Journal of Experimental Marine Biology and Ecology 404 (2011) 108–115

Contents lists available at ScienceDirect

Journal of Experimental Marine Biology and Ecology

journal homepage: www.elsevier.com/locate/jembe

Comparative phylogeography and connectivity of sibling species of the marine copepod *Clausocalanus* (Calanoida)

Leocadio Blanco-Bercial [a,b,*], Florentina Álvarez-Marqués [b], Ann Bucklin [a]

[a] Department of Marine Sciences, University of Connecticut, 1080 Shennecossett Road, Groton, 06340 CT, USA
[b] Departamento de Biología de Organismos y Sistemas, Universidad de Oviedo, C/Catedrático Rodrigo Uría, s/n, 33071 Oviedo, Asturias, Spain

ARTICLE INFO

Article history:
Received 4 February 2011
Received in revised form 12 May 2011
Accepted 13 May 2011
Available online 7 June 2011

Keywords:
Clausocalanus
Connectivity
Global phylogeography
Isthmus of Panama
mtDNA

ABSTRACT

Connectivity of the marine epipelagic environment is subject to presence of subtle barriers that can be difficult to identify and to signals from the geological history of the oceans. This study examines the effects of species' geographical distribution on their population structure as mediated by differential effects of the recent geological history of the oceans. For this purpose, we studied the sequence variation of the mitochondrial cytochrome oxidase *c* subunit I (COI) gene in samples of two sibling species of the calanoid copepod genus *Clausocalanus*. Analyses included molecular population genetic, phylogeographic, and phylogenetic approaches. The cosmopolitan *Clausocalanus arcuicornis* is shown to have a single panmictic population across this species' extensive geographic range, with sufficient gene flow – despite vast distances and geological and oceanographic barriers – to maintain genetic cohesion. In contrast, the biantitropical *Clausocalanus lividus* exhibits clear differentiation between Atlantic and Pacific Ocean populations, suggesting a vicariance process that started after the rise of the Isthmus of Panama.

1. Introduction

Connectivity among the geographic populations of a species is necessary to maintain genetic integrity throughout its distributional range. Barriers to gene flow in the marine pelagic environment are often not clearly identified (Cowen et al., 2007; Goetze, 2005; Knowlton, 2000) and the effect of these barriers will differ depending on the spatial distribution of the species. To examine evidence of barriers to gene flow and cohesion among pelagic populations in different ocean basins, the biogeography of species must be studied in relation to the geological history of the oceans (Bradbury et al., 2008; Hellberg, 2009). Comparative studies of species that overlap in both time and space can provide better understanding of the existence of boundaries and their variable effects (Avise, 2009; Hickerson et al., 2010; McGovern et al., 2010; Peijnenburg et al., 2005; Reece et al., 2010).

Heterogeneous approaches to phylogeographic analysis and dispersal estimates (e.g., tag-and-release) have yielded inconsistent results in the ocean (Cowen and Sponaugle, 2009). In particular, it is uncertain whether previous studies may accurately represent global patterns (Bradbury et al., 2008). The development of molecular markers allowed genetic studies in phylogeography (Avise, 1998, 2009) and yielded improved understanding in the marine environment (Hedgecock et al., 2007; Selkoe et al., 2008). Many studies have used mitochondrial DNA (mtDNA) genes as markers (Avise, 2009; Avise et al., 1987), although multigene analysis including also nuclear loci is a clear and desirable trend in phylogeographic studies (Avise, 2009; Brito and Edwards, 2009). Despite recent doubts about the adequacy of mtDNA for this purpose, these genes have proved to be useful tools to address recent – up to several million years ago – phylogeographic and speciation phenomena (Avise, 2009; Galtier et al., 2009; Knowlton, 2000; Lessios, 2008).

Many studies have focused on species with planktonic larval stages (Bradbury et al., 2008; Weersing and Toonen, 2009), with fewer studies on holoplanktonic taxa (e.g., Caudill and Bucklin, 2004; Eberl et al., 2007; Goetze and Ohman, 2010; Goodall-Copestake et al., 2010; Milligan et al., 2011; Peijnenburg et al., 2006; Unal et al., 2006).

To investigate the presence and effect of barriers to population connectivity and gene flow of marine holoplanktonic species, we chose the calanoid copepod genus *Clausocalanus* (Giesbrecht, 1888). This genus comprises thirteen geographically widespread, epiplanktonic species of small copepods (Frost and Fleminger, 1968) and includes dominant zooplankton taxa from low to high latitudes (Pakhomov and Perissinotto, 1997; Peralba et al., 2010; Peralba and Mazzocchi, 2004; Schnack-Schiel et al., 2010). Despite their overlapping distributions, the species are ecologically, biogeographically, and genetically distinct (Bucklin and Frost, 2009; Cornils et al., 2007; Peralba et al., 2010; Saiz and Calbet, 1999; Schnack-Schiel et al., 2010). This study focuses on two species, *C. lividus* and *C. arcuicornis*. Both species are epipelagic, with distributions in both shelf and open ocean waters (Frost and Fleminger, 1968; Peralba et al., 2010; Peralba and

* Corresponding author at: Department of Marine Sciences, University of Connecticut, 1080 Shennecossett Road, Groton, 06340 CT, USA. Tel.: +1 860 405 9027; fax: +1 860 405 9153.
E-mail addresses: leocadio@uconn.edu (L. Blanco-Bercial), fmarques@uniovi.es (F. Álvarez-Marqués), ann.bucklin@uconn.edu (A. Bucklin).

0022-0981/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jembe.2011.05.011

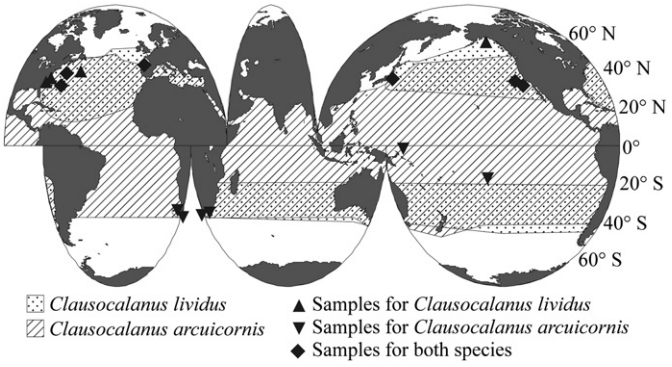

Fig. 1. Range of distribution and sample locations for *Clausocalanus lividus* and *Clausocalanus arcuicornis*.

