THE ROYAL SOCIETY

Received 3 May 2003
Accepted 30 June 2003
Published online 25 September 2003

Cryptic speciation on the high seas; global phylogenetics of the copepod family Eucalanidae

Erica Goetze

Integrative Oceanography Division, 0218, 9500 Gilman Drive, Scripps Institution of Oceanography, University of California, San Diego La Jolla, CA 92093-0218, USA (*egoetze@ucsd.edu*)

Few genetic data are currently available to assess patterns of population differentiation and speciation in planktonic taxa that inhabit the open ocean. A phylogenetic study of the oceanic copepod family Eucalanidae was undertaken to develop a model zooplankton taxon in which speciation events can be confidently identified. A global survey of 20 described species (526 individuals) sampled from 88 locations worldwide found high levels of cryptic diversity at the species level. Mitochondrial (16S rRNA, CO1) and nuclear (ITS2) DNA sequence data support 12 new genetic lineages as highly distinct from other populations with which they are currently considered conspecific. Out of these 12, at least four are new species. The circumglobal, boundary current species *Rhincalanus nasutus* was found to be a cryptic species complex, with genetic divergence between populations unrelated to geographic distance. 'Conspecific' populations of seven species exhibited varying levels of genetic differentiation between Atlantic and Pacific basins, suggesting that continental landmasses form barriers to dispersal for a subset of circumglobal species. A molecular phylogeny of the family based on both mitochondrial (16S rRNA) and nuclear (ITS2, 18S rRNA) gene loci supports monophyly of the family Eucalanidae, all four eucalanid genera and the 'pileatus' and 'subtenuis' species groups.

Keywords: copepods; open ocean; speciation; cryptic species; Eucalanidae; *Rhincalanus nasutus*

1. INTRODUCTION

Marine invertebrate species often exhibit high levels of gene flow between populations owing to effective transport of planktonic larvae on ocean currents (e.g. Palumbi 1992; Lessios *et al.* 1997; Bierne *et al.* 2003*a*,*b*). This high gene flow would be expected to oppose the processes of population differentiation and speciation. Large-scale studies to examine marine speciation in near-shore invertebrates include research on sea urchins (Palumbi *et al.* 1997; Lessios *et al.* 2001), gastropods (Reid *et al.* 1996), mussels (Ladoukakis *et al.* 2002; Bierne *et al.* 2003*a*), giant clams (Benzie & Williams 1997) and starfish (Williams 1997; Williams & Benzie 1998) among others. However, little attention has been focused on marine species that inhabit the vast expanse of the open ocean. How does speciation take place in the open ocean, where barriers to dispersal are particularly difficult to discern, and where species are planktonic throughout their entire life cycle? Few data are currently available to answer this question, despite its importance to understanding the universality of evolutionary processes in high gene flow systems. Recent work on oceanic Foraminifera (de Vargas *et al.* 1999, 2001, 2002; Stewart *et al.* 2001), coccolithophores (Sáez *et al.* 2003) and deep sea fishes (Miya & Nishida 1997) suggests that greater specificity in ecological and oceanographic habitat preferences than previously supposed may be an important component of differentiation in the open ocean.

Pelagic copepods are remarkably diverse despite many biological characteristics that should inhibit the speciation process. Oceanic copepod species typically have large population sizes, which extend over vast geographic ranges, and are planktonic throughout their entire life cycle. Many have the ability to tolerate large vertical gradients in environmental properties, which may predispose them to being able to survive in a variety of oceanographic environments. Despite these characteristics they are by far the most diverse taxon of the marine zooplankton, with approximately 1800 marine calanoid species reported worldwide (Mauchline 1998). However, sibling species are common in marine taxa (Knowlton 1993, 2000; de Vargas *et al.* 2003), and recent observations of copepod 'populations' exhibiting high levels of genetic divergence despite morphological conservatism suggests that for this group reproductive isolation may be uncoupled from morphological divergence (Bucklin *et al.* 1996, 1998; Rocha-Olivares *et al.* 2001; Lee & Frost 2002). This raises the possibility that copepod species, particularly those with circumglobal biogeographic distributions, may include multiple lineages that are evolutionarily distinct.

The calanoid copepod family Eucalanidae is an oceanic taxon whose member species are likely to have arisen in the open sea. Member species are ecologically prominent in subtropical, tropical, equatorial and temperate–boreal waters of the world ocean. Eucalanids can be extremely abundant, particularly in low-oxygen regions, and sometimes constitute almost 100% of the calanoid zooplankton fauna (Muniza & Kazmi 1995). The family comprises 23 described species in four circumglobal genera (*Eucalanus*, *Rhincalanus*, *Pareucalanus*, *Subeucalanus*). Early work by Lang (1965) and Fleminger (1973) established global descriptions for biogeographic distributions of species within the family and recognized four species groups among the 17 species within the genus *Eucalanus s. l.* These included the 'subtenuis', 'pileatus', 'elongatus' and 'attenuatus' species groups, all distinguished by characteristic features of the distribution of integumental pores on the exoskeleton as well as by the shape and position of seminal receptacles. Geletin (1976) elevated two of these

Table 1. Eucalanid taxa considered in this analysis, with their approximate biogeographic distributions (modified from Lang (1965) and Fleminger (1973)) and number of individuals analysed for DNA sequences of 16S rRNA, ITS2,18S rRNA and CO1. (For species in which multiple genetic lineages were identified, the original species descriptor and date are included with the first reference to the described species in the table. Type specimens were not examined and it is not currently known which of the genetic lineages corresponds to the original species description. Abbreviations for locations: C., central; CA, California Current; NA, North Atlantic; SWP, southwest Pacific; Sulu, Sulu Sea; K/PH, Kuroshio Current and Philippine Sea; PAC, Pacific; EP, eastern Pacific; WP, western Pacific.)

			no. of individuals sequenced			
genus	species (clade)	biogeographic distributions	16S	ITS2	18S	CO1
Eucalanus	*bungii* Giesbrecht, 1892	boreal, sub-polar, N. Pacific	18	4		10
	californicus Johnson, 1938	transition zone, N. Pacific	34	4		11
	elongatus (Dana, 1849)	tropical, Pacific and Indian	4	3		
	hyalinus (1) (Claus, 1866)	tropical–subtropical, circumglobal	45	3	2	> 250
	hyalinus (2)	tropical–subtropical, circumglobal	26	3		> 250
	inermis Giesbrecht, 1892	eastern tropical Pacific	20	4	2	9
Rhincalanus	*cornutus* Dana, 1849	tropical–subtropical, Atlantic	17	4		
	rostrifrons (EP) Dana, 1852	tropical–subtropical, E. + C. Pacific	18	4	2	20
	rostrifrons (WP)	tropical–subtropical, W. Pacific	19	5		
	gigas Brady, 1883	southern ocean, circumpolar	19	3	2	
	nasutus (CA) Giesbrecht, 1888	California Current	12	4		31
	nasutus (Peru)	Humboldt Current	10	4		6
	nasutus (Sulu)	Indo-west Pacific	8	4		5
	nasutus (K/PH)	Kuroshio Current, Philippine Sea	18	3		10
	nasutus (SWP)	southwest Pacific, subtropical	35	3		8
	nasutus (NA)	northern N. Atlantic	12	3	2	3
Pareucalanus	sp.	tropical–subtropical, circumglobal	22	6		
	sewelli (NA) (Fleminger, 1973)	tropical–subtropical, Atlantic	20	4		
	sewelli (PAC)	tropical–subtropical, Pacific	15	4		
	attenuatus (Dana, 1849)	equatorial, Pacific and Indian	23	7	2	
	parki (Fleminger, 1973)	temperate, N. Pacific	12	3		
	langae (Fleminger, 1973)	temperate, southern ocean	10	3	2	
Subeucalanus	*crassus* (NA) (Giesbrecht, 1888)	tropical–subtropical, Atlantic	8	3		
	crassus (SWP)	tropical–subtropical, Pacific	3	—	1	
	sp.	Korean Strait, East China Sea	8	4		
	longiceps (Matthews, 1925)	boreal–temperate, southern ocean	10	—	2	
	monachus (Giesbrecht, 1888)	tropical–subtropical, Atlantic	15	2		
	subtenuis (Giesbrecht, 1888)	tropical, circumglobal	35	2		
	mucronatus (Giesbrecht, 1888)	tropical, Indian and W. Pacific	5	2		
	pileatus (NA) (Giesbrecht, 1888)	tropical–subtropical, Atlantic	4	4		
	pileatus (PAC)	tropical–subtropical, Pacific	5	4		
	subcrassus (Giesbrecht, 1888)	tropical, Indo-Pacific	21	4		

species groups to generic status: the genera *Subeucalanus* (comprising 'subtenuis' and 'pileatus' species groups) and *Pareucalanus*. Bradford-Grieve (1994) subsequently included *Pareucalanus peruanus* (Volkov 1971), and Prusova *et al.* (2001) described a new species, *Subeucalanus flemingeri*, from the Persian Gulf. Including the four species in *Rhincalanus*, this brings the current total to 23 described species in the family (table 1). Bjornberg (1972, 1986) also proposed that the family Eucalanidae may be polyphyletic, based on observations of naupliar swimming behaviour and general morphology. She concluded that there were two distinct lineages in the family, and proposed that *Pareucalanus* and *Subeucalanus* be placed in their own family, the Subeucalanidae, in the superfamily Centropagoidea. Such a placement could imply colonization of oceanic waters by an ancestrally coastal/neritic group rather than *in situ* diversification in the open ocean, and is therefore an important hypothesis to test in the present study.

