering all samples ($r^2 = 0.298$, $P < 0.001$) and also when excluding the USA ($r^2 = 0.222$, $P < 0.001$).

Clade A displayed a typical unimodal distribution (Fig. 4a), closely matching the expectations of the growth population model, while clade B was clearly bimodal (Fig. 4b). However, SSD and raggedness probability values did not allow rejection of the sudden expansion model for both clades and Fu's F_S

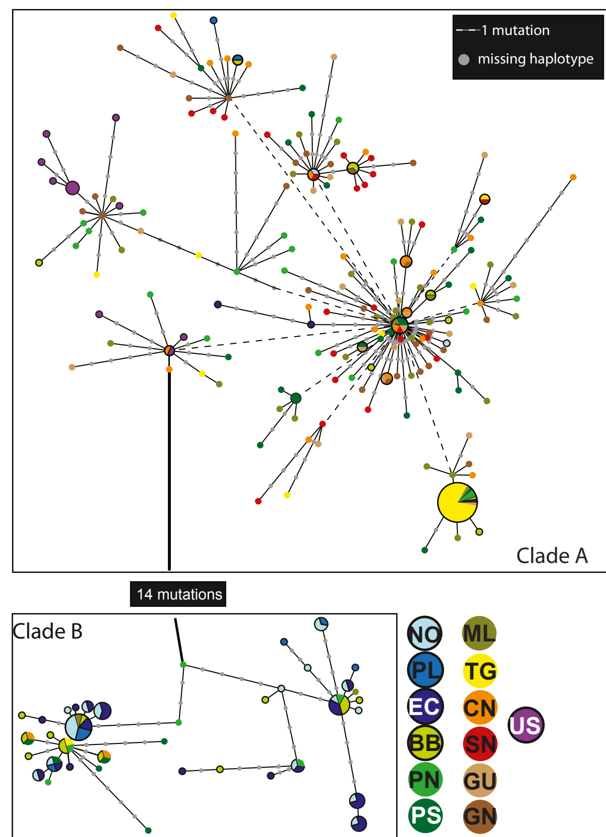

Figure 2 Mitochondrial DNA cytochrome *b* haplotype network of *Engraulis encrasicolus* and *E. eurystole* constructed with a median-joining algorithm. Haplotypes are coloured according to Fig. 1 labels and sample abbreviations are defined in Table 1. A black outline of haplotypes indicates their origin from recently colonized areas (USA and the north of Europe, English Channel and Bay of Biscay). Dashed lines indicate one mutation.

© 2014 John Wiley & Sons Ltd

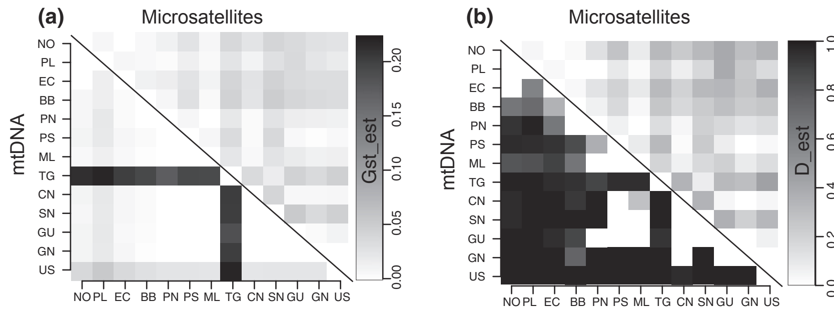

Figure 3 Pairwise-location G_{st_est} (a) and D_{est} (b) values for microsatellites (above diagonal) and mtDNA sequences (below diagonal) of *Engraulis encrasicolus* and *E. eurystole*. Site abbreviations defined in Table 1.

values were negative and significant, which might also indicate population expansion. Bayesian skyline plots revealed remarkable population size stability throughout most of the evolutionary history of the lineages (Fig. 4c,d).

Microsatellite DNA

Multilocus genotypes from 462 anchovies were obtained from 13 locations. The number of alleles per locus varied from 14 (locus 135) to 152 (locus 41.2) over all locations (Table 2, Appendix S3). Mean allelic richness, standardized for comparison across a minimum common sample size of nine individuals, ranged from 6.1 (Tangier) to 9.5 (Malaga and USA). Expected heterozygosity (H_E) varied between 0.797 (Senegal) and 0.894 (USA) and the observed heterozygosity (H_O) varied between 0.653 (Guinea-Bissau) and 0.825 (Portugal north) (Table 1). North Atlantic locations south of the English Channel have higher levels of allelic richness than northern European