Mazzocchi, 2004; Schnack-Schiel et al., 2010). The species show different geographical ranges (Fig. 1): *C. lividus* has a disjunct antitropical distribution, while *C. arcuicornis* is a cosmopolitan species (Frost and Fleminger, 1968) that is found in all ocean basins between 40° N and 40° S. Although Frost and Fleminger (1968) described the *C. lividus* distributional range as spanning the Pacific and Indian Oceans in the Southern Hemisphere, the species has not to our knowledge subsequently been reported in these regions (M.D. Viñas, INIDEP, Argentina; P. Ayón, IMARPE, Peru; R. Escribano, Universidad de Concepcion, Chile; D. McKinnon, AIMS, and F. Coman, CSIRO, Australia; pers. comm.).

In this work, we examine DNA sequence variation of the mitochondrial cytochrome oxidase *c* subunit I (COI) gene in samples of two species of copepods collected throughout their distributional ranges. We present here our findings of the comparative effects of barriers to gene flow and recent geological history of the oceans on the phylogeography and connectivity of these epipelagic zooplankton species.

2. Methods

2.1. Samples

A total of 87 individuals for *C. lividus* and 96 for *C. arcuicornis* were analyzed, including samples from cruises in the Atlantic, Pacific and Indian Oceans (Fig. 1; Table 1). Samples were preserved in 95% ethanol, which was changed after 24 h. The material analyzed for this study included samples stored in the Census of Marine Zooplankton (CMarZ) archives located at the Department of Marine Sciences, University of Connecticut.

2.2. DNA extraction, PCR amplification and sequencing

DNA extractions were carried out using the DNeasy Blood & Tissue Kit (QIAGEN) from whole individuals and with an elution volume of 50 to 300 μl AE buffer. PCR amplifications for a fragment of COI were performed in a total volume of 25 μl, including 5 μl of 5× Green GoTaq® Flexi Buffer, 2.5 μl of 25 mM $MgCl_2$, 1 μl of dNTPs (final concentration 0.2 mM each), 1 μl of each primer (10 μM), 0.75 units of GoTaq® Flexi DNA Polymerase (Promega) and 3 μl of DNA sample. Two primer sets were used: the consensus primers LCO1490 and HCO2198 (Folmer et al., 1994) and the specific primers for *Clausocalanus* spp., COI 5′ GAGCCTGGTCAGGAATAATCG 3′ (forward) and 5′ GGTCTCCTCCTCCTCCAACAT 3′ (reverse) (Blanco-Bercial and Álvarez-Marqués, 2007). The PCR protocol included an initial denaturation step of 94 °C (4 min), followed by denaturation at 94 °C for 35 s, annealing at 53 °C (*Clausocalanus* primers) or 45 °C (1490F/2198R) for 45 s, and extension at 69 °C for 45 s for 35 cycles. A final extension phase at 69 °C for 20 min was followed by storage at 4 °C. PCR products were checked by electrophoresis on a 1% agarose/TBE gel; positive results were purified using UltraClean® PCR Clean-Up Kit (Mo Bio). The purified PCR products were sequenced using the same set of primers as in the original amplification and Big Dye Terminator Ver. 3.1 (Applied Biosystems Inc., ABI), and run on an ABI 3130 Genetic Analyzer capillary DNA

Table 1
Location, number of individuals (N), collection, date, geographical coordinates, and grouping pattern for the AMOVA pool and for coalescence (Coals., *C. lividus*) and isolation with migration (IwM, *C. arcuicornis*) analyses for all samples.

Clausocalanus lividus

Location	N	Cruise	Date	Latitude	Longitude	Pool	Coals.
Alaska	9	Seward Line	Sep 2005	58°09.65′ N	147°47.60′ W	Alk	Pacific
California Current	11	RR-9610	Oct 1996	32°25.05′ N	123°08.50′ W	CC1	Pacific
	23	RR-9610	Oct 1996	32°20.00′ N	123°18.70′ W	CC2	Pacific
Japan	3	KT 09-4	Apr 2009	35°01.18′ N	139°22.04′ E	-	Pacific
Bay of Biscay	22	Radial Cudillero	Feb 2004	43°42.00′ N	06°09.00′ W	NEA	Atlantic
NW Atlantic	6	RHB-0603	Apr 2006	33°31.47′ N	69°57.68′ W	NWA	Atlantic
	4	DE-0808	Aug 2008	38°16.30′ N	74°24.40′ W	NWA	Atlantic
	1	DE-0711	Oct 2007	38°26.00′ N	73°46.10′ W	NWA	Atlantic
	6	OCE-258	Apr 1993	40°45.80′ N	62°28.10′ W	NWA	Atlantic
	2	OCE-258	Apr 1993	39°46.30′ N	54°44.70′ W	NWA	Atlantic

Clausocalanus arcuicornis

Location	N	Cruise	Date	Latitude	Longitude	Pool	IwM
California Current	11	RR-9610	Oct 1996	32°25.05′ N	123°08.50′ W	CC	CC
	11	RR-9610	Oct 1996	32°20.00′ N	123°18.70′ W	CC	CC
Tahiti	21	S202A	Jan 2006	17°10.60′ S	150°24.50′ W	Tah	Tah
Central Pacific	5	EUC-FE-2006	Sep 2006	01°20.91′ N	155°59.71′ E	–	–
Japan	1	KT 09-4	Apr 2009	35°01.18′ N	139°22.04′ E	–	–
South Africa	4	SARP-OM-EN26/12	Dec 2005	34°13.74′ S	18°04.46′ E	SA	SA
	4	Pel. Sp. Biomass-225	Nov 2006	33°26.74′ S	17°48.63′ E	SA	SA
	8	Pel. Sp. Biomass-225	Nov 2006	34°43.68′ S	24°56.94′ E	SA	SA
	6	Pel. Sp. Biomass-225	Nov 2006	34°25.91′ S	25°59.01′ E	SA	SA
Bay Biscay	17	Radial Cudillero	Jul 2004	43°42.00′ N	06°09.00′ W	NEA	NAt
NW Atlantic	2	OCE-258	Apr 1993	40°45.80′ N	62°28.10′ W	NWA	NAt
	6	RHB-0603	Apr 2006	33°31.47′ N	69°57.68′ W	NWA	NAt

sequencer. Sequences were edited using MEGA Ver. 4.1 (Tamura et al., 2007) and aligned with the ClustalW (Thompson et al., 1994) included in MEGA. The aligned sequences were trimmed to a length of 465 base pairs for analysis.

2.3. Molecular diversity indices and population structure

Nucleotide (π) and haplotype (H_d) diversities and neutrality tests were calculated with DnaSP Ver. 5 (Librado and Rozas, 2009). Neutrality was tested with Tajima's *D* (Tajima, 1989), Fu's F_S (Fu, 1997) and R_2 (Ramos-Onsins and Rozas, 2002). It should be noted that R_2 performs better than the other two metrics with small sample sizes (Ramos-Onsins and Rozas, 2002). The significance of the tests was estimated through 10,000 coalescence simulations.