One eucalanid species of particular interest is the cosmopolitan *R. nasutus*. *Rhincalanus nasutus* is often very abundant in the surface waters of boundary currents and upwelling zones in all three major ocean basins, and is largely absent from central oligotrophic waters (Schmaus & Lehnhofer 1927; Lang 1965; Castro *et al.* 1993). The species is eurybathic (0–4800 m), and has centres of abundance at lower epipelagic and upper mesopelagic depths (Grice & Hulsemann 1965; Lang 1965; Roe 1972; Ohman *et al.* 1998). The species has never been examined on a global spatial scale.

The present study seeks to develop an oceanic zooplankton taxon for which the phylogeny is well resolved and cryptic species have been identified. This is a necessary first step to examining speciation in the open sea, and this study will serve as the foundation for ongoing work on evolution and speciation within the Eucalanidae. I sequenced mitochondrial and nuclear DNA from 20 species of eucalanid copepods from around the world to

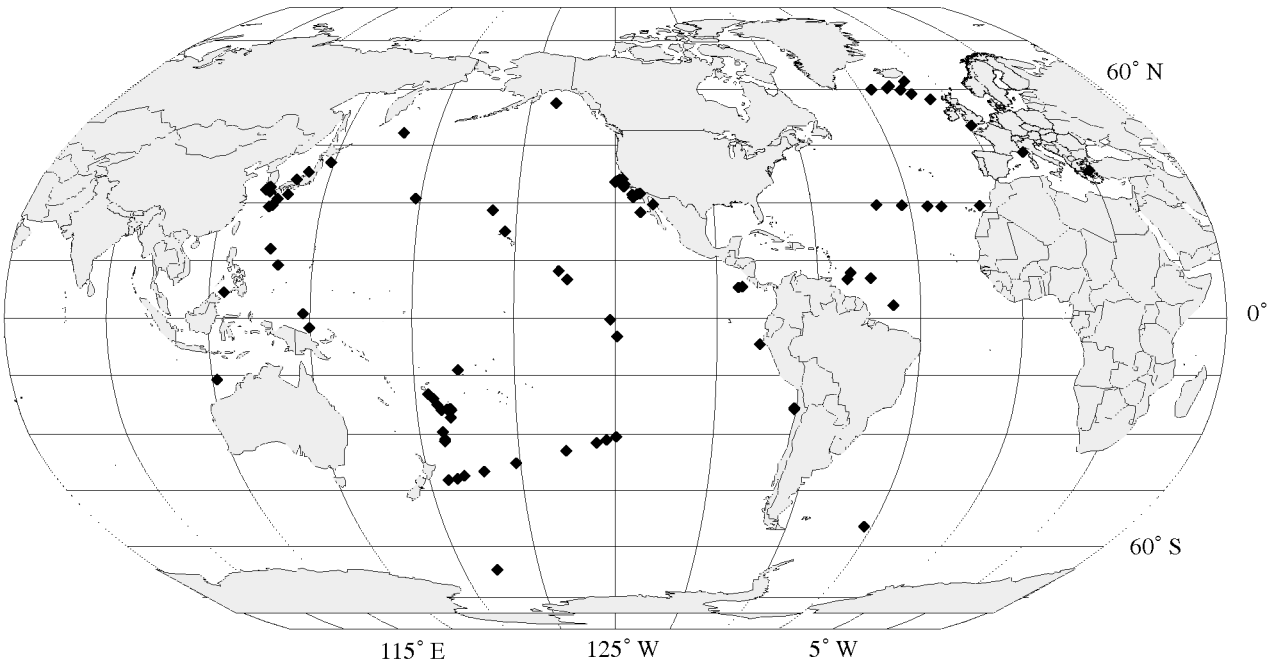

Figure 1. Collection sites for eucalanid specimens sequenced in this study.

address the following questions. (i) Do morphological species correspond to genetic clades? And are there genetic subdivisions within species? (ii) Do continents function as barriers to gene flow in circumglobal species? (iii) Do evolutionary relationships among eucalanid species match expectations based on morphological similarity? (iv) Are the Eucalanidae monophyletic?

2. MATERIAL AND METHODS

(a) *Sample collection and specimen identification*

Eucalanids were collected from a total of 88 locations worldwide (figure 1). Table 1 lists the species analysed, their approximate biogeographic distributions, and sample sizes for mitochondrial 16S ribosomal RNA (16S rRNA), nuclear internal transcribed spacer 2 (ITS2), nuclear 18S ribosomal RNA (18S rRNA) and mitochondrial cytochrome oxidase 1 (CO1) sequences. Specific sampling localities can be found in electronic Appendix A (available on The Royal Society's Publications Web site). The number of collecting localities for each species varied according to its biogeographic distribution, ranging to a maximum of 23 sampling locations for the circumglobal species *R. nasutus*. A total of 526 individuals were sequenced (16S rRNA) from 20 out of the 23 described species in the family. Samples on four major cruises were collected with 333 μm mesh plankton nets towed obliquely to either 200 m or 800–1000 m depth. The remaining samples were collected with a variety of sampling gear. Samples were either preserved in 95% ethyl alcohol, or frozen immediately in liquid nitrogen or a −80 °C freezer.

Specimens were individually sorted and identified to species following Fleminger (1973), Bradford-Grieve (1994) and Lang (1965). Species requiring integumental pore analysis for accurate identification, including *Pareucalanus attenuatus*, *P. sewelli*, *Subeucalanus subtenuis*, *S. mucronatus*, *S. crassus*, *S. pileatus* and *S. subcrassus*, were identified by removing the anterior portion of the prosome for DNA sequencing, and subjecting the posterior prosome and urosome to integumental pore analysis (as described in Fleminger 1973). In a few cases, digestion of the whole specimen of *S. pileatus* and *subcrassus* was necessary to obtain sufficient DNA for amplification. In these cases, voucher specimens of these species from the same samples were identified by integumental pore analysis.

(b) *DNA extraction, PCR and sequencing*

The 16S rRNA locus was used to screen individuals within a species for the presence of cryptic genetic lineages. Any lineages identified were then included in the ITS2 dataset by sequencing 2–7 individuals from the new lineage. Limited data for CO1 are presented here, primarily to enable comparison with other copepod groups. A pilot study demonstrated limited usefulness of CO1 as a phylogenetic marker owing to high levels of saturation.

DNA was extracted from individual copepods using either a lysis buffer protocol (Lee & Frost 2002) or the QIAGEN DNeasy tissue kit. Primers and conditions used in PCR amplification reactions can be found in electronic Appendix B. DNA sequencing was carried out on either an ABI 373 or MegaBACE 500 automated DNA sequencer.

Additional RNA/cDNA experiments were undertaken for CO1 with *E. hyalinus* 2 ($n = 2$) and *R. nasutus* (CA, $n = 3$, SWP, $n = 4$) to ensure that sequences obtained from genomic DNA based PCR reactions were expressed gene products. Synthesized cDNA was PCR amplified, cloned, sequenced and compared with sequences obtained from genomic DNA based PCR amplifications. ITS2 cloning experiments were also undertaken with two eucalanid species (*S. subtenuis* and *S. mucronatus*), in which three clones containing post-PCR ITS2 inserts were sequenced for each of two individuals. New sequence data obtained in this study can be found under GenBank accession numbers AY335822–AY335899 and AY371083–AY371094.