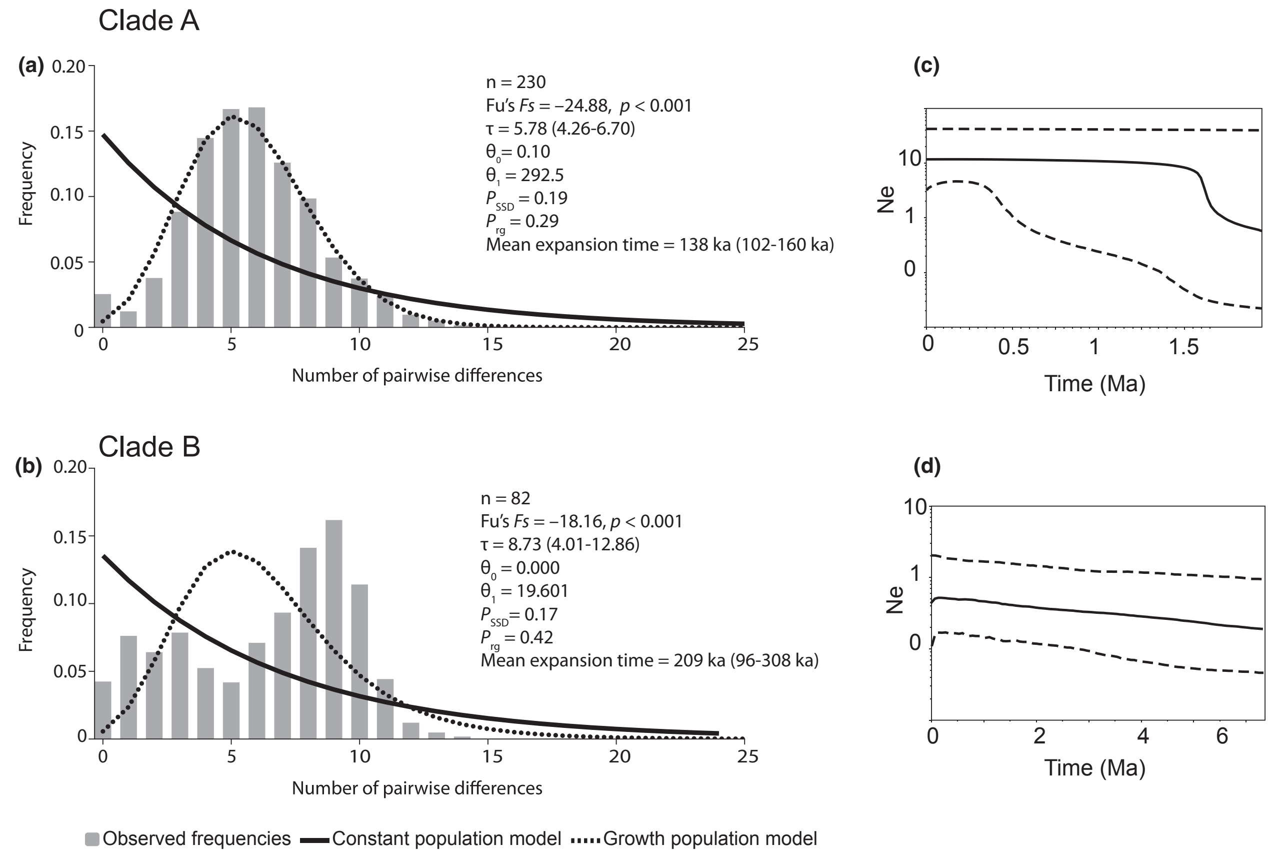

Figure 4 Mismatch distribution of *Engraulis encrasicolus* and *E. eurystole* in the two panels on the left (a and b), indicating number of individuals in the analysis (n), Fu's F_S test of selective neutrality and population expansion, evolutionary expansion age in mutational units (τ), effective population size before (θ_0) and after (θ_1) population expansion, and mean expansion time in units of thousand years (ka). Note that the range in expansion age corresponds to the 95% confidence interval of (τ). P_{SSD} represents the significance of sum of squares deviations and P_{rg} the significance of the raggedness statistics. Bayesian skyline plots are shown in the two panels to the right (c and d). The *y*-axis (note logarithmic scale) indicates effective population size estimates multiplied by generation time. Operational time-scales (*x*-axis) are based on per site mutation rates and a 1-year generation interval. Solid lines are median estimates of effective population size (N_e), dotted lines are the 95% posterior density limits. Time at the origin represents the present day.

© 2014 John Wiley & Sons Ltd

Table 2 Summary statistics across nine microsatellite loci of the anchovies *Engraulis encrasicolus* and *E. eurystole*. *Na*, total number of alleles; *Ar*, average number of alleles across locations; *Effnum*, effective number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity within populations; H_t, total heterozygosity; H'_t, corrected total heterozygosity; G_{IS}, inbreeding coefficient; Null allele$_F$, null allele frequency by FREENA; Null allele$_M$, null allele frequency by MICRO-CHECKER; F_{ST}, global per locus F_{ST} of Weir (1996); F_{ST}(ENA), global per locus F_{ST} of Weir (1996) using the ENA correction described in Chapuis & Estoup (2007).