Parsimony haplotype networks were constructed with TCS Ver. 1.2.1 (Clement et al., 2000), in order to visualize the diversity and phylogenetic relationships among the different haplotypes and provide qualitative assessment of their geographic distributions. The best-fitting substitution model was determined with jModelTest (Posada, 2008). Once the appropriate model was selected, the estimated parameters and the most similar model from the available were set in Arlequin Ver. 3.5.1.2 (Excoffier and Lischer, 2010) to evaluate the population structure for each species. Samples that had fewer than 5 individuals and distant from others were excluded from these analyses. Samples from the NW Atlantic were pooled for analysis of both species. For *C. arcuicornis*, samples from the California Current (CC) were also pooled. Pairwise F_{ST} distances among the different population were calculated and tested for significance through 100,172 permutations, with $\alpha = 0.05$ under strict Bonferroni correction (Bonferroni, 1936). An hierarchical Analysis of MOlecular VAriance or AMOVA (Excoffier et al., 1992) was performed with different *a priori* groupings, which were tested based on the geographical distributions of the samples and the previously obtained F_{ST} distances. The significance of the Arlequin statistics for the variance corresponding to among groups (Φ_{CT}), among populations within groups (Φ_{SC}), and within populations (Φ_{ST}) was determined based on 100,172 permutations. Also, average pairwise distances (i.e., numbers of nucleotide differences) were calculated for within- and between-groups in MEGA Ver. 4.1 (Tamura et al., 2007).

2.4. Clausocalanus lividus coalescence analysis

Since sequences with substantial levels of saturation are not suitable for molecular clock estimates (Wilke et al., 2009), the degree of substitution saturation was examined prior to analysis using the test of Xia et al. (2003), as implemented in the package DAMBE (Xia and Lemey, 2009). Divergence times between the Atlantic and Pacific Ocean populations of *C. lividus* were estimated by haplotype coalescence with the BEAST package Ver. 1.5.4 (Drummond and Rambaut, 2007). Markov Chain Monte Carlo chains (Drummond et al., 2002) were run for 10,000,000 generations, sampling every 1000 generations after a burn-in period of 1000 sampled generations. The analysis was run under the strict molecular clock model and HKY + I substitution model, as obtained from jModelTest, and under two different partition designs: a) without partitioning the data; and b) with two partitions following the codon positions, 1 + 2 and 3. The tree prior distribution selected was of constant size, as was estimated from Tajima's *D*, Fu's F_S and R_2 results. The prior for proportion of invariable sites (p-inv) was set according to the value obtained from jModelTest, whereas priors for the rest of parameters were left as default. The posterior distribution was examined in Tracer Ver. 1.5 (Rambaut and Drummond, 2007). The final tree with divergence estimates was computed in TreeAnnotator Ver. 1.5.4, with a burn-in value of 1000 generations, and analyzed with FigTree Ver. 1.3.1 (http://tree.bio.ed.ac.uk/software/figtree). To our knowledge, no direct estimates of mutation rates for COI are available for copepods; thus, the time of the most recent common ancestor (tmrca) was estimated from a range of mutation rates corresponding to other related crustaceans, ranging from 1.2 to 1.5% per million year (Knowlton and Weigt, 1998; Lessios, 2008).

2.5. Clausocalanus arcuicornis isolation with migration analysis

For *C. arcuicornis*, coalescence analysis for isolation and migration was performed in order to complement the F_{ST} and AMOVA analyses. Coalescent genealogies can provide more realistic results for many biological systems than can summary statistics, such as F_{ST} (Kuhner, 2009). The migration rates among the different ocean regions [i.e., North Atlantic, South Atlantic/SW India Ocean boundary (South Africa), North Pacific (California Current) and South Pacific (Tahiti)] were estimated with the software IMa2 (Hey, 2010). The advantage of the Isolation with Migration model implemented in this software is that the only assumptions (in our case) are neutrality and no recombination within the locus, allowing non-equilibrium between genetic drift and migration (Hey and Nielsen, 2007; Nielsen and Wakeley, 2001). The samples from the Central Pacific and Japan were excluded from this analysis; the NW and NE Atlantic samples were pooled. Two parallel procedures were followed on the basis of the tree prior distribution. In the first case, an UPGMA tree was constructed using the Neighbor program of PHYLIP Ver. 3.69 (Felsenstein, 2004) based on a matrix of F_{ST} distances among the four groups established, using the same procedure as described for the population structure study. The second tree prior was based on coalescence methods and was constructed in BEAST following the same procedure as described for *C. lividus*, but under the HKY + I + Γ substitution model and with priors indicated by jModelTest. This approach helps recognize possible bias introduced by the tree prior probability distribution. The mutation model chosen for IMa2 was HKY, as recommended for mitochondrial data by the IMa2 documentation. Short runs were initially performed to empirically obtain a proper upper bound on the prior distribution for the population parameters (i.e., time of divergence, migration rates and population size). Once these parameter priors were established, long runs of the Markov chain Monte Carlo (MCMC) simulations were done; in each case, burn-in period was run until reaching the stationary state; then every 100th genealogy was sampled until at least 100,000 genealogies were saved. Only migration among sampled populations was considered (i.e., ancestral populations had migration rates of zero) under two different modeled scenarios: a) full migration, and, b) zero migration between North Atlantic and any Pacific Ocean population.

3. Results

3.1. Molecular diversity and neutrality

Analysis of the 465 base-pair fragment of COI identified 43 haplotypes for *C. arcuicornis* and 33 for *C. lividus* (GenBank accession nos. JF279610–JF279685). All the haplotypes were shown to code for the same protein based on the amino acid sequence.

Total and per-sample haplotype (H_d) and nucleotide (π) diversity are indicated in Table 2. When considering all individuals, *C. lividus* showed lower H_d but higher π values than *C. arcuicornis*. Both H_d and π were higher for *C. arcuicornis* when considering single samples or groups of samples from the same ocean basin (Table 2). Test results showed non-significant departures from neutrality for Tajima's *D*, Fu's F_S or R_2 for either species. The exception was Fu's F_S for *C. arcuicornis*, which showed a significant negative value ($F_S = -16.385$, $p < 0.005$), perhaps indicating a recent population expansion. However, for all analyses we assumed constant size for the demographic models.

The haplotype networks showed different patterns for the two species. *C. arcuicornis* showed a core with a large number of haplotypes separated by one to three substitutions, with a small

Table 2
Total and per-sample number of individuals (N) and haplotypes (No.H.), and haplotype (H_d) and nucleotide (π) diversities.