(c) *Sequence alignment and phylogenetic reconstruction*

Multiple sequence alignments were performed using ClustalW (Thompson *et al.* 1994), followed by manual editing as necessary in MacClade (Maddison & Maddison 2000). A series

of five alignments of the 16S rRNA data were generated using a variety of gap opening and extension costs. Alignments were then compared with published secondary structures for arthropod mitochondrial 16S rRNA genes, including *Eurytemora affinis* (Lee 1997), *Artemia salina*, *Drosophila melanogaster* (Cannone *et al.* 2002), as well as insect structures published by Buckley *et al.* (2000). Indels present within the eucalanid sequences were largely confined to helices 61, 75 and 84 in Domains IV and V (notation as in Buckley *et al.* 2000). A total of 81 sites that could not be confidently aligned were deleted, resulting in a 245 bp final alignment. The final 16S rRNA dataset contained 131 variable sites, 125 of which were parsimony informative. A genus-level 16S rRNA alignment was also generated for *Rhincalanus* sequences, from which no characters were deleted. The final ITS2 alignment (511 bp in length) contained 16 indels, 13 of which were a single nucleotide in length. The 18S rRNA sequences were highly conserved, and the final alignment was unambiguous.

Phylogenetic analyses for all datasets were conducted with maximum-parsimony (MP), maximum-likelihood (ML) and Bayesian methods. MP and ML analyses were performed with PAUP* 4.10b (Swofford 2002). MP analysis for each dataset was repeated 1000 times with random sequence addition, to explore tree space for multiple optima. For ML and Bayesian analyses, the appropriate model of molecular evolution was selected by the Akaike Information Criterion as implemented in MODELTEST (Posada & Crandall 1998). Parameter values from MODELTEST were used as a starting point, with subsequent refinement through an iterative process including: (i) heuristic searching of tree space; and (ii) re-estimation of parameter values based on the new tree topology. This process was repeated until tree topologies were stable. Final models and parameters can be found in table 3 in electronic Appendix B. Node stability was estimated in MP and ML analyses by performing 1000 replicates of the non-parametric bootstrap with 10 or 100 random sequence additions per replicate.

Bayesian analyses were conducted with MRBAYES (Huelsenbeck & Ronquist 2001). All analyses were performed with uninformative priors. Four chains were used per run (three heated and one cold), and each analysis was repeated three times, twice for two million generations, with the final analysis running for 10 million generations. All three analyses of the same dataset produced identical tree topologies.

Phylogenetic analyses were conducted on each of the datasets individually. The 16S rRNA and ITS2 data were then combined, tested for incongruence by the incongruence length difference (ILD) test, and analysed again as a combined dataset. The combined dataset consisted of 756 aligned nucleotide positions, 282 of which were parsimony informative. The ILD test was conducted with 1000 test replicates, with 100 random sequence additions per replicate. The ILD test was only marginally non-significant ($p = 0.062$). Combined results are presented in figure 2*a*, and results for separate analyses can be found in electronic Appendix D.

Several calanoid species were included here as outgroup taxa. New sequence data for *Labidocera trispinosa*, *Labidocera jollae*, *Centropages bradyi*, *Candacia bipinnata* and *Candacia* sp. were included as representative taxa of the Centropagoidea. Gene sequences of *Calanus pacificus* (AF295333), *Calanus hyperboreus* (AF227971), *Calanus finmarchicus* (AF367719), *Calanus propinquus* (AY118066), *Calanoides acutus* (AY118071), *Metridia lucens* (AF293440), *Haloptilus ocellatus* (AY118069) and *Ctenocalanus citer* (AY118078) were retrieved from GenBank to serve as outgroups in the Megcalanoidea, Arietelloidea and Clausocalanoidea (Bucklin *et al.* 1995, 2003; S. Grabbert, A. C. Bucklin, S. B. Smolenack and H. U. Dahms, unpublished data).

3. RESULTS

(a) *Molecular phylogeny of the Eucalanidae*

The three gene loci (16S rRNA, ITS2, 18S rRNA) varied considerably in their rate of molecular evolution, and each proved useful for resolving different nodes within the family phylogeny. The 16S rRNA locus exhibited the highest levels of divergence, ranging between 1.6% and 42% (uncorrected p-distances; see electronic Appendix C), and was best able to differentiate species and subspecies lineages. All previously described species exhibited fixed DNA substitutions at 16S rRNA, differentiating them from all other species. Intraspecific variation at 16S rRNA was observed, with haplotypes ranging from 0.1% to 1.7% divergent. Values close to this maximal value of 1.7% were only observed in circumglobal species *Eucalanus hyalinus* 1, *Pareucalanus* sp. and *Subeucalanus subtenuis*. Intraspecific variation within a region was typically observed to be between 0.1% and 1.0%. The youngest pair of previously described sister species, *E. californicus* and *E. bungii*, exhibited a 3% difference at 16S rRNA (table 2*d*). No variation was observed at ITS2 within individuals or within species, and the locus was most useful for resolving intermediate–deep nodes within the family.

Molecular evolutionary relationships among eucalanid species were consistent across the three gene loci and three phylogenetic reconstruction methods employed. The four genera currently included in the family Eucalanidae, *Eucalanus*, *Pareucalanus*, *Rhincalanus* and *Subeucalanus*, were all recovered as monophyletic groups in both mitochondrial and nuclear analyses. The highest levels of bootstrap support for monophyly at the genus level can be observed in the combined dataset (figure 2*a*), with all bootstrap and posterior probability support values at 100%. The 'pileatus' and 'subtenuis' species groups originally designated by Fleminger (1973) within the *Subeucalanus* genus were also consistently recovered as distinct groups across loci and methods employed. A high level of confidence for this result can be observed in all datasets, with bootstrap values ranging between 95% and 100% for these groupings across all phylogenetic methods.

Relationships among species within genera were also largely consistent with expectations based on morphological similarity. High levels of statistical support were found for sister species relationships of *E. californicus* and *bungii* (100%), *P. langae* and *parki* (98%), *S. subtenuis* and *mucronatus* (92%), *P. sewelli* and *attenuatus* (70%), and *R. rostrifrons* and *cornutus* (98%, all values for ML, combined data), as expected given greater morphological similarity between these species pairs. The ITS2 and 16S rRNA results differ only in their resolution of these sister species nodes. However, the relationships between the latter two species pairs are complicated by the discovery of two new genetic lineages, which may reflect more recent speciation events. To the best of the author's knowledge, there were no previous morphological hypotheses for relationships at deeper nodes within species groups, so the results in figure 2 represent the first assessment of patterns of divergence for these species.