Locus	*Na*	*Ar* = 9	*Effnum*	H_O	H_E	H_t	H'_t	G_{IS}	Null allele$_F$	Null allele$_M$	F_{ST}	F_{ST}(ENA)
L10	72	11.4	7.2	0.842	0.877	0.922	0.925	0.040	0.013	†	0.047	0.046
L135	14	7.6	5.8	0.756	0.845	0.862	0.864	**0.105**	0.036	0.068	0.024	0.024
L407	55	10.1	6.5	0.828	0.863	0.911	0.915	0.040	0.016	†	0.059	0.058
L452	48	12.3	11.0	0.893	0.926	0.952	0.954	0.036	0.011	†	0.029	0.028
L508	26	10.2	6.6	0.601	0.871	0.922	0.927	0.309	0.142	0.032	0.068	0.064
L291a	28	8.2	6.5	0.628	0.866	0.881	0.882	**0.275**	0.119	0.031	0.015	0.013
L41.1	57	8.5	5.0	0.604	0.819	0.858	0.861	**0.262**	0.102	0.015	0.051	0.043
L41.2	152	16.0	20.8	0.617	0.977	0.987	0.988	**0.369**	*	0.005	0.010	*
L291b	17	6.1	4.6	0.807	0.798	0.809	0.810	−0.011	0.002	†	0.017	0.017
Overall	52	10.0	8.2	0.731	0.871	0.900	0.903	**0.161**				

Significant values are in bold, $P < 0.05$.
*Number of alleles exceeds software capability.
†Software does not correct these loci.

locations. Allelic diversities in the Mediterranean sample (A_{avg} = 16.6) were similar to those in nearby Atlantic locations (Portugal north, Portugal south, Tangier and Canaries (A_{avg} = 14.0–17.9). Average expected heterozygosities showed no geographical pattern (Appendix S2). Locus 41.2 was identified as possibly having null alleles and/or large allele dropout, as it presented high overall G_{IS} (Table 2). Null allele uncorrected and corrected estimated per locus F_{ST} values are very similar (Table 2), but because of software limitations those comparisons could not be performed for locus 41.2. Therefore, we removed this locus from further analysis to ensure that erroneous estimations were not introduced by the hypervariable nature of the locus.

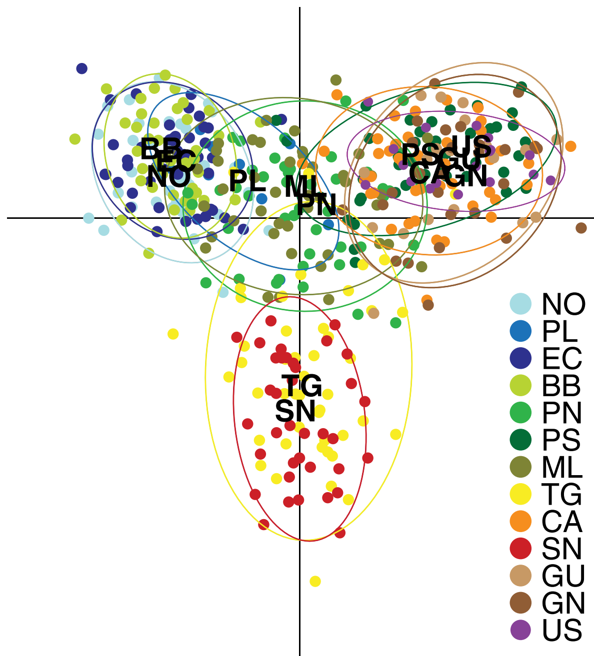

Figure 5 Discriminant analysis of principal components (DAPC) of multi-locus *Engraulis encrasicolus* and *E. eurystole* genotype data for all study locations. Individual genotypes appear as circles, and sample code (defined in Table 1) represents the centre of dispersion of each group. Horizontal and vertical axes are the first two principal components, respectively.

SAShA suggests that anchovies are not panmictic (ODD = 2870 km, EDD = 3094 km, $P < 0.001$). Also, IBD was significant only when the western Atlantic was included (with USA: $r^2 = 0.053$, $P = 0.019$ and without USA: $r^2 = 0.058$, $P = 0.210$). No patterns for genetic differentiation of microsatellite loci (assessed both by D_{est} and G_{st_est}) were evident (Fig. 3). Confidence intervals (data not shown) all excluded zero, indicating significant pairwise differentiation.

In the initial step of DAPC, 83 principal components of the PCA were retained as input to discriminant analysis, which accounted for more than 90% of the total variance. Based on the BIC, we chose models with 5–12 clusters to provide useful data summaries. Visual inspection of scatterplots identified the following $K = 4$ clusters (Fig. 5): (1) Norway, English Channel, Bay of Biscay; (2) Portugal north and Malaga; (3) Portugal south, Canaries, Guinea-Bissau, Ghana and USA; (4) Tangier and Senegal. Poland had an intermediate position, between group one and group two. The horizontal axis separated Tangier and Senegal from the rest, while the vertical axis set northern locations (Norway, Bay of Biscay, English Channel and Poland) apart from more southern sites (Portugal south, Canaries, Guinea-Bissau and Ghana) and USA. The mean assignment rate for each individual to the correct genetic cluster was 78%. However, after the cross-validation, the assignment power dropped to 31%.

DISCUSSION

The distribution of genetic diversity and isolation patterns among populations of the European anchovy indicates post-gla-

© 2014 John Wiley & Sons Ltd

cial dispersals both to the western and north-eastern Atlantic. Based on our results we suggest that anchovies inhabiting the north-western Atlantic, nominally designated as *E. eurystole*, are conspecific with *E. encrasicolus* and do not merit species status. These anchovies represent an extension of the range of European anchovies across the Atlantic. The north-eastern and north-western Atlantic may have been rapidly recolonized since the LGM by large numbers of anchovies, such that leading-edge populations retained the genetic diversity of parent populations.