Species	Sample	N	No.H.	H_d	π
C. arcuicornis		**96**	**43**	**0.958**	**0.0180**
	South Africa	22	10	0.857	0.0057
	NW Atlantic	8	8	1.000	0.0127
	NE Atlantic	17	10	0.919	0.0326
	California Current	22	12	0.887	0.0174
	Tahiti	21	13	0.910	0.0101
	Equatorial Pacific	5	3	0.800	0.0138
C. lividus		**87**	**33**	**0.874**	**0.0337**
	NW Atlantic	19	11	0.877	0.0072
	NE Atlantic	22	8	0.732	0.0043
	California C1	11	5	0.618	0.0034
	California C2	23	8	0.581	0.0020
	Alaska	9	4	0.806	0.0032

number of very divergent haplotypes (Fig. 2). In contrast, *C. lividus* showed two cores corresponding to the Atlantic and Pacific Oceans (Fig. 3). The cores were separated by ~30 substitutions with each containing a group of very closely related haplotypes separated by one or two substitutions.

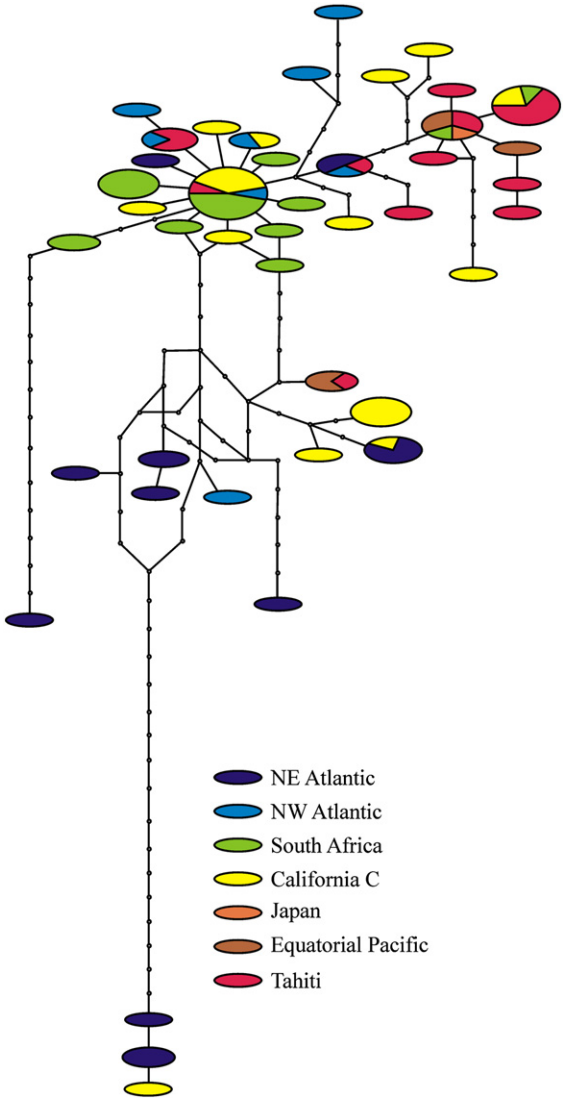

Fig. 2. Haplotype network of *Clausocalanus arcuicornis*. Each branch represents a mutational step, and small black circles, missing haplotypes needed to connect observed haplotypes. The size of the ovals is proportional to the number of sampled individuals with that haplotype.

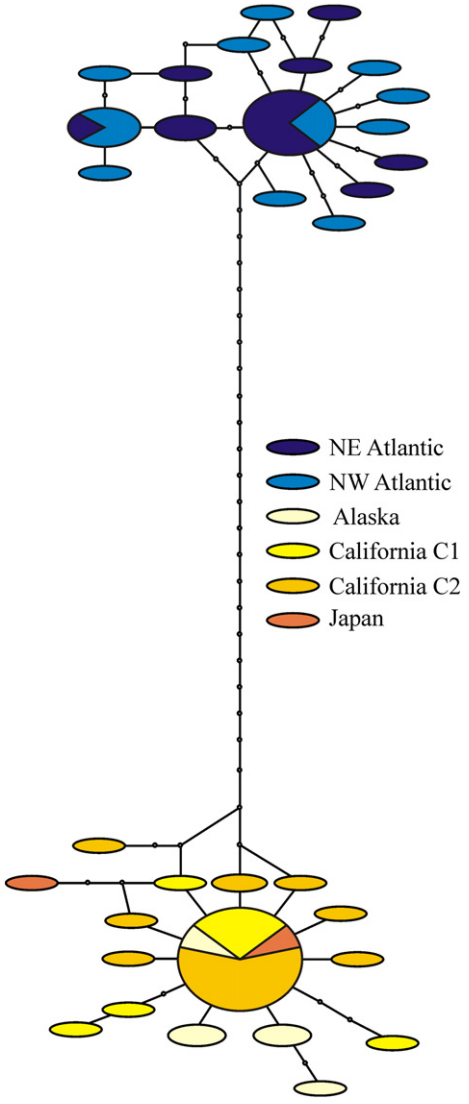

Fig. 3. Haplotype network of *Clausocalanus lividus*. Each branch represents a mutational step, and small black circles, missing haplotypes needed to connect observed haplotypes. The size of the ovals is proportional to the number of sampled individuals with that haplotype.

3.2. Population structure

The model of sequence evolution selected for both species using jModelTest under the Akaike Information Criterion (AIC) was the Tamura and Nei model (Tamura and Nei, 1993). For *C. arcuicornis*, the proportion of invariable sites was 0.406 with a gamma shape of 0.201 (+ I + G). For *C. lividus*, the gamma shape (+ G model) was 0.016. Under these models, *C. arcuicornis* F_{ST} values were relatively low (0.06 to 0.3) and significant for pairs of samples from different ocean basins, except the NW Atlantic/California value that was not significant (Table 3). In contrast, *C. lividus* showed high F_{ST} values between pairs of samples from different ocean basins, and differences between sample pairs from the same basin were much lower (Table 3). The largest values for with-in ocean comparisons were between Alaska and the two California Current samples, but both comparisons were non-significant after Bonferroni correction.

The *a priori* AMOVA grouping established for *C. arcuicornis* divided the samples into three groups, each corresponding to a major ocean basin: Atlantic Ocean (NW and NE samples), Pacific Ocean (CC and Tahiti), and Indian Ocean (South Africa). This configuration also agreed with the pairwise F_{ST} values (see above). The AMOVA results

Table 3
Pairwise F_{ST} between samples. Bold numbers indicate significant values after strict Bonferroni correction ($\alpha_B = 0.005$).