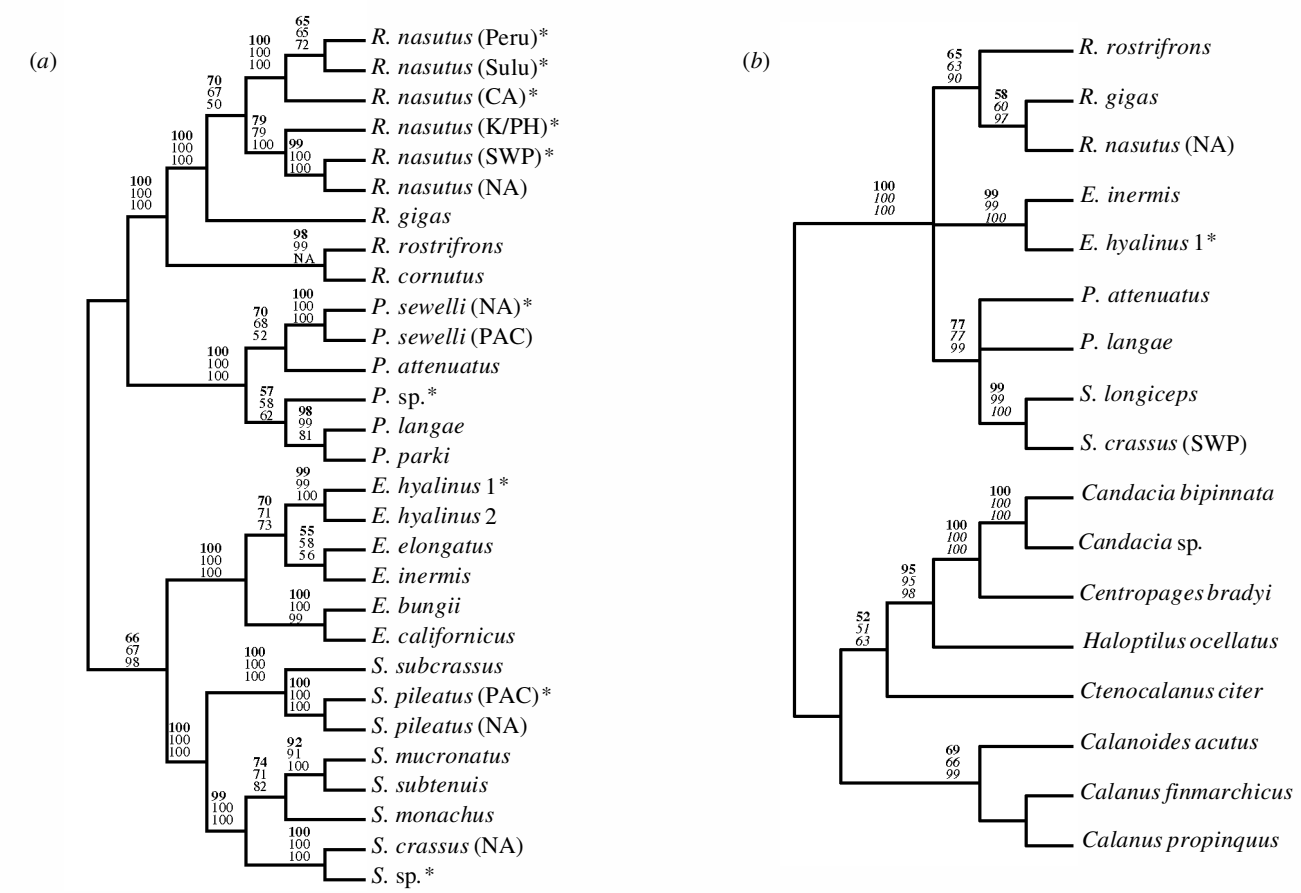

Figure 2. Results for phylogenetic analysis of (*a*) combined data including 16S rRNA + ITS2, and (*b*) 18S rRNA. Values above each node correspond, from top to bottom, to bootstrap support from ML and MP analyses, with the final values corresponding to posterior probability support from the Bayesian analysis. Values in bold are ML bootstrap values. Asterisks denote new lineages. The position of two additional new lineages, *S. crassus* (SWP) and *R. rostrifrons* (WP), can be found in the 16S rRNA analysis in figure 4*a*, electronic Appendix D.

(b) *Testing for monophyly of the Eucalanidae*

Results from outgroup analysis with all three gene loci were consistent with monophyly of the Eucalanidae. In the 18S rRNA analysis (figure 2*b*), all eucalanid and centropagoid species (*Candacia bradyi*, *Candacia bipinnata* and *Centropages* sp.) were included as ingroup taxa, and species from calanoid superfamilies Arietelloidea, Megacalanoidea and Clausocalanoidea were designated as outgroup taxa. One hundred per cent support was observed for monophyly of the Eucalanidae across all three analytical methods. Centropagoid species *C. bipinnata*, *C.* sp. and *C. bradyi* were found to group together in a monophyletic clade with other calanoid superfamily representatives, rejecting the hypothesis of Bjornberg (1972, 1986) of close relationships among Centropagoid, *Pareucalanus* and *Subeucalanus* species. Outgroup analysis results for 16S rRNA and ITS2 datasets were also found to be consistent with monophyly of the Eucalanidae.

(c) *Cryptic taxa*

Twelve new genetic lineages were discovered within the family Eucalanidae. The term 'lineage' is applied here to any population characterized by unexpectedly high levels of genetic divergence from other conspecifics. All such lineages exhibit fixed differences at mitochondrial loci from other conspecific populations, and all but one also display differences at nuclear ITS2 that appear to be fixed. The observed genetic distances between these new lineages and their closest relatives ranged between 1.6% and 23.2% (uncorrected p-distance) at 16S rRNA (table 2). These distances range from slightly less divergent than the youngest sister species pair known to be reproductively isolated, *E. californicus* and *E. bungii* (3%; table 2*d*), to divergences comparable to the deepest intrageneric nodes within the family (e.g. *E. inermis* and *E. californicus*, 19.7%; electronic Appendix C). Genetic distances were also determined at the CO1 locus, but only for species within *Eucalanus* and *Rhincalanus*. New lineages ranged from 5.1% to 24.3% divergent at CO1 from other conspecific populations (table 2). Genetically distinct lineages were found in all four eucalanid genera, and in both Atlantic and Pacific ocean basins.

Out of the 12 new genetic lineages identified here, at least four are cryptic species. Owing to the current absence of supporting ecological and morphological data, the criteria used here to designate new lineages as cryptic species are very conservative. All four cryptic species: (i) show fixed differences at both mitochondrial and nuclear loci; (ii) occur in close geographic proximity to their closest congeners; and (iii) exhibit genetic divergences greater than 3% at 16S rRNA from all other conspecific populations. Criterion (iii) was chosen to reflect the observed genetic divergence between *E. californicus* and *E. bungii*, the least genetically divergent pair of previously described sister

Table 2. Genetic differentiation of new lineages in the Eucalanidae, based on 326 bp of 16S rRNA, 511 bp of ITS2 and 518 bp of CO1.
((*a*) Divergence of cryptic species from their closest congener, (*b*) divergence of lineages with moderate to high fixed differences, (*c*) genetic differentiation between conspecific 'populations' in Atlantic and Pacific ocean basins, and (*d*) genetic differentiation of *Eucalanus californicus* and *Eucalanus bungii*, the least genetically divergent pair of previously described sister species in the family. Location abbreviations as in table 1. A dash represents no data.)

	uncorrected p-distance estimates		
genus species	16S (%)	ITS2 (%)	CO1 (%)
(*a*)			
Pareucalanus sp.–*P. sewelli* (NA)	18.1–23.2	1.8	—
Subeucalanus sp.–*S. crassus*	18.6–18.9	3.4	—
Eucalanus hyalinus 1–*E. hyalinus* 2	7.9–8.7	0.2	13.5–16.6
Rhincalanus nasutus clade 1–clade 2	15.4–18.2	3.2	20.0–24.3
(*b*)			
R. rostrifrons (WP)–*rostrifrons* (EP)	6.7–7.6	0.0	—
Rhincalanus nasutus (CA)–*nasutus* (Peru)	1.6–2.7	0.2	5.1–6.6
R. nasutus (CA/Peru)–*nasutus* (Sulu)	6.2–6.4	0.4	10.1–14.7
R. nasutus (SWP)–*nasutus* (NA)	3.6–4.0	0.2	8.8–10.1
R. nasutus (SWP/NA)–*nasutus* (K/PH)	8.7–10.3	0.8	18.3–19.7
(*c*)			
Pareucalanus sewelli	2.4	0.6	—
Rhincalanus nasutus (NA–SWP)	3.6–4.0	0.2	8.8–10.1
Subeucalanus pileatus	9.6–10.3	0.4	—
Subeucalanus crassus (NA–SWP)	3.1–4.0	—	—
Eucalanus hyalinus 1	0.5–1.7[a]	—	0.0–2.5[a]
Eucalanus hyalinus 2	0.3–0.6[a]	—	0.5–4.2[a]
Pareucalanus sp.	0.0–1.7[a]	—	—
(*d*)			
E. californicus–bungii	3	0.8	13.5

[a] Not fixed differences between populations.

species in the Eucalanidae (table 2*d*). This value is used as a general guideline for the level of genetic differentiation that can accompany speciation within this family. Criterion (ii) is included as a guideline for populations likely to have come into secondary contact, and which therefore do not fall into the ambiguous category of allopatrically distributed populations that may or may not be 'potentially interbreeding' populations *sensu* Mayr (1963*a*,*b*).

The four species identified here as cryptic species are highly divergent from their closest relatives at both 16S rRNA and CO1 (table 2*a*). The new species *Pareucalanus* sp. is inferred to have diverged before the *P. sewelli–P. attenuatus* split (figure 2; electronic Appendix D), and in the 16S rRNA analysis is found with moderate bootstrap support to be a sister species to the clade containing the three lineages, *P. sewelli* (NA), *P. sewelli* (PAC) and *P. attenuatus*. The position of *Pareucalanus* sp. is poorly resolved in the combined dataset (figure 2*a*). It was found to co-occur with *P. sewelli* in the tropical Atlantic, and it co-occurs with both *sewelli* and *attenuatus* in the western tropical Pacific (electronic Appendix A). *Eucalanus hyalinus* 2, the sister species to *hyalinus* 1 (99% support, ML), was found to co-occur with *hyalinus* 1 in subtropical and temperate waters of both Atlantic and Pacific basins (electronic Appendix A; E. Goetze, unpublished data).

A new genetic clade was discovered in the *Subeucalanus* 'subtenuis' species group which includes three genetic lineages within the nominal species *S. crassus*. One 'crassus' lineage, *Subeucalanus* sp., is considered a cryptic species, and is genetically very divergent from the remaining 'crassus' lineages (18.6%, 16S rRNA). This species is identified on the basis of specimens collected only in the Korean Strait region, and little is currently known about its biogeographic distribution. It is, however, clear that both *S. crassus* (PAC) and *Subeucalanus* sp. occur in the western Pacific (electronic Appendix A). The two most divergent clades within the *R. nasutus* species complex (see below) are also included in table 2*a*, as it is certain that there are a minimum of two species within the complex. Both major clades are present in both Atlantic and Pacific ocean basins (figure 3*a*,*c*), and the observed genetic differentiation between the two clades is high (15.4–18.2% at 16S rRNA, 20–24.3% at CO1).