Genetic population structure

In the present work, population genetic structure of North Atlantic anchovies was not concordant between mtDNA and nuclear microsatellites. Mitochondrial DNA displays a deep divergence between two clades, with little apparent geographical structure, apart from the clinal frequencies of clades, while microsatellites provide evidence of the existence of four main clusters (Norway–Bay of Biscay–English Channel; Portugal north–Malaga; Portugal south–Canaries–Guinea-Bissau–Ghana–USA; Tangier–Senegal; Poland occupies an intermediate position between the first and the second group). These results were consistent with other studies, so far as the locations of samples overlapped with those in our work (e.g. Magoulas *et al.*, 2006; Bouchenak-Khelladi *et al.*, 2008; Sanz *et al.*, 2008; Zarraonaindia *et al.*, 2012). The distributions of mtDNA, allozyme, microsatellite and nuclear gene markers in the various studies consistently showed a complex population structure for European anchovies in the Atlantic Ocean and Mediterranean Sea.

The inclusion of fish in non-spawning condition from 8 out of the 13 sample locations requires the potential caveat of underestimating population differentiation, as transient migrants, not representative of the assumed local population origins, may have been sampled. Indeed, we detected some degree of connectivity between populations (e.g. Norway, English Channel and Bay of Biscay) that possibly reflects ephemeral migrants between those locations. Both marker types showed clear differences in population structure, with mtDNA displaying a strong geographical cline for two clades. This is in contrast to microsatellite-based inferences of four genetic clusters corresponding loosely to large geographical regions. Combining mtDNA and microsatellites markers has the benefit of simultaneously assessing both historical and contemporary imprints on population structure (Limborg *et al.*, 2012). However, besides the effect of differences in mutation rates, we cannot eliminate other possible causes for mitonuclear discrepancies, such as sex-biased dispersal or selection.

Recent colonization of northern European seas

Many of the north-eastern Atlantic mtDNA haplotypes are shared with southern locations (Figs 1a & 2) and microsatellites also show a close relationship between northern populations and those inhabiting the Bay of Biscay (Fig. 5). Because north-eastern Atlantic areas would have been uninhabitable by anchovies during LGM conditions, our data are best explained by a southern origin of recently established northern populations. The geographical range expansion to the northern Atlantic Ocean must have occurred post-LGM, after the retreats of the British and the Fennoscandian ice sheets, because the minimum temperature presently tolerated by the European anchovy is 0–2 °C (Alheit *et al.*, 2012) (Fig. 1b). During the LGM this temperature range could only be found further south, in the English Channel (de Vernal & Hillaire-Marcel, 2006) (Fig. 1c). Besides temperature, the abundance of anchovies is also dependent on local productivity, and during the LGM the nearshore areas of the North Atlantic experienced significantly reduced productivity (de Vernal & Hillaire-Marcel, 2006), which could have also had an impact on the persistence of the species in that region.

We detected a clear signal of demographic expansion for clade A (Fig. 4a,c). Yet, for a population expansion of clade A to be more recent than the LGM, the mutation rate of cyt *b* in European anchovy would have to be at least 8×10^{-8}. Calibrated mutation rates for the cyt *b* locus in most marine taxa (Bowen *et al.*, 2001; Lessios, 2008) and other engraulids (Grant *et al.*, 2012) are closer to 2×10^{-8}. Therefore, population expansion in clade A is probably much older than the LGM. The signal of demographic expansion is less pronounced in clade B. However, like clade A, much of the genetic diversity observed in clade B is most likely older than the LGM (Fig. 4b,d) and there is no evidence to suggest a population bottleneck or pronounced demographic expansion during the last 20 kyr. Therefore, this species appears to have been able to track its thermal habitat preferences along coastlines and shift its geographical range latitudinally during recent climatic oscillations, without experiencing population bottlenecks or severe losses of genetic diversity, as has been the case in other species. For example, in Pacific herring (*Clupea pallasii*) Bayesian skyline plots display rapid population growth after the LGM and a flat population history during previous climatic fluctuations, suggesting the erosion of genetic signals from prior climate-related population disturbances (Grant *et al.*, 2012). In the European anchovy the signal of relatively ancient expansions is still visible and has not been eroded during more recent climatological oscillations.

The European anchovy does not display genetic signatures of recent, pronounced population contractions or expansions, particularly in the northern range of the distribution, where those effects would be most expected. This species displays high mtDNA and microsatellite diversity levels at all locations at its distributional margins. This observation probably reflects mass dispersal at the colonization front. In fact, there are records of swift appearances of anchovies in the North Sea during the 1990s, after a 30-year absence, and more recently the distribution range has extended to the Shetland Islands, southern Norway and into the Baltic Sea (Petitgas *et al.*, 2012). Fast colonizations of northern regions during favourable conditions are therefore plausible events and the dispersal capacities are pivotal for preserving the species genetic diversity under range contractions or shifts (Arenas *et al.*,

© 2014 John Wiley & Sons Ltd

2011). We posit that European anchovies are capable of fast-tracking optimal habitat conditions, and their swift mass movements prevent the loss of genetic diversity that would otherwise result from climate-related habitat disturbances, such as occurred during the LGM.