Species	Sample	SA	NWA	NEA	CC
C. arcuicornis	South Africa	–			
	NW Atlantic	**0.14352**	–		
	NE Atlantic	**0.30092**	0.14020	–	
	California C	**0.14348**	0.06181	**0.15510**	–
	Tahiti	**0.29597**	**0.22944**	**0.27472**	0.08799
		NWA	NEA	CC1	CC2
C. lividus	NW Atlantic	–			
	NE Atlantic	0.00000	–		
	California C1	**0.99999**	**0.99999**	–	
	California C2	**0.99999**	**0.99999**	0.01789	–
	Alaska	**0.99999**	**1.00000**	0.09299	0.25490

indicated that the among-group variation (Φ_{CT}) was the smallest component of the variance and not significant; the largest component of the variance corresponded to with-in populations (Φ_{ST}). The between populations with-in group variance (Φ_{SC}) was lower, but significant (Table 4). For *C. lividus*, the AMOVA grouping tested the Atlantic (NE and NW samples) versus the Pacific samples (Alaska, CC1 and CC2). The analysis showed that, although almost all of the variance was explained by differences between ocean basins (Φ_{CT}), it was not significant ($p = 0.098$). The differences between samples from the same ocean basin (Φ_{SC}) were low and not significant; variance explained by the within samples component (Φ_{ST}) was low, but significant.

Average number of nucleotide substitutions within and between the AMOVA groups (Table 5) was relatively low for *C. arcuicornis* in all cases (<3%). For *C. lividus*, nucleotide differences were very low for all within group comparison, but they were very high (>6%) when comparing individuals from the Atlantic to the Pacific Oceans.

3.3. Clausocalanus lividus coalescence analysis

The test of substitution in DAMBE showed little saturation, indicating that the data are suitable for this analysis. The BEAST analysis showed good mixing and effective sample sizes (ESS) higher than 1000 for every parameter. The analysis found two very divergent clades: the North Atlantic and North Pacific Oceans. The time of coalescence of the two clades (or the tmrca for the species) was determined to be 3.8–4.7 My (Million years) for the unpartitioned data set and 3.6–4.5 My for the codon-partition model.

3.4. Clausocalanus arcuicornis isolation with migration analysis

The trees obtained from the F_{ST} matrix (not shown) and the BEAST analysis comparing the different ocean sub basins were slightly different (Fig. 4). The F_{ST}-based tree showed a close relationship between the North Atlantic and the North Pacific (California Current), while the coalescent-based BEAST analysis indicated the same, except for the south hemisphere ones (South Africa and Tahiti). The IMa2 analysis with full migration pattern was not able to resolve the model of migration when considering the BEAST tree as phylogenetic tree prior (results not shown). Reduced migration models gave similar patterns for migration rates for every pair, although the range was narrower for the BEAST analysis (Table 6). When considering the full migration model under the F_{ST} tree prior conditions, the analysis was able to resolve the pattern. The resulting migration rates tended to reduce the differences among the rates, when compared to the reduced models results; the maximum rate corresponded to the migration from the North Atlantic to the California Current sample. For the reduced models, the tendency was much higher migration from Tahiti and South Africa toward the other locations than the reverse, a high rate from Tahiti to South Africa, and a very low rate from South Africa toward Tahiti (Table 6).

Table 4
Analysis of molecular variance (AMOVA) for *C. arcuicornis* and *C. lividus* (d.f.: degrees of freedom).

Source of variation	d.f.	Variance components	Φ statistics	p-value
Clausocalanus arcuicornis				
Among groups	2	0.3529	$\Phi_{CT} = 0.0628$	0.2657
Among populations within groups	2	0.8047	$\Phi_{SC} = 0.1528$	0.0079
Within populations	85	4.4618	$\Phi_{ST} = 0.2060$	<0.0001
Clausocalanus lividus				
Among groups	1	335193.1368	$\Phi_{CT} = 0.9999$	0.0985
Among populations within groups	3	0.0418	$\Phi_{SC} = 0.0180$	0.1998
Within populations	79	2.2929	$\Phi_{ST} = 0.9999$	<0.0001

Table 5
Percentage of pairwise distances (as nucleotide differences) between AMOVA groups (average ± standard deviation) for a 465 base pair region of COI.

Species	Sample	SA	Atl	Pac
C. arcuicornis	South Africa	0.57 ± 0.48		
	Atlantic	2.13 ± 1.70	2.85 ± 0.48	
	Pacific	1.25 ± 0.88	2.45 ± 1.75	1.47 ± 1.13
		Atl	Pac	
C. lividus	Atlantic	0.57 ± 0.39		
	Pacific	6.27 ± 0.27	0.28 ± 0.29	

4. Discussion

This study elucidates the effects of species' biogeographical distribution on population connectivity, genetic structure, and vicariance over recent geological history. Comparisons among congeneric sibling species are particularly useful to allow stronger inferences of the underlying biological and geological processes driving patterns of connectivity, since many aspects of species' ecology and evolution are held in common among such species. The cosmopolitan distribution of *C. arcuicornis* is shown to allow high connectivity among different ocean basins. In contrast, the disjunct antitropical distribution of *C. lividus* is hypothesized to be influenced by vicariance processes after the closure of the Isthmus of Panama and the loss of direct communication between the North Atlantic and North Pacific Oceans.

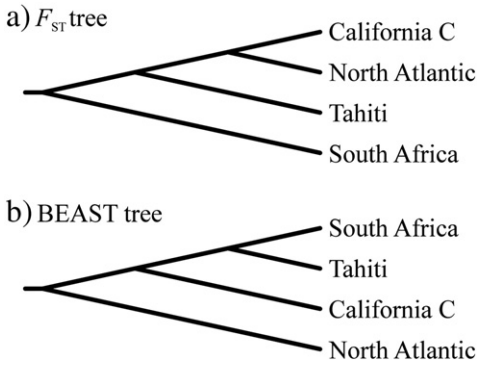

Fig. 4. Prior probability distribution F_{ST} based (a) and BEAST coalescence-based (b) trees for the isolation and migration analysis.

Table 6
Migration rates per generation between the North Atlantic (NAt), California Current (CC), Tahiti (Tah) and South Africa (SA) samples under the three models considered (see Methods and Results).