Out of the remaining eight genetically distinct lineages, five exhibit moderate to high levels of fixed differences from their closest relatives, but, in the absence of morphological and ecological data, are not classified as cryptic species here (table 2*b*). In the first case, *Rhincalanus rostrifrons*, two divergent lineages are identified at 16S rRNA (6.7–7.6%), but no differences between the lineages are observed at ITS2. The two lineages occur in close proximity in the western tropical Pacific (electronic Appendix A). The four remaining lineages are members of the *R. nasutus* species complex described below.

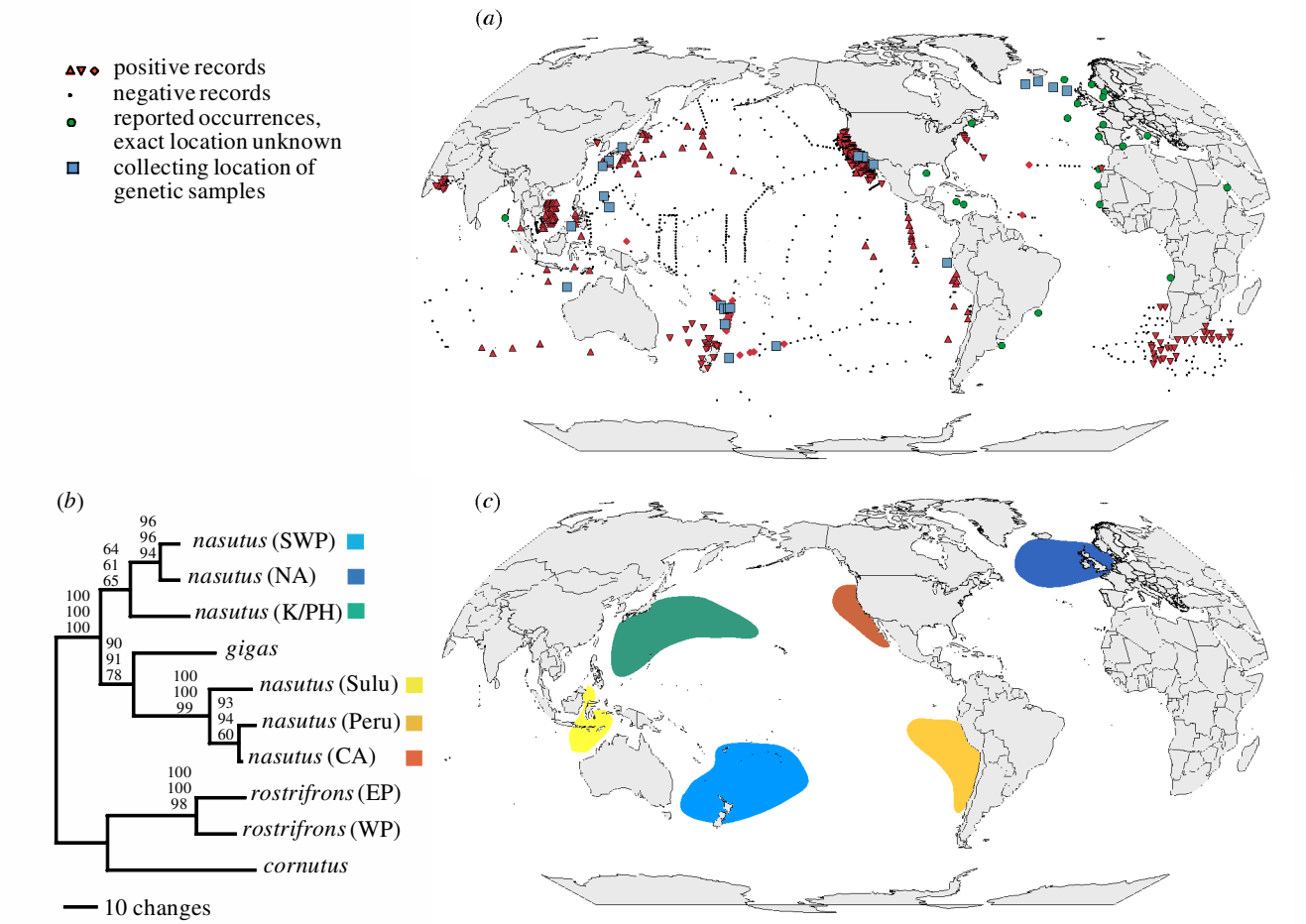

Figure 3. Geographic distribution of genetic lineages within the *Rhincalanus nasutus* species complex. (*a*) Geographic distribution of recorded locations for *R. nasutus*, and collection locations of specimens sequenced in this study. Blue squares represent genetic sample locations, red triangles and diamonds represent positive records for which exact locations are known, green circles represent positive records where *R. nasutus* has been reported to occur, but for which no exact positions are known, and black dots are negative records for *R. nasutus*. Literature sources for recorded locations include: Lang, (1965, red triangles); Bradford-Grieve (1994), CalCOFI atlases (Bowman & Johnson 1973; Fleminger 1964), Grice & Hart (1962), De Decker (1984), Deevey & Brooks (1977), Muniza & Kazmi (1995) all of which are marked by red inverted triangles. Green circles mark records from Madhupratap *et al.* (1981), Campos Hernandez (1980), Razouls & de Bovée (1999), Scotto di Carlo *et al.* (1984), Weikert & Koppelmann (1993), Poulet *et al.* (1996), and red diamonds mark new distributional data from this study. (*b*) Phylogram from ML analysis of the genus-level *Rhincalanus* sequence alignment (16S rRNA). Bootstrap values above each node listed as in figure 2. Location abbreviations as in table 1. (*c*) A preliminary interpretation of the geographic distribution of genetic lineages, based on the locations of known population centres and evolutionary relationships among sequenced specimens. Colour coding for lineages included in (*b*).

(d) *The* Rhincalanus nasutus *species complex*

Six genetically distinct lineages were identified within circumglobal *R. nasutus*, corresponding to populations in the California Current, the Humboldt Current, the Sulu Sea, the southwestern subtropical Pacific, the Kuroshio Current and Philippine Sea and the northern North Atlantic (figure 3*b*,*c*). All lineages show unique fixed differences at both mitochondrial and nuclear loci, with genetic distances between lineages ranging from 1.6% to 18.2% at 16S rRNA (table 2*a*,*b*). Levels of genetic divergence between *nasutus* lineages in all cases but one are as great or greater than that observed between *E. californicus* and *E. bungii*.

The *R. nasutus* species complex contains two main clades (figure 3*b*). One clade consists of the California, Humboldt and Sulu Sea lineages, and the second clade includes lineages from the Kuroshio Current/Philippine Sea, southwestern subtropical Pacific and the northern North Atlantic. Levels of genetic divergence between the two clades are high (15–18% 16S rRNA, 20–24% CO1), and are roughly equivalent to that between *Rhincalanus gigas* and either clade (electronic Appendix C). In the genus-level phylogenetic analysis (figure 3*b*), high bootstrap support (90–100%) was found for the placement of the Antarctic species *R. gigas* within the *nasutus* species complex, as a member of the California–Peru–Sulu lineage. In the family-level analyses (figure 2) the placement of *R. gigas* was ambiguous, with weak bootstrap support for *gigas* outside the *nasutus* complex in the ITS2 results.

(e) *Gene flow between ocean basins*

The presence of gene flow between 'populations' currently considered conspecific in Atlantic and Pacific basins appears to differ between eucalanid species. Eight species in the Eucalanidae have circumglobal distributions, of which seven were sampled in both Atlantic and Pacific

Oceans. Four of these species, *P. sewelli*, *S. crassus*, *S. pileatus* and *R. nasutus*, all show fixed differences at 16S rRNA between ocean basins (table 2*c*), with divergences ranging between 2.4% and 10.3%. Fixed DNA substitutions were also observed in the nuclear genomes of *S. pileatus*, *P. sewelli* and *R. nasutus*, the three species for which ITS2 data are also available. These divergences signify a total absence of gene flow between lineages in Atlantic and Pacific basins for these four species.

The three remaining circumglobal species, *E. hyalinus* 1, *E. hyalinus* 2 and *Pareucalanus* sp., all appear to have either exchanged genes in the recent past, or to experience ongoing gene flow between ocean basins. These three species exhibit levels of genetic differentiation between ocean basins (0.0–1.7%, 16S rRNA) that are typical of intraspecific variation.