Western North Atlantic anchovies

Our results show that populations of the putative species *E. eurystole* in the north-western Atlantic belong to the same clade A observed in European anchovies. The cyt *b* haplotype network places most clade A haplotypes of *E. eurystole* in the European subclades, with one haplotype shared between the eastern Atlantic (Canary Islands) and the USA. Moreover, divergences between western and eastern Atlantic locations are smaller than among eastern Atlantic locations. Depending on effective population size and mutation rate, populations persisting in the western Atlantic over one or more glacial cycles might be expected to show considerable divergence from European source populations and to have deeply coalescing haplotype genealogies consistent with long persistence (Wares & Cunningham, 2001). Yet, as this is not the case, not only does *E. eurystole* appear to be conspecific with *E. encrasicolus* (virtually no differences in morphology; Whitehead *et al.*, 1988) but north-western Atlantic populations are likely to have been derived from the eastern Atlantic after the LGM.

We suggest that north-western Atlantic populations originated from genetically diverse sources (putatively western and central African populations), producing a non-monophyletic population in the USA, which displays similar divergences from all other locations. Furthermore, divergence is probably recent, given the close haplotype relationship with the eastern counterparts (Fig. 2). Our results are most consistent with a massive wave of colonists, comparable to those moving into northern European waters in response to sea surface temperature increases. Given that we estimated a high genetic diversity in putative source populations, the diversity in the western Atlantic could be the result of a single colonization event and not necessarily the outcome of multiple colonizations.

Species with amphi-Atlantic distributions imply present or historical trans-Atlantic migration(s) within periods of favourable ecophysiological conditions. Western Atlantic anchovies may have reached the Americas along one of two possible current-mediated dispersal routes, the 'northern' route or the 'equatorial' route. Opportunities for dispersal across the North Atlantic are likely to be limited to interglacial periods, when warming allows the stepping-stone colonizations of mid-North Atlantic islands that are typical of east-to-west dispersals (Vermeij, 2005). North Atlantic warming provided the conditions required for anchovy spawning (14 °C: Motos, 1996) and larval growth (16 °C: Urtizberea *et al.*, 2008) in the waters around the Shetland Islands, which may have acted as stepping-stone areas. However, the north-eastern Atlantic samples consist mostly of clade B individuals, contrasting with the western Atlantic, which is composed entirely of clade A. Moreover, European anchovies have never been recorded in the Faroe Islands, Iceland or Greenland, and the differentiation between the north-eastern and the western Atlantic and the lack of clade B individuals renders a 'northern' route unlikely.

Alternatively, an 'equatorial' route implies that anchovy populations off West and Central Africa have reached the Americas with the North Equatorial Current (NEC), running from West Africa to north-eastern South America. The NEC flows presently at a maximum speed of 15 cm s^{-1}, meaning that by purely passive drift, anchovies would take *c.* 200 days, probably much less if adult swimming is taken into account, to cross over the 2500 km from Guinea to Brazil continental platforms. A closer genetic relationship between western Atlantic and West Africa locations and the shared haplotype between the western Atlantic and Canary Islands further support the hypothesis of an 'equatorial' dispersal route. Dispersals from African populations in which clade A haplotypes predominate preclude the need to invoke clade sorting or founder effects to explain the absence of clade B haplotypes in north-western samples.

After crossing the Atlantic, anchovies could have followed northward currents (e.g. the Caribbean, Antilles and Florida and the Gulf Stream) through the Caribbean Sea to higher latitudes, up to their temperature tolerances. Apparently, high temperatures would have not limited dispersals of anchovies in these regions because temperatures over the last 100 ka were never higher than today (Emiliani, 1966). Therefore, anchovies could have crossed this hypothetical barrier at any time, not being restricted to specific periods. However, anchovy dispersals might be favoured during cooling periods by the enhancement of tropical currents (Fratantoni *et al.*, 2000). These fish may have found more suitable conditions to cross the Atlantic Ocean and the Caribbean Sea after the LGM, more likely during the Dryas stadials (18–15 ka; 14–13.7 ka; 12.8–11.5 ka) (Roberts, 1998) or during cooling periods in the late Holocene, from after the Hypsithermal period to the Little Ice Age (*c.* 0.5 ka; Denton & Karlén, 1973).

Examples of trans-Atlantic dispersal events include west-to-east migrations (e.g. the rock hind *Epinephelus adscensionis*; Carlin *et al.*, 2003) and east-to-west dispersals observed in four tropical fish, *Acanthurus monroviae* and *Parablennius pilicornis* (Luiz-Júnior *et al.*, 2004), *Epinephelus marginatus* (Joyeux *et al.*, 2001) and *Aulostomus strigosus* (Bowen *et al.*, 2001). However, most of the colonization routes proposed are fairly uncontroversial as these species have latitudinally narrower geographical distributions. For instance, *Clupea harengus*, with a geographical distribution restricted to the northern Atlantic Ocean, exhibits no genetic differentiation between the north-western Atlantic (Nova Scotia) and north-eastern Atlantic (North Sea) (Bekkevold *et al.*, 2005). Although no attempt was made to propose a colonization route to or from either side of the Atlantic, a hypothetical route for *C. harengus* would certainly include a northern route. The difference with anchovies is that although we would expect the colonization route to be mostly along the continental platform, the evidence points instead to a southern equatorial route.