	NAt to CC	CC to NAt	Tah to NAt	NAt to Tah	SA to NAt	NAt to SA
Full – F_{ST}	0.7763	0.1762	0.0013	0.0013	0.0013	0.0013
Reduced						
F_{ST}					0.9350	0.0008
BEAST					0.7025	0.1380
	SA to CC	CC to SA	SA to Tah	Tah to SA	CC to Tah	Tah to CC
Full – F_{ST}	0.7238	0.0013	0.0088	0.0013	0.0013	0.5513
Reduced						
F_{ST}	1.9040	0.0010	0.0010	1.6420	0.0010	0.8712
BEAST	1.1920	0.0025	0.0025	0.6775	0.0025	0.2975

Although the analysis of multiple genes (especially nuclear) may have enhanced or provided additional insights from this study, especially for resolving the deep node between ocean basin clades of *C. lividus*, the single mitochondrial marker used provides sufficient and compelling support for the conclusions

C. lividus is shown to have a high number of closely-related haplotypes within each ocean basin population, which are very distant between different ocean basins. On the contrary, *C. arcuicornis* shares and mixes haplotypes from samples scattered along a more than 40,000 km route that crosses four ocean basins and many recognized biogeographical barriers. It was not possible to identify a single cluster of haplotypes that could be assigned to a single ocean basin. These differences between the two species can explain the mismatch between the global H_d and π. The high π values found for *C. lividus* reflects a high number of variable sites, but many of the variable sites are not shared across the whole range and the final H_d is lower than that obtained for *C. arcuicornis*, which has a lower π. Within each ocean basin, there is a pattern of lower H_d and π for *C. lividus* than for *C. arcuicornis*. The higher diversities of *C. arcuicornis* may also reflect the general connectivity through the whole distributional range of the species, resulting in larger effective population size, and thus higher number and diversity of haplotypes.

4.1. Clausocalanus lividus

Clausocalanus lividus is shown to comprise two genetically and geographically distinct populations that have been subjected to a vicariance process, which according to our results would have started between 4.7 and 3.6 Mya (Million years ago). This estimate matches with the interval between the depletion of the water flow through the Central American Seaway and reorganization of the general ocean circulation around 4.6 Mya (Haug and Tiedemann, 1998; Schneider and Schmittner, 2006) and the final and definitive rise of the Isthmus of Panama 3.5 Mya (Bartoli et al., 2005; Coates et al., 1992) (Fig. 5). These events contributed to the isolation and divergence of many marine species (Knowlton and Weigt, 1998; Lessios, 2008; Marko, 2002), and also to the Atlantic and Pacific populations of this neritic to open ocean epipelagic copepod. *C. lividus* results indicate that almost all the molecular variance is due to the differences between ocean basins. Within ocean basins, the F_{ST} distances were not significantly different from zero, but F_{ST} distances of ~1 between samples from different ocean basins supports their complete isolation (Hedgecock et al., 2007; Hellberg, 2009). In addition, the high divergences for *C. lividus* were slightly higher than values found between populations of other calanoid copepods (Bucklin et al., 2003; Hill et al., 2001) or sibling species of other crustaceans (Knowlton and Weigt, 1998) from the different ocean basins, but were much lower than the ~20% differences found between *Calanus* species occurring on either side of the Isthmus of Panama (Hill et al., 2001).

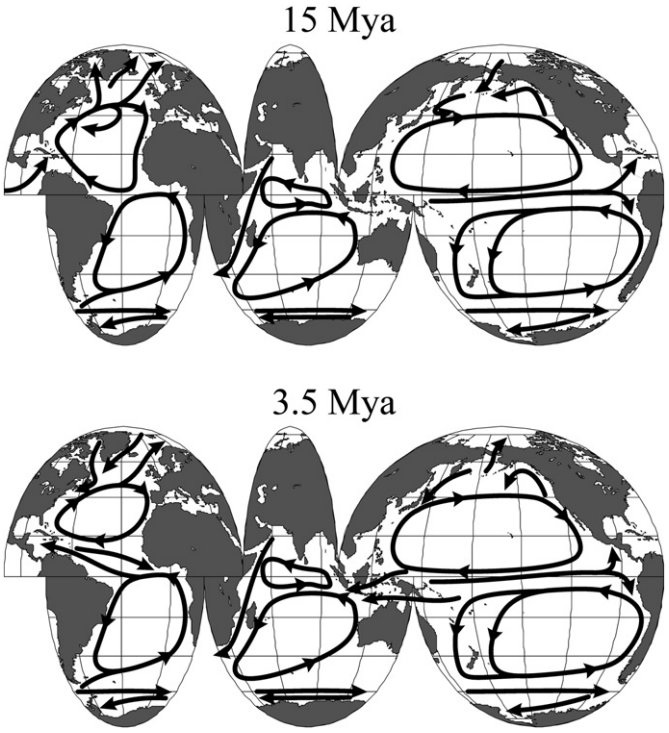

Fig. 5. Map of the ocean surface currents before and after the closure of the Isthmus of Panama (after Haug and Tiedemann, 1998; Luyendyk et al., 1972; Maier-Reimer et al., 1990; Schneider and Schmittner, 2006; Schweitzer, 2001).

4.2. Clausocalanus arcuicornis

Clausocalanus arcuicornis populations are shown to lack persistent and definitive barriers to gene flow between ocean basins, with widespread connectivity throughout the geographical range of the species. This pattern of genetic structure of the population is similar to the found for other marine copepods (Eberl et al., 2007; Goetze, 2005). The absence of significant differences between the NW Atlantic and the East Pacific (CC) and the high theoretical migration rates between North Atlantic and CC may reflect the past direct connection between the ocean basins. Migration of planktonic taxa can be considered to be primarily driven by ocean currents (Cowen and Sponaugle, 2009). Before the closure of the Central American Seaway, the current flow migration between basins would have been from the Pacific into the Atlantic through this passage. At that time, the weak flux from the Pacific into the Indian Ocean (Schneider and Schmittner, 2006) and from the North to South Atlantic (Maier-Reimer et al., 1990) might have resulted in the relative isolation of the Indian Ocean. This would explain the high proportion of individuals in the South Africa sample with private haplotypes, despite its low H_d and the high differentiation of the South Africa station compared to others. Water flow from the Pacific into the Indian Ocean would have increased after the closure of the Isthmus of Panama (Fig. 5), when the net flux from the Pacific into the North Atlantic was only through the Indonesian Archipelago and the Indian Ocean (Schneider and Schmittner, 2006). This would match with the relatively strong migration rates found from Tahiti to South Africa and from South Africa into the North Atlantic, when there was no flow of migrants between the Atlantic and the Pacific Oceans.