4. DISCUSSION

(a) *Morphological species versus genetic clades*

This is one of the first phylogenetic studies to screen systematically and globally for the presence of cryptic species within an oceanic zooplankton taxon. Our results demonstrate that indeed, sibling species are common in the sea (Knowlton 1993). Validating species boundaries with molecular markers will be an essential first step to any study of a morphologically conservative marine group in which a taxonomically complete phylogeny is required. The high levels of inter-population genetic divergence and intra-population genetic cohesiveness of the 12 new genetic lineages identified here suggest that these 'populations' are distinct on both ecological and evolutionary time-scales. Results demonstrate a total absence of gene flow between these 12 lineages and other populations with which they are currently considered conspecific.

While it is not a trivial matter to determine which of these lineages should be considered new species based on genetic data alone, several species concepts can provide guidance as to what the requirements should be for a population to be recognized as a valid species. Under the Phylogenetic Species Concept (Cracraft 1989), all 12 lineages identified here would be considered new species. Designation of new species under the Biological Species Concept (Mayr 1942, 1963*a*) is more ambiguous given the absence of data on reproductive compatibility. However, at least five of these lineages occur in close enough geographic proximity to allow interbreeding. The fixed genetic differences observed in mitochondrial and nuclear genomes of four lineages demonstrate that they have not done so. In addition to these four cryptic species, eight new lineages are also likely to deserve species-level recognition following further consideration of differentiation in morphological or ecological characters. In the meantime, these lineages should be recognized as demographically and evolutionarily distinct from other conspecific populations, regardless of their current taxonomic status. Future work on the Eucalanidae will include examination of all genetically distinct lineages for differentiation in morphological characters.

The biogeographic distributions of these new lineages are incompletely known, given the limitations of sampling in the current dataset. However, it appears that examples can be found in which they occur in sympatry (*E. hyalinus* 1 and 2, *Pareucalanus* sp. and *sewelli*/*attenuatus*), parapatry (*R. rostrifrons* (EP and WP)) and allopatry (*R. nasutus*, *S. pileatus* (NA–PAC), *P. sewelli* (NA–PAC), *S. crassus* (NA–SWP)) relative to their closest congeners. Preliminary results suggest that the 10 lineages that appear in allo- and parapatry fractionate what was originally thought to be a circumglobal or large-scale biogeographic range. For example, new lineages discovered in *P. sewelli*, *S. pileatus* and *R. rostrifrons* may have distributions restricted to one ocean basin, with a sister lineage found in a second ocean basin. The two cryptic species that co-occur in sympatry with their closest congeners may also partition oceanographic habitat by depth or water mass preferences. For example, the presence of a sibling species previously cryptic within *P. sewelli* may help to explain curious observations of bimodal vertical distributions of populations in the eastern North Atlantic (Roe 1972, fig. 1, under species name *attenuatus*), as well as the presence of two distinct size groups within adult females (Fleminger 1973, fig. 16) and especially adult males of *P. sewelli* (Roe 1972, fig. 4).

Six of the genetically distinct lineages mentioned above were identified within *R. nasutus*, which has previously been considered a single circumglobal species (Schmaus & Lehnhofer 1927; Lang 1965; Castro *et al.* 1993; Bradford-Grieve 1994). The high levels of interpopulation genetic divergence observed in this taxon suggests that it is very probably a species complex. Although putative cosmopolitan marine species are increasingly being identified as cryptic species complexes (Scholin *et al.* 1995; Klatau *et al.* 1999; de Vargas *et al.* 1999, 2001, 2002; Lee 2000; Lazoski *et al.* 2001), there have been few examples for species that extend well offshore of the coastal zone. A literature compilation of recorded locations for *R. nasutus* (figure 3*a*) demonstrates that the species can and does occur in oceanic regions as far as 4000 km offshore. Given such an oceanographically broad distribution, it would be reasonable to expect high levels of gene flow at ocean basin scales. The genetic data, however, strongly support the conclusion that lineages centred in different coastal boundary currents do not exchange genes.

Furthermore, genetic relationships among *nasutus* lineages do not follow expectations based on surface ocean circulation. Dominant gyre flow fields in subtropical waters of each hemisphere might be expected to induce a more recent shared evolutionary history for lineages in eastern and western boundary currents of each hemisphere than with lineages in other parts of the globe. However, observed sister lineage pairs include the northern North Atlantic and southwestern subtropical Pacific, and the California and Peru populations (figure 3*b*,*c*). The close California–Peru relationship demonstrates more recent connectivity between these eastern boundary current lineages in the Pacific than between east–west boundary currents within each hemisphere. The sister lineage relationship between the northern North Atlantic and southwestern subtropical Pacific supports a more recent shared evolutionary history across both hemispheres and ocean basins than between clade 2 lineages in the Pacific (Kuroshio/Philippine Sea and southwest subtropical Pacific). Results from the *R. nasutus* species complex emphasize that marine species need not be neritic or coastal in distribution for allopatric populations to diverge genetically on large spatial scales, and secondly, that

evolutionary relationships between lineages may not reflect present-day surface ocean currents.

(b) *Phylogeny of the Eucalanidae, and congruence with morphological taxonomy*

Relationships among previously described eucalanid species, as inferred from both mitochondrial and nuclear DNA data, were highly congruent with results from previous morphological studies of the family. Analysis of DNA sequence data from three gene loci consistently recovered strong support for four monophyletic genera in the family Eucalanidae, with member species following designations by Fleminger (1973), Geletin (1976) and Bradford-Grieve (1994). The subgeneric clades of the 'pileatus' and subtenuis' species groups (Fleminger 1973) within *Subeucalanus* were also very well supported by the current genetic dataset. Mitochondrial and nuclear DNA data support specific-level differentiation for all 20 previously described species in the family that were sampled here.

The family Eucalanidae was also found to be monophyletic, with no support for the hypothesis of Bjornberg (1972, 1986) of two major lineages within the family. Results for all three gene loci suggest that the genera *Subeucalanus* and *Pareucalanus* are more closely related to the genera *Rhincalanus* and *Eucalanus* than to members of the calanoid superfamily Centropagoidea. Implications from this result are twofold; first, that diversification of the *Pareucalanus* and *Subeucalanus* genera was not associated with expansion of a predominantly coastal group, the Centropagoidea, into oceanic waters. Second, that naupliar morphological and behavioural evolution may be more plastic on evolutionary time-scales than previously thought (Fryer 1984; Dahms 2000). Results presented here suggest that phylogenetic relationships inferred from adult characters more closely approximate the evolutionary history of this family.

(c) *Barriers in the plankton: formidable or permeable?*

Oceanographic barriers to dispersal have long been thought to play an important role in the process of speciation in oceanic zooplankton, despite the fact that it is often remarked that it is difficult to envision how this occurs in the open ocean (McGowan 1971; Pierrot-Bults & van der Spoel 1979). Recent genetic data from planktonic Foraminifera have called this view into question, with observations of SSU or ITS haplotypes shared between ocean basins in *Orbulina universa*, *Globigerinella siphonifera* II and *Globorotalia truncatulinoides* (de Vargas *et al*. 1999, 2002, 2003), and between poles in *Globigerina bulloides*, *Turborotalia quinqueloba* and *Neogloboquadrina pachyderma* (Darling *et al.* 2000). Such observations have suggested that barriers to gene flow, such as subtropical fronts, tropical waters (for cold water species) or continental landmasses, may be highly permeable or non-existent for planktonic organisms, with the ability to establish successful populations outside the distributional range as the only key factor controlling expansion of a biogeographic range (Norris 2000; Norris & de Vargas 2000).

Results presented here suggest a qualification of this view, whereby barriers to dispersal are permeable for some species and formidable to others. Levels of genetic differentiation between Atlantic and Pacific 'populations' of seven circumglobal eucalanid species vary from 10.3% fixed sequence divergence (*S. pileatus* NA, PAC) to intraspecific haplotype sharing of 16S rRNA or CO1 haplotypes between ocean basins (*E. hyalinus* 1, *Pareucalanus* sp.; table 2*c*). Four out of seven species exhibit a total absence of gene flow between 'populations' in the Atlantic and Pacific, while the remaining three demonstrate relatively recent or ongoing genetic exchange between basins. Given that all seven species have successful populations established in both ocean basins, it appears unlikely that individuals effectively disperse, but fail to recruit in the local population. Rather, continental landmasses do appear to act as a barrier to dispersal in some, but not all species. Similar results are observed for circumglobal *Clausocalanus jobei* and *C. pergens* (Bucklin *et al.* 2003), suggesting that species specificity in patterns of gene flow on circumglobal scales may be a general phenomenon in zooplanktonic species. Additionally, the four species that exhibit an absence of gene flow between ocean basins all have distributions that extend to 40° S latitude, suggesting that the extent of the southern edge of the distributional range does not appear to be the key factor determining success in dispersal around Cape Horn (cf. Fleminger & Hulsemann 1973). It is currently unclear what characteristic of the life history, ecological specificity, evolutionary history and/or biogeography of a species determines its ability to disperse across such semi-permeable barriers as the continental landmasses of the Americas.