© 2014 John Wiley & Sons Ltd

CONCLUSIONS

The results from this study provide valuable insights into the dynamics of anchovy population expansions and dispersals in response to ocean-climatic fluctuations in the Atlantic Ocean. Alternating cooling and warming periods in the North Atlantic over the course of the Quaternary might have displaced the ranges of these populations and set the stage for dispersal and recolonization. We found similar allele frequencies on either side of the English Channel and no significant decrease in genetic diversity in populations north of the English Channel. As anchovies are coastal species, climatic fluctuations were expected to enhance latitudinal range movements, but not longitudinal migrations across the Atlantic. However, climatic oscillations affected water mass circulation and trade winds, and specific cooling periods may have favoured these trans-Atlantic migrations. Whereas pre-LGM trans-Atlantic colonizations cannot be excluded, results are more compatible with post-LGM colonizations. Colonizations of the European anchovy in the northern part of the East Atlantic are unlikely to reflect first-time occupations, but instead represent the last of a series of range displacement cycles that reflect ongoing gene flow with southern core populations that have prevented divergence. The case of European anchovies contributes to our knowledge of how climate cycles conditioned species range contractions and expansions and show that the genetic characteristics of marginal populations, relative to central populations, can be determined by the dispersal dynamics of species.

ACKNOWLEDGEMENTS

This work was only possible due to the collaboration of many colleagues in collecting and shipping samples: E. Torstensen (IMR, Norway), M. Wyszynski (SFI, Poland), S. Vaz, F. Coppin, Y. Verin and D. Leroy (IFREMER, France), I. Zarraonaindia (UPV/EHU, Spain), K. Erzini (CCMAR, Portugal), A. Dias (Portugal), J. Tornero and J. Alarcón (IEO, Spain), M. Santamaria (Canary Islands), A. Hervé (Senegal), L. Fonseca (Guinea-Bissau), H. Dankwa (WRI, Ghana), J. Vasslides and R. Hagan (RUMFS, USA) and E. Wiley and A. Bentley (KNHM, USA). F. Lima produced Fig. 1b and c. We thank W. S. Grant for discussion and comment on earlier versions of this manuscript, N. Coelho and M. Valente for technical help, A. Yu for sharing unpublished microsatellites, J. Neiva, R. L. Cunha, S. Teixeira, K. Nicastro and G. Zardi for discussions, Nina O. Therkildsen for sharing the DAPC code, and Christine Maggs, M. Limborg and several anonymous referees for valuable comments and suggestions. This study was funded in part by the ESF MarinERA-NET-project/189570 'Marine phylogeographic structuring during climate change: the signature of leading and rear edge of range shifting populations'. G.S.'s research was supported by a doctoral grant from FCT-Portugal (SFRH/BD/36600/2007).

REFERENCES

Agostini, V.N. & Bakun, A. (2002) "Ocean Triads" in the Mediterranean Sea: physical mechanisms potentially structuring reproductive habitat suitability (example application to European anchovy, *Engraulis encrasicolus*). *Fisheries Oceanography*, **11**, 129–142.

Alheit, J., Pohlmann, T., Casini, M., Greve, W., Hinrichs, R., Mathis, M., O'Driscoll, K., Vorberg, R. & Wagner, C. (2012) Climate variability drives anchovies and sardines into the North and Baltic Seas. *Progress in Oceanography*, **96**, 128–139.

Arenas, M., Ray, N., Currat, M. & Excoffier, L. (2011) Consequences of range contractions and range shifts on molecular diversity. *Molecular Biology and Evolution*, **29**, 207–218.

Bekkevold, D., Andre, C., Dahlgren, T.G., Clausen, L.A.W., Torstensen, E., Mosegaard, H., Carvalho, G.R., Christensen, T.B., Norlinder, E. & Ruzzante, D.E. (2005) Environmental correlates of population differentiation in Atlantic herring. *Evolution*, **59**, 2656–2668.

Borrell, Y.J., Piñera, J.A., Sánchez Prado, J.A. & Blanco, G. (2012) Mitochondrial DNA and microsatellite genetic differentiation in the European anchovy *Engraulis encrasicolus* L. *ICES Journal of Marine Science: Journal du Conseil*, **69**, 1357–1371.

Bouchenak-Khelladi, Y., Durand, J.D., Magoulas, A. & Borsa, P. (2008) Geographic structure of European anchovy: a nuclear-DNA study. *Journal of Sea Research*, **59**, 269–278.

Bowen, B.W., Bass, A.L., Rocha, L.A., Grant, W.S. & Robertson, D.R. (2001) Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global scale. *Evolution*, **55**, 1029–1039.

Briers (2003) *VBA add-in Mantel.xla 1.2.4*. Available at: http://www.lifesciences.napier.ac.uk/staff/rob/software/index.html (accessed August 2012).

Carlin, J.L., Robertson, D.R. & Bowen, B.W. (2003) Ancient divergences and recent connections in two tropical Atlantic reef fishes *Epinephelus adscensionis* and *Rypticus saponaceous* (Percoidei: Serranidae). *Marine Biology*, **143**, 1057–1069.

Chapuis, M.-P. & Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

CLIMAP (1981) *Seasonal reconstruction of the Earth's surface at the last glacial maximum*. Geological Society of America Map Chart Series, Geological Society of America, Boulder, CO.

Denton, G.H. & Karlén, W. (1973) Holocene climatic variations – their pattern and possible cause. *Quaternary Research*, **3**, 155–205.