4.3. A historical view of pelagic biogeography

Before the closure of the Isthmus of Panama, the flow from the Pacific into the Atlantic would have maintained the connections between conspecific planktonic populations from the Atlantic and the Pacific Oceans through the Central American Seaway (Fig. 5). The

connection through the Southern Hemisphere via the Indonesian Archipelago would have been weaker. After 15 Mya, the processes that originated the Isthmus of Panama began, filling the gap completely about 3.5 Mya. But much earlier, around 4.6 Mya, the currents were definitively changed (Fig. 5). For epiplanktonic species without a cosmopolitan distribution, this would have meant the separation of the species' continuous range into two disjunct regions, with populations then becoming genetically isolated. The closure of the Central America Seaway also interrupted the direct contact between Atlantic and Pacific populations of cosmopolitan species. However, the connection for those species was maintained from the Pacific to the Indian and from the Indian to the Atlantic Oceans, with the enhanced flow through the Indonesian Archipelago, and the contact and genetic integrity of the species persisted throughout the global range.

Acknowledgements

We would like to acknowledge Janet Bradford-Grieve (National Institute of Water & Atmospheric Research, New Zealand), Ryuji Machida (The Smithsonian Institution, USA), David McKinnon (Australian Institute of Marine Science, Australia), Brian D. Ortman (The University of the British Columbia, Canada), Ebru Unal (University of Connecticut, USA), Hans M. Verheye (Marine and Coastal Management, South Africa) and Maria D. Viñas (Instituto Nacional de Investigación y Desarrollo Pesquero, Argentina) for providing or trying to provide samples for this study. We especially acknowledge Nancy J. Copley and Peter H. Wiebe (Woods Hole Oceanographic Institution, USA) for their support providing samples and cruise data. Bruce W. Frost (University of Washington, USA) offered insightful comments on early stages of the research. We gratefully acknowledge programmatic support from the Alfred P. Sloan Foundation. Additional support for this project was provided by the NOAA Office of Ocean Exploration (Award no. 24009100) and by the "Control a largo plazo de las condiciones químico-biológicas de la plataforma continental de Asturias, 2002–2006" contract (IEO - University of Oviedo). L.B.B. was partially supported by a PhD grant from the "III Plan Regional de I + D + I del Principado de Asturias" (FICYT). This study is a contribution from the Census of Marine Zooplankton (CMarZ, see www.CMarZ.org), an ocean realm field project of the Census of Marine Life. **[SS]**

References

Avise, J.C., 1998. The history and purview of phylogeography: a personal reflection. Mol. Ecol. 7, 371–379.
Avise, J.C., 2009. Phylogeography: retrospect and prospect. J. Biogeogr. 36, 3–15.
Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18, 489–522.
Bartoli, G., Sarnthein, M., Weinelt, M., Erlenkeuser, H., Garbe-Schönberg, D., Lea, D.W., 2005. Final closure of Panama and the onset of Northern Hemisphere Glaciation. Earth Planet. Sci. Let. 237, 33–44.
Blanco-Bercial, L., Álvarez-Marqués, F., 2007. RFLP procedure to discriminate between *Clausocalanus* Giesbrecht, 1888 (Copepoda, Calanoida) species in the Central Cantabrian Sea. J. Exp. Mar. Biol. Ecol. 344, 73–77.
Bonferroni, C.E., 1936. Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8, 3–62.
Bradbury, I.R., Laurel, B., Snelgrove, P.V.R., Bentzen, P., Campana, S.E., 2008. Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. Proc. R. Soc. B: Biol. Sci. 275, 1803–1809.
Brito, P., Edwards, S., 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica 135, 439–455.
Bucklin, A., Frost, B.W., 2009. Morphological and molecular phylogenetic analysis of evolutionary lineages within *Clausocalanus* (Copepoda: Calanoida). J. Crust. Biol. 29, 111–120.
Bucklin, A., Frost, B.W., Bradford-Grieve, J., Allen, L.D., Copley, N.J., 2003. Molecular systematic and phylogenetic assessment of 34 calanoid copepod species of the Calanidae and Clausocalanidae. Mar. Biol. 142, 333–343.
Caudill, C.C., Bucklin, A., 2004. Molecular phylogeography and evolutionary history of the estuarine copepod, *Acartia tonsa*, on the Northwest Atlantic coast. Hydrobiologia 511, 91–102.
Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1659.
Coates, A.G., Jackson, J.B.C., Collins, L.S., Cronin, T.M., Dowsett, H.J., Bybell, L.M., Jung, P., Obando, J.A., 1992. Closure of the Isthmus of Panama: the near-shore marine record of Costa Rica and western Panama. Geol. Soc. Am. Bull. 104, 814–828.
Cornils, A., Schnack-Schiel, S.B., Böer, M., Graeve, M., Struck, U., Al-Najjar, T., Richter, C., 2007. Feeding of Clausocalanids (Calanoida, Copepoda) on naturally occurring particles in the northern Gulf of Aqaba (Red Sea). Mar. Biol. 151, 1261–1274.
Cowen, R.K., Sponaugle, S., 2009. Larval dispersal and marine population connectivity. Annu. Rev. Mar. Sci. 1, 443–466.
Cowen, R.K., Gawarkiewicz, G., Pineda, J., Thorrold, S.R., Werner, F.E., 2007. Population connectivity in marine systems: an overview. Oceanography 20, 14–21.
Drummond, A., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Drummond, A.J., Nicholls, G.K., Rodrigo, A.G., Solomon, W., 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161, 1307–1320.
Eberl, R., Cohen, S., Cipriano, F., Carpenter, E.J., 2007. Genetic diversity of the pelagic harpacticoid copepod *Macrosetella gracilis* on colonies of the cyanobacterium *Trichodesmium* spp. Aquat. Biol. 1, 33–43.
Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 10, 564–567.
Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479–491.
Felsenstein, J., 2004. PHYLIP (Phylogeny Inference Package) version 3.69. Department of Genome Sciences, University of Washington, Seattle.
Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299.
Frost, B.W., Fleminger, A., 1968. A revision of the genus *Clausocalanus* (Copepoda: Calanoida) with remarks on distributional patterns in diagnostic characters. Bull. Scripps Inst. Oceanogr. Univ. Calif. 12, 1–235.
Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.
Galtier, N., Nabholz, B., Glémin, S., Hurst, G.D.D., 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18, 4541–4550.
Giesbrecht, W., 1888. Elenco dei Copepodi pelagici raccolti dal Tenete di vascello Gaetano Chierchia durante il viaggio della R. Corvetta "Vettor Pisani" negli anni 1882–1885 e dal Tenente di vascello Francesco Orsini nel Mar Rosso, nel 1884. Atti Acad. naz. Lincei Rd., Cl. Sci. fis. mat. nat., (4) 4 sem. 2, 284–287, 330–338.
Goetze, E., 2005. Global population genetic structure and biogeography of the oceanic copepods *Eucalanus hyalinus* and *E. spinifer*. Evolution 59, 2378–2398.
Goetze, E., Ohman, M.D., 2010. Integrated molecular and morphological biogeography of the calanoid copepod family Eucalanidae. Deep-Sea Res. II 57, 2110–2129.
Goodall-Copestake, W.P., Perez-Espona, S., Clark, M.S., Murphy, E.J., Seear, P.J., Tarling, G.A., 2010. Swarms of diversity at the gene *cox1* in Antarctic krill. Heredity 104, 513–518.
Haug, G.H., Tiedemann, R., 1998. Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. Nature 393, 673–676.
Hedgecock, D., Barber, P.H., Edmands, S., 2007. Genetic approaches to measuring connectivity. Oceanography 20, 70–79.
Hellberg, M.E., 2009. Gene flow and isolation among populations of marine animals. Annu. Rev. Ecol. Evol. Syst. 40, 291–310.
Hey, J., 2010. Isolation with migration models for more than two populations. Mol. Biol. Evol. 27, 905–920.
Hey, J., Nielsen, R., 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Acad. Sci. 104, 2785–2790.
Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F., Yoder, A.D., 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Mol. Phylogenet. Evol. 54, 291–301.
Hill, H.S., Allen, L.D., Bucklin, A., 2001. Multiplexed species-specific PCR protocol to discriminate four N. Atlantic *Calanus* species, with an mtCOI gene tree for ten *Calanus* species. Mar. Biol. 139, 279–287.
Knowlton, N., 2000. Molecular genetic analyses of species boundaries in the sea. Hydrobiologia 420, 73–90.
Knowlton, N., Weigt, L.A., 1998. New dates and new rates for divergence across the Isthmus of Panama. Proc. R. Soc. Lond. Ser. B: Biol. Sci. 265, 2257–2263.
Kuhner, M.K., 2009. Coalescent genealogy samplers: windows into population history. Trends Ecol. Evol. 24, 86–93.
Lessios, H.A., 2008. The Great American Schism: divergence of marine organisms after the rise of the Central American Isthmus. Annu. Rev. Ecol. Evol. Syst. 39, 63–91.
Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.
Luyendyk, B.P., Forsyth, D., Phillips, J.D., 1972. Experimental approach to the paleocirculation of the oceanic surface waters. Geol. Soc. Am. Bull. 83, 2649–2664.
Maier-Reimer, E., Mikolajewicz, U., Crowley, T., 1990. Ocean general circulation model sensitivity experiment with an open Central American Isthmus. Paleoceanography 5, 349–366.
Marko, P.B., 2002. Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. Mol. Biol. Evol. 19, 2005–2021.
McGovern, T.M., Keever, C.C., Saski, C.A., Hart, M.W., Marko, P.B., 2010. Divergence genetics analysis reveals historical population genetic processes leading to contrasting phylogeographic patterns in co-distributed species. Mol. Ecol. 19, 5043–5060.