This paper is dedicated to Abraham Fleminger, whose life's work on the evolution of calanoid copepods continues to be an inspiration to the author. His careful morphological analysis of the Eucalanidae serves as a foundation for the work presented here. The following people very kindly provided specimens from disparate regions of the global ocean: G. Mitchell and S. Storms, P. Fiedler and V. Philbrick, M. Galbraith, B. Frost, S. Ohtsuka, S. Nishida and R. Machida, Y.-S. Kang, M. Landry and C. Allen, J. Grieve, V. Andersen, S. Gasparini, X. Irigoien, E. Carpenter and A. Subramaniam, R. Harris and T. Smith, R. Escribano and P. Hidalgo Diaz, P. Ayon, and R. Wilson and T. Baiz. A. Yayanos and D. Capone welcomed E.G. aboard their cruises during 2001. Grace Wyngaard and Sarah Gerken kindly provided PCR primers for ITS2. M. D. Ohman, J. M. Flowers, C. S. Willet, R. Burton, N. Knowlton, C. de Vargas, W. Newman and J. Kohn provided discussion, encouragement, technical advice and comments on the manuscript. The work was supported by NSF grants OCE02-21063 and OCE01-10300, a UC ship funds grant and an EPA STAR graduate fellowship.

REFERENCES

Benzie, J. A. H. & Williams, S. T. 1997 Genetic structure of giant clam (*Tridacna maxima*) populations in the West Pacific is not consistent with dispersal by present-day ocean currents. *Evolution* **51**, 768–783.

Bierne, N., Borsa, P., Daguin, C., Jollivet, D., Viard, F., Bonhomme, F. & David, P. 2003*a* Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *M. galloprovincialis*. *Mol. Ecol.* **12**, 447–461.

Bierne, N., Bonhomme, F. & David, P. 2003*b* Habitat preference and the marine-speciation paradox. *Proc. R. Soc. Lond.* B **270**, 1399–1406. (DOI 10.1098/rspb.2003.2404.)

Bjornberg, T. K. S. 1972 Developmental stages of some tropical and subtropical planktonic marine copepods. In *Studies on the fauna of Curacao and other Caribbean Islands*, vol. XL

(ed. P. W. Hummelinck & L. J. van der Steen), pp. 1–185. Dordrecht, The Netherlands: Martinus Nijhoff.

Bjornberg, T. K. S. 1986 The rejected nauplius: a commentary. Second International Conference on Copepoda, Ottawa. *Syllogeus* **58**, 232–236.

Bowman, T. E. & Johnson, M. W. 1973 *Distributional atlas of calanoid copepods in the California Current region, 1949 and 1950*. La Jolla, CA: Marine Life Research Program.

Bradford-Grieve, J. M. 1994 *The marine fauna of New Zealand: pelagic calanoid Copepoda: Megacalanidae, Calanidae, Paracalanidae, Mecynoceridae, Eucalanidae, Spinocalanidae, Clausocalanidae*. New Zealand Oceanographic Institute Memoir 102. Wellington: NIWA (National Institute of Water and Atmospheric Research).

Buckley, T. R., Simon, C., Flook, P. K. & Misof, B. 2000 Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. *Insect Mol. Biol.* **9**, 565–580.

Bucklin, A., Frost, B. W. & Kocher, T. D. 1995 Molecular systematics of six *Calanus* and three *Metridia* species (Calanoida: Copepoda). *Mar. Biol.* **121**, 655–664.

Bucklin, A., LaJeunesse, T. C., Curry, E., Wallinga, J. & Garrison, K. 1996 Molecular diversity of the copepod, *Nannocalanus minor*: genetic evidence of species and population structure in the North Atlantic. *J. Mar. Res.* **54**, 285–310.

Bucklin, A., Bentley, A. M. & Franzen, S. P. 1998 Distribution and relative abundance of *Pseudocalanus moultoni* and *P. newmani* (Copepoda: Calanoida) on Georges Bank using molecular identification of sibling species. *Mar. Biol.* **132**, 97–106.

Bucklin, A., Frost, B., Bradford-Grieve, J., Allen, L. & Copley, N. 2003 Molecular systematic and phylogenetic assessment of 34 calanoid copepod species of the Calanidae and Clausocalanidae. *Mar. Biol.* **142**, 333–343.

Campos Hernandez, A. 1980 Distribucion y abundancia relativa de los copepodos planctonicos en el Golfo de Mexico y el Caribe. *Cienc. Biol.* (*La Habana*) **5**, 57–74.

Cannone, J. J. (and 13 others) 2002 The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BioMed Central Bioinformatics* **3**, 15.

Castro, L. R., Bernal, P. A. & Tronsco, V. A. 1993 Coastal intrusion of copepods: mechanisms and consequences on the population biology of *Rhincalanus nasutus*. *J. Plankton Res.* **15**, 501–515.

Cracraft, J. 1989 Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In *Speciation and its consequences*, vol. xiii (ed. D. Otte & J. A. Endler), p. 679. Sunderland, MA: Sinauer Associates.

Dahms, H.-U. 2000 Phylogenetic implications of the crustacean nauplius. *Hydrobiologia* **417**, 91–99.

Darling, K. F., Wade, C. M., Stewart, I. A., Kroon, D., Dingle, R. & Leigh Brown, A. J. 2000 Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminfers. *Nature* **405**, 43–47.

De Decker, A. H. B. 1984 Near-surface copepod distribution in the south-western Indian Ocean and the south-eastern Atlantic Ocean. *Ann. S. Afr. Mus.* **93**, 303–370.

Deevey, G. B. & Brooks, A. L. 1977 Copepods of the Sargasso Sea off Bermuda: species composition, and vertical and seasonal distribution between the surface and 2000 m. *Bull. Mar. Sci.* **27**, 256–291.

de Vargas, C., Norris, R., Zaninetti, L., Gibb, S. W. & Pawlowski, J. 1999 Molecular evidence of cryptic speciation in planktonic foraminifera and their relation to oceanic provinces. *Proc. Natl Acad. Sci. USA* **96**, 2864–2868.

de Vargas, C., Renaud, S., Hilbrecht, H. & Pawlowski, J. 2001 Pleistocene adaptive radiation in *Globorotalia truncatulinoides*: genetic, morphologic, and environmental evidence. *Paleobiology* **27**, 104–125.

de Vargas, C., Bonzon, M., Rees, N. W., Pawlowski, J. & Zaninetti, L. 2002 A molecular approach to biodiversity and biogeography in the planktonic foraminifer *Globigerinella siphonifera* (d/Orbigny). *Mar. Micropaleontol.* **45**, 101–116.

de Vargas, C., Sáez, A. G., Medlin, L. K. & Thierstein, H. R. 2003 Super species in the calcareous plankton. In *Coccolithophores—from molecular processes to global impact* (ed. H. R. Thierstein & J. R. Young). Berlin: Springer. (In the press.)

Fleminger, A. 1964 *Distributional atlas of calanoid copepods in the California Current region*, vol. 1. La Jolla, CA: California Manne Research Committee.

Fleminger, A. 1973 Pattern, number, variability, and taxonomic significance of integumental organs (sensilla and glandular pores) in the genus *Eucalanus* (Copepoda, Calanoida). *Fish. Bull. Wash. DC* **71**, 965–1010.

Fleminger, A. & Hulsemann, K. 1973 Relationship of Indian Ocean epiplanktonic calanoids to the world oceans. *Ecol. Stud. Anal. Synth.* **3**, 339–347.

Fryer, G. 1984 Structure, function, and behavior, and the elucidation of evolution in copepods and other crustaceans. *Syllogeus* **58**, 150–157.

Geletin, Y. V. 1976 The ontogenetic abdomen formation in copepods of genera *Eucalanus* and *Rhincalanus* (Calanoida, Eucalanidae) and new system of these copepods. *Issledovaniya Fauny Morei* **18**, 75–93.

Grice, G. D. & Hart, A. D. 1962 The abundance, seasonal occurrence and distribution of the epizooplankton between New York and Bermuda. *Ecol. Monogr.* **32**, 287–309.

Grice, G. D. & Hulsemann, K. 1965 Abundance, vertical distribution and taxonomy of calanoid copepods at selected stations in the northeast Atlantic. *J. Zool.* **146**, 213–262.

Huelsenbeck, J. P. & Ronquist, F. R. 2001 MrBayes: Bayesian inference of phylogenetic trees. *Biometrics* **17**, 754–755.