Drummond, A., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T. & Wilson, A. (2011) *Geneious v. 5.1*. Available at: http://www.geneious.com/.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Emiliani, C. (1966) Isotopic paleotemperatures. *Science*, **154**, 851–857.

Excoffier, L. & Lischer, H.E.L. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics

© 2014 John Wiley & Sons Ltd

analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Fratantoni, D.M., Johns, W.E., Townsend, T.L. & Hurlburt, H.E. (2000) Low-latitude circulation and mass transport pathways in a model of the tropical Atlantic Ocean. *Journal of Physical Oceanography*, **30**, 1944–1966.

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

Goudet, J. (1995) FSTAT version 1.2: a computer program to calculate F-statistics. *The Journal of Heredity*, **86**, 485–486.

Grant, W.S., Lecomte, F. & Bowen, B.W. (2010) Biogeographical contingency and the evolution of tropical anchovies (genus *Cetengraulis*) from temperate anchovies (genus *Engraulis*). *Journal of Biogeography*, **37**, 1352–1362.

Grant, W.S., Liu, M., Gao, T. & Yanagimoto, T. (2012) Limits of Bayesian skyline plot analysis of mtDNA sequences to infer historical demographies in Pacific herring (and other species). *Molecular Phylogenetics and Evolution*, **65**, 203–212.

Harpending, H. (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Hedrick, P.W. & Goodnight, C. (2005) *A standardized genetic differentiation measure. Evolution*, **59**, 1633–1638.

Jombart, T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart, T. (2013) *An introduction to adegenet 1.4-0*. Available at: ftp://ie.freshrpms.net/pub/CRAN/web/packages/adegenet/vignettes/adegenet-basics.pdf.

Jombart, T., Devillard, S., Dufour, A.B. & Pontier, D. (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, **101**, 92–103.

Jost, L. (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Joyeux, J.C., Floeter, S.R., Ferreira, C.E.L. & Gasparini, J.L. (2001) Biogeography of tropical reef fishes: the South Atlantic puzzle. *Journal of Biogeography*, **28**, 831–841.

Keenan, K., McGinnity, P., Cross, T.F., Crozier, W.W. & Prodöhl, P.A. (2013) diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, **4**, 782–788.

Kelly, R.P., Oliver, T.A., Sivasundar, A. & Palumbi, S.R. (2010) A method for detecting population genetic structure in diverse, high gene-flow species. *Journal of Heredity*, **101**, 423–436.

Lambeck, K., Esat, T.M. & Potter, E.K. (2002) Links between climate and sea levels for the past three million years. *Nature*, **419**, 199–206.

Lessios, H.A. (2008) The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 63–91.

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

Limborg, M., Hanel, R., Debes, P., Ring, A., André, C., Tsigenopoulos, C. & Bekkevold, D. (2012) Imprints from genetic drift and mutation imply relative divergence times across marine transition zones in a pan-European small pelagic fish (*Sprattus sprattus*). *Heredity*, **109**, 96–107.

Luiz-Júnior, O.J., Floeter, S.R., Gasparini, J.L., Ferreira, C.E.L. & Wirtz, P. (2004) The occurrence of *Acanthurus monroviae* (Perciformes: Acanthuridae) in the south-western Atlantic, with comments on other eastern Atlantic reef fishes occurring in Brazil. *Journal of Fish Biology*, **65**, 1173–1179.

Maggs, C.A., Castilho, R., Foltz, D., Henzler, C., Jolly, M.T., Kelly, J., Olsen, J., Perez, K.E., Stam, W., Vainola, R., Viard, F. & Wares, J. (2008) Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology*, **89**, S108–S122.

Magoulas, A., Tsimenides, N. & Zouros, E. (1996) Mitochondrial DNA phylogeny and the reconstruction of the population history of a species: the case of the European anchovy (*Engraulis encrasicolus*). *Molecular Biology and Evolution*, **13**, 178–190.

Magoulas, A., Castilho, R., Caetano, S., Marcato, S. & Patarnello, T. (2006) Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (*Engraulis encrasicolus*). *Molecular Phylogenetics and Evolution*, **3**, 734–746.

McInerny, G.J., Turner, J.R.G., Wong, H.Y., Travis, J.M.J. & Benton, T.G. (2009) How range shifts induced by climate change affect neutral evolution. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1527–1534.

van der Meer, M.H., Jones, G.P., Hobbs, J.P.A. & van Herwerden, L. (2012) Historic hybridization and introgression between two iconic Australian anemonefish and contemporary patterns of population connectivity. *Ecology and Evolution*, **2**, 1592–1604.

Molfetti, É., Torres Vilaça, S., Georges, J.-Y., Plot, V., Delcroix, E., Le Scao, R., Lavergne, A., Barrioz, S., dos Santos, F.R. & de Thoisy, B. (2013) Recent demographic history and present fine-scale structure in the northwest Atlantic leatherback (*Dermochelys coriacea*) turtle population. *PLoS ONE*, **8**, e58061.

Motos, L. (1996) Reproductive biology and fecundity of the Bay of Biscay anchovy population (*Engraulis encrasicolus* L.). *Scientia Marina*, **60**, 195–207.

van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Parada, C., van der Lingen, C.D., Mullon, C. & Penven, P. (2003) Modelling the effect of buoyancy on the transport of anchovy (*Engraulis capensis*) eggs from spawning to nursery grounds in the southern Benguela: an IBM approach. *Fisheries Oceanography*, **12**, 170–184.