Milligan, P., Stahl, E., Schizas, N., Turner, J., 2011. Phylogeography of the copepod *Acartia hudsonica* in estuaries of the northeastern United States. Hydrobiologia 666, 155–165.
Nielsen, R., Wakeley, J., 2001. Distinguishing migration from isolation: a Markov Chain Monte Carlo approach. Genetics 158, 885–896.
Pakhomov, E.A., Perissinotto, R., 1997. Mesozooplankton community structure and grazing impact in the region of the Subtropical Convergence south of Africa. J. Plankton Res. 19, 675–691.
Peijnenburg, K.T.C.A., van Haastrecht, E.K., Fauvelot, C., 2005. Present-day genetic composition suggests contrasting demographic histories of two dominant chaetognaths of the North-East Atlantic, *Sagitta elegans* and *S. setosa*. Mar. Biol. 147, 1279–1289.
Peijnenburg, K.T.C.A., Fauvelot, C., Breeuwer, J.A.J., Menken, S.B.J., 2006. Spatial and temporal genetic structure of the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. Mol. Ecol. 15, 3319–3338.
Peralba, A., Mazzocchi, M.G., 2004. Vertical and seasonal distribution of eight *Clausocalanus* species (Copepoda: Calanoida) in oligotrophic waters. ICES J. Mar. Sci. 61, 645–653.
Peralba, A., Mazzocchi, M.G., Harris, R.H., 2010. Niche separation of *Clausocalanus* spp. in the Mediterranean Sea and the Atlantic Ocean. ICES Cooperative Research Report, 300, pp. 51–55.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
Rambaut, A., Drummond, A.J., 2007. Tracer v 1.5. Available from http://beast.bio.ed.ac.uk/Tracer.
Ramos-Onsins, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19, 2092–2100.
Reece, J.S., Bowen, B.W., Joshi, K., Goz, V., Larson, A., 2010. Phylogeography of two moray eels indicates high dispersal throughout the Indo-Pacific. J. Hered. 101, 391–402.
Saiz, E., Calbet, A., 1999. On the free-spawning reproductive behaviour of the copepod *Clausocalanus lividus* (Frost & Fleminger 1968). J. Plankton Res. 21, 599–602.
Schnack-Schiel, S.B., Mizdalski, E., Cornils, A., 2010. Copepod abundance and species composition in the Eastern subtropical/tropical Atlantic. Deep-Sea Res. II 57, 2064–2075.
Schneider, B., Schmittner, A., 2006. Simulating the impact of the Panamanian seaway closure on ocean circulation, marine productivity and nutrient cycling. Earth Planet. Sci. Let. 246, 367–380.
Schweitzer, C.E., 2001. Paleobiogeography of Cretaceous and Tertiary decapod crustaceans of the North Pacific Ocean. J. Paleontol. 75, 808–826.
Selkoe, K.A., Henzler, C.M., Gaines, S.D., 2008. Seascape genetics and the spatial ecology of marine populations. Fish Fish. 9, 363–377.
Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.
Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.
Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Unal, E., Frost, B.W., Armbrust, V., Kideys, A.E., 2006. Phylogeography of *Calanus helgolandicus* and the Black Sea copepod *Calanus euxinus*, with notes on *Pseudocalanus elongatus* (Copepoda, Calanoida). Deep-Sea Res. II 53, 1961–1975.
Weersing, K., Toonen, R.J., 2009. Population genetics, larval dispersal, and connectivity in marine systems. Mar. Ecol. Prog. Ser. 393, 1–12.
Wilke, T., Schultheiß, R., Albrecht, C., 2009. As time goes by: a simple fool's guide to molecular clock approaches in invertebrates. Am. Malacol. Bull. 27, 25–45.
Xia, X., Lemey, P., 2009. Assessing substitution saturation with DAMBE. In: Lemey, P., Salemi, M., Vandamme, A.-M. (Eds.), The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge University Press, Cambridge, UK, pp. 615–630.
Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26, 1–7.