Klatau, M., Russo, C. A. M., Lazoski, C., Boury-Esnault, N., Thorpe, J. P. & Sole-Cava, A. M. 1999 Does cosmopolitanism result from overconservative systematics? A case study using the marine sponge *Chondrilla nucula*. *Evolution* **53**, 1414–1422.

Knowlton, N. 1993 Sibling species in the sea. *A. Rev. Ecol. Syst.* **24**, 189–216.

Knowlton, N. 2000 Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* **420**, 73–90.

Ladoukakis, E. D., Saavedra, C., Magoulas, A. & Zouros, E. 2002 Mitochondrial DNA variation in a species with two mitochondrial genomes: the case of *Mytilus galloprovincialis* from the Atlantic, the Mediterranean and the Back Sea. *Mol. Ecol.* **11**, 755–769.

Lang, B. T. 1965 Taxonomic review and geographical survey of the copepod genera *Eucalanus* and *Rhincalanus* in the Pacific Ocean. PhD thesis, Scripps Institution of Oceanography, University of California, San Diego, CA, USA.

Lazoski, C., Sole-Cava, A., Boury-Esnault, N., Klatau, M. & Russo, C. 2001 Cryptic speciation in a high gene flow scenario in the oviparous marine sponge *Chondrosia reniformis*. *Mar. Biol.* **139**, 421–429.

Lee, C. E. 1997 Global phylogeography of the copepod species complex *Eurytemora affinis* and recent invasions of fresh water. PhD thesis, Department of Oceanography, University of Washington, Seattle, WA, USA.

Lee, C. E. 2000 Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate 'populations'. *Evolution* **54**, 2014–2027.

Lee, C. E. & Frost, B. W. 2002 Morphological stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae). *Hydrobiologia* **480**, 111–128.

Lessios, H. A., Kessing, B. D. & Robertson, D. R. 1997 Massive gene flow across the world's most potent marine biogeographic barrier. *Proc. R. Soc. Lond.* B **265**, 583–588. (DOI 10.1098/rspb.1998.0334.)

Lessios, H. A., Kessing, B. D. & Pearse, J. S. 2001 Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. *Evolution* **55**, 955–975.

McGowan, J. A. 1971 Oceanic biogeography of the Pacific. In *The micropaleontology of the oceans* (ed. B. M. Funnel & W. R. Reidel), pp. 3–74. Cambridge University Press.

Maddison, D. R. & Maddison, W. P. 2000 MacClade *4: analysis of phylogeny and character evolution. Version 4.0*. Sunderland, MA: Sinauer Associates.

Madhupratap, M., Sreekumaran Nair, S. R., Achuthankutty, C. T. & Nair, V. R. 1981 Major crustacean groups and zooplankton diversity around Andaman-Nicobar Islands. *Indian J. Mar. Sci.* **10**, 266–269.

Mauchline, J. 1998 *The biology of calanoid copepods*. San Diego, CA: Academic Press.

Mayr, E. 1942 *Systematics and the origin of species*. New York: Columbia University Press.

Mayr, E. 1963*a* *Animal species and evolution*. Cambridge, MA: Harvard University Press.

Mayr, E. 1963*b* *Populations, species, and evolution*. Cambridge, MA: Harvard University Press.

Miya, M. & Nishida, M. 1997 Speciation in the open ocean. *Nature* **389**, 803–804.

Muniza, F. & Kazmi, Q. B. 1995 Occurrence and abundance of two genera: *Eucalanus* and *Rhincalanus* of the family Eucalanidae (Copepoda, Calanoida) in the samples of Nasser cruise 1, January 1992. *Sci. Int.* (*Lahore*) 7, 395–400.

Norris, R. D. 2000 Pelagic species diversity, biogeography, and evolution. *Paleobiology* **26**, 236–258.

Norris, R. D. & de Vargas, C. 2000 Evolution all at sea. *Nature* **405**, 23–24.

Ohman, M. D., Drits, A. V., Clarke, M. E. & Plourde, S. 1998 Differential dormancy of co-occurring copepods. *Deep-sea Res.* **45**, 1709–1740.

Palumbi, S. R. 1992 Marine speciation on a small planet. *Trends Ecol. Evol.* **7**, 114–118.

Palumbi, S. R., Grabowsky, G., Duda, T., Geyer, L. & Tachino, N. 1997 Speciation and population genetic structure in tropical Pacific sea urchins. *Evolution* **51**, 1506–1517.

Pierrot-Bults, A. C. & van der Spoel, S. (ed.) 1979 Speciation in macrozooplankton. In *Zoogeography and diversity of plankton*, pp. 144–167. New York: Halstead Press.

Posada, D. & Crandall, K. A. 1998 Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Poulet, S. A., Laabir, M. & Chaudron, Y. 1996 Characteristic features of zooplankton in the Bay of Biscay. *Sci. Mar.* **60**, 79–95.

Prusova, I. Y., Al-Yamani, F. & Al-Mutairi, H. 2001 *Subeucalanus flemingeri* sp. n. from the Arabian Gulf (Copepoda: Eucalanidae). *Zoosystematica Rossica* **10**, 47–54.

Razouls, C. & de Bovée, F. 1999 Diversité et répartion géographique chez les Copépodes planctoniques marins. See http://www.obs-banyuls.fr/Copepodes/Razouls1.html.

Reid, D. G., Rumbak, E. & Thomas, R. H. 1996 DNA, morphology and fossils: phylogeny and evolutionary rates of the gastropod genus *Littorina*. *Phil. Trans. R. Soc. Lond.* B **351**, 877–895.

Rocha-Olivares, A., Fleeger, J. W. & Foltz, W. 2001 Decoupling of molecular and morphological evolution in deep lineages of a meiobenthic harpacticoid copepod. *Mol. Biol. Evol.* **18**, 1088–1102.

Roe, H. S. J. 1972 The vertical distributions and diurnal migrations of calanoid copepods collected on the SOND cruise, 1965. II. Systematic account: families Calanidae up to and including the Aetideidae. *J. Mar. Biol. Assoc. UK* **52**, 315–343.

Sáez, A. G., Probert, I., Geisen, M., Quinn, P., Young, J. R. & Medlin, L. K. 2003 Pseudocryptic speciation in coccolithophores. *Proc. Natl Acad. Sci. USA* **100**, 7163–7168.

Schmaus, P. H. & Lehnhofer, K. 1927 Copepoda (4): Rhincalanus Dana 1852 der Deutschen Tiefsee Expedition. Systematik und Verbreitung der Gattung. *Wissenschaftliche Ergebnisse der Deutschen Tiefsee-Expedition auf dem Dampfer 'Valdivia' 1898–1899* **23**, 355–400.

Scholin, C. A., Hallegraeff, G. M. & Anderson, M. 1995 Molecular evolution of the *Alexandrium tamarense* 'species complex' (Dinophyceae): dispersal in the North American and west Pacific regions. *Phycologia* **34**, 472–485.

Scotto di Carlo, B., Ianora, A., Fresi, E. & Hure, J. 1984 Vertical zonation patterns for Mediterranean copepods from the surface to 3000 m at a fixed station in the Tyrrhenian Sea. *J. Plankton Res.* **6**, 1031–1056.

Stewart, I. A., Darling, K. F., Kroon, D., Wade, C. M. & Troelstra, S. R. 2001 Genotypic variability in subarctic Atlantic planktic foraminifera. *Mar. Micropaleontol.* **43**, 143–153.

Swofford, D. L. 2002 *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*, v. 4. Sunderland, MA: Sinauer Associates.

Thompson, J. D., Higgins, D. G. & Gibson, J. 1994 ClustalW: improving the sensitivity of progressive mulitple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

Volkov, A. F. 1971 New species of Eucalanus genera (Crustacea: Copepoda) from Peruan Current. *Izvestiya Tikhookeanskogo nauchno-issledovatel'skogo Instituta rybnogo khozyaistva i okeanorafii* **79**, 141–143. [In Russian.]

Weikert, H. & Koppelmann, R. 1993 Vertical structural patterns of deep-living zooplankton in the NE Atlantic, the Levantine Sea and the Red Sea: a comparison. *Oceanol. Acta* **16**, 163–177.

Williams, S. T. 1997 Indo-west Pacific patterns of genetic differentiation in the high dispersal starfish, *Linckia laevigata*. *Mol. Ecol.* **6**, 559–573.

Williams, S. T. & Benzie, J. A. H. 1998 Evidence of a biogeographic break between populations of a high dispersal starfish: congruent regions within the Indo-west Pacific defined by color morphs, mtDNA, and allozyme data. *Evolution* **52**, 87–99.

As this paper exceeds the maximum length normally permitted, the author has agreed to contribute to production costs.

Visit http://www.pubs.royalsoc.ac.uk to see electronic appendices to this paper.