Pennings, P.S., Achenbach, A. & Foitzik, S. (2011) Similar evolutionary potentials in an obligate ant parasite and its two host species. *Journal of Evolutionary Biology*, **24**, 871–886.

© 2014 John Wiley & Sons Ltd

Petitgas, P., Alheit, J., Peck, M., Raab, K., Irigoien, X., Huret, M., van der Kooij, J., Pohlmann, T., Wagner, C., Zarraonaindia, I. & Dickey-Collas, M. (2012) Anchovy population expansion in the North Sea. *Marine Ecology Progress Series*, **444**, 1–13.

Posada, D. (2004) *Collapse 1.2: describing haplotypes from sequence alignments*. Available at: http://darwin.uvigo.es/software/collapse.html (accessed August 2012).

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure from multilocus genotype data. *Genetics*, **155**, 945–959.

Rambaut, A. & Drummond, A.J. (2007) *Tracer 1.5*. Available at: http://beast.bio.ed.ac.uk/Tracer.

Ramos-Onsins, S.E. & Rozas, J. (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Renssen, H. & Vandenberghe, J. (2003) Investigation of the relationship between permafrost distribution in NW Europe and extensive winter sea-ice cover in the North Atlantic Ocean during the cold phases of the Last Glaciation. *Quaternary Science Reviews*, **22**, 209–223.

Roberts, N. (1998) *The Holocene: an environmental history*. John Wiley & Sons, Malden, MA.

Rogers, A. & Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Salicru, M., Menendez, M.L., Morales, D. & Pardo, L. (1993) Asymptotic distribution of (h, ϕ)-entropies. *Communications in Statistics – Theory and Methods*, **22**, 2015–2031.

Sambrook, J. & Russell, D. (2001) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, New York.

Sanz, N., Garcia-Marín, J.L., Viñas, J., Roldán, M. & Pla, C. (2008) Spawning groups of European anchovy: population structure and management implications. *ICES Journal of Marine Science: Journal du Conseil*, **65**, 1635–1644.

Teacher, A.G.F. & Griffiths, D.J. (2011) HapStar: automated haplotype network layout and visualization. *Molecular Ecology Resources*, **11**, 151–153.

Toonen, R. & Hughes, S. (2001) Increased throughput for fragment analysis on ABI Prism 377 automated sequencer using a membrane comb and STRand software. *BioTechniques*, **31**, 1320–1324.

Urtizberea, A., Fiksen, Ø., Folkvord, A. & Irigoien, X. (2008) Modelling growth of larval anchovies including diel feeding patterns, temperature and body size. *Journal of Plankton Research*, **30**, 1369–1383.

Vermeij, G.J. (2005) From Europe to America: Pliocene to recent trans-Atlantic expansion of cold-water North Atlantic molluscs. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2545–2550.

de Vernal, A. & Hillaire-Marcel, C. (2006) Provincialism in trends and high frequency changes in the northwest North Atlantic during the Holocene. *Global and Planetary Change*, **54**, 263–290.

Wares, J.P. & Cunningham, C.W. (2001) Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution*, **55**, 2455–2469.

Weir, B.S. (1996) *Genetic data analysis II*. Sinauer Associates, Sinauer Associates, Sunderland, MA.

Whitehead, P.J., Nelson, W.S. & Wongratana, T. (1988) *Clupeoid fishes of the world (suborder Clupeoidei)*. Food and Agriculture Organization of the United Nations, Rome.

Zarraonaindia, I., Pardo, M.A., Iriondo, M., Manzano, C. & Estonba, A. (2009) Microsatellite variability in European anchovy (*Engraulis encrasicolus*) calls for further investigation of its genetic structure and biogeography. *ICES Journal of Marine Science: Journal du Conseil*, **66**, 2176–2182.

Zarraonaindia, I., Iriondo, M., Albaina, A., Pardo, M.A., Manzano, C., Grant, W.S., Irigoien, X. & Estonba, A. (2012) Multiple SNP markers reveal fine-scale population and deep phylogeographic structure in European anchovy (*Engraulis encrasicolus* L.). *PLoS ONE*, **7**, e42201.

SUPPORTING INFORMATION

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

Appendix S1 Primer sequences, polymerase chain reaction conditions, sources and summary statistics of molecular markers.

Appendix S2 Plot of diversity measures along the latitudinal gradient.

Appendix S3 *Engraulis encrasicolus* microsatellite allele frequencies for each locus.

BIOSKETCHES

Gonçalo Silva is a PhD student at CCMAR (University of Algarve). This research was conducted as part of his dissertation, which investigates the phylogeography, population genetic structure, and genetic variation of the anchovies *Engraulis* spp.

John B. Horne is a post-doctoral fellow at CCMAR focusing on the study of evolutionary ecology and population biology of marine organisms, including phylogeography, biogeography, phylogenetics, conservation genetics and ecological genomics.

Rita Castilho is an assistant professor at the University of Algarve, and a researcher at CCMAR, and leads a team with broad interests in phylogeography and population genetics of marine and terrestrial vertebrate and invertebrate species.

Editor: Christine Maggs

© 2014 John Wiley & Sons Ltd