

Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoan taxon

Africa Gómez¹, Peter J. Wright², David H. Lunt¹, Juan M. Cancino³, Gary R. Carvalho^{1,2} and Roger N. Hughes^{2,*}

¹Department of Biological Sciences, University of Hull, Cottingham Road, Hull HU6 7RX, UK

²School of Biological Sciences, University of Wales, Bangor, Gwynedd LL57 2UW, UK

³Facultad de Ciencias, Universidad Católica de la Santísima Concepción, Casilla 297, Concepción, Chile

Despite increasing threats to the marine environment, only a fraction of the biodiversity of the oceans has been described, owing in part to the widespread occurrence of cryptic species. DNA-based barcoding through screening of an orthologous reference gene has been proposed as a powerful tool to uncover biological diversity in the face of dwindling taxonomic expertise and the limitations of traditional species identification. Although DNA barcoding should be particularly useful in the sea, given the prevalence of marine cryptic species, the link between taxa identified through DNA barcodes and reproductively isolated taxa (biological species) has rarely been explicitly tested. Here, we use an integrated framework comparing breeding compatibility, morphology and mitochondrial (cytochrome *c* oxidase 1) and nuclear (elongation factor-1-alpha) DNA sequence variation among globally distributed samples of the cosmopolitan marine bryozoan *Celleporella hyalina* (L.). Our results reveal that *C. hyalina* comprises numerous deep, mostly allopatric, genetic lineages that are reproductively isolated, yet share very similar morphology, indicating rampant cryptic speciation. The close correspondence between genetic lineages and reproductively isolated taxa in the context of minimal morphological change suggests that DNA barcoding will play a leading role in uncovering the hidden biodiversity of the oceans and that the sole use of morphologically based taxonomy would grossly underestimate the number of marine species.

Keywords: cytochrome *c* oxidase 1; cryptic species; DNA barcoding; elongation factor 1-alpha; marine biodiversity; mitochondrial DNA

1. INTRODUCTION

The evolution of biological species diversity has often been accompanied by a corresponding expression of morphological variation, epitomized, for example, by the African lake cichlids, Darwin's finches or Hawaiian *Drosophila*, in which rapid, recent evolutionary radiations have generated diverse colour patterns, mouth and beak shapes (Givnish *et al.* 2000). However, species do not always look distinct. Some are remarkably difficult to identify morphologically, often requiring subtle morphological or morphometric analyses, yet they comply with the classical definition of biological species through reproductive incompatibility (Mayr 1942). Such taxa can display substantial cladogenesis with little or no morphological change (Avice *et al.* 1994; Hebert 1998; Jackson & Cheetham 1999), often forming cryptic species complexes (Knowlton 1993, 2000; Gómez *et al.* 2002). Failure to detect such cryptic species not only risks underestimating biodiversity (Knowlton 2000), but also hinders interpretation of ecological relationships (Dodson *et al.* 2003), definition of biogeographical barriers (Kucera & Darling 2002) and understanding of

speciation patterns (Jackson & Cheetham 1999; Knowlton 2000; Hebert *et al.* 2004a).

DNA barcoding (Hebert & Gregory 2005) has been proposed as a novel and powerful taxonomic tool. Hebert *et al.* (2003a) first called for DNA 'barcodes' (a standardized DNA sequence) of all species to be established, with the goals of: (i) assigning unknown individuals, life-history stages and forensic or subfossil material to already described species, and (ii) aiding the discovery of cryptic biodiversity, including new species. Although stirring controversy, with passionate opponents in some taxonomic circles (Moritz & Cicero 2004; Will & Rubinoff 2004; Ebach & Holdrege 2005), the value of DNA barcoding has been widely recognized and enthusiastically embraced in many quarters; as a result, copious new sequence data are being generated (Hajibabaei *et al.* 2005). The effectiveness of DNA barcoding as a species-identification tool has been demonstrated in a variety of groups from copepods (Bucklin *et al.* 2003) and springtails (Hogg & Hebert 2004) to birds (Hebert *et al.* 2004b; Johnson & Cicero 2004) and primates (Lorenz *et al.* 2005).

In well-known organisms, the existing taxonomy, mainly based on traditional approaches, offers a useful framework for testing the validity of DNA barcoding. Other groups, including many marine phyla, are rife with cryptic species (Knowlton 1993, 2000), and are therefore not amenable to traditional taxonomic methodology

* Author for correspondence (r.n.hughes@bangor.ac.uk).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2006.3718> or via <http://www.journals.royalsoc.ac.uk>.

(e.g. nematodes, picoplankton; Moon-van der Staay *et al.* 2001; Floyd *et al.* 2002). In such cases, the value of DNA barcoding is to reveal hidden biological diversity; in this way, the DNA database becomes a reference framework, generating hypotheses of species status that can be tested and supplemented with other information. Indeed, across the eukaryotic spectrum, taxa previously recognized as single morphospecies are being resolved by molecular techniques into complexes of cryptic species (Knowlton 1993; Kucera & Darling 2002; Hebert *et al.* 2004a).

The successful application of DNA barcoding to difficult or poorly known groups, however, has been impeded by a major uncertainty. Levels of DNA sequence divergence within and between species vary widely among taxa (Hewitt 1996; Avise & Johns 1999; Hebert *et al.* 2003b). For example, average sister-species sequence divergences in cytochrome *c* oxidase 1 (COI or *cox1*), the standard mitochondrial DNA (mtDNA) barcoding gene, vary between 1% in cnidarians and over 15% in annelids and crustaceans (Hebert *et al.* 2003b). The correspondence of species boundaries defined solely by DNA and by more traditional means therefore remains unclear. As Blaxter (2003) notes: 'a major unresolved issue is how closely these molecular taxa correspond to what traditional biologists would recognize as species'. At present, there is no objective criterion that guarantees recognition of species boundaries when groups have been retrieved through barcoding. Methods employed to assign individuals to known species include using the ratio of intraspecific to interspecific variation as a screening threshold for species status (Hebert *et al.* 2004b) or using a given threshold of sequence similarity (e.g. 99%) to cluster sequences into molecular operational taxonomic units (MOTUs; Floyd *et al.* 2002). However, both methods have drawbacks. The 'threshold' method is heavily influenced by our ability to recognize species using traditional methods, and it requires *a priori* knowledge of both species identity and phylogenetic relationships in a well-sampled group (Meyer & Paulay 2005). Even in these ideal circumstances, this method is prone to errors (Meyer & Paulay 2005). The 'MOTU' method can be an arbitrary exercise that attempts no contextual link with traditionally based species. Consequently, there is a pressing need to establish correspondence between barcode groups and biological species (Eyualem & Blaxter 2003; Sáez & Lozano 2005) in order to achieve effective integration of traditional and molecular methodologies.

Given the ubiquity of cryptic species in the marine environment (Knowlton 1993, 2000), the application of DNA barcoding might prove to be crucial for uncovering true levels of biodiversity (Schander & Willassen 2005). Here, we test the ability of DNA barcoding to recover groups consistent with biological species in a model marine taxon. We screened global samples of a common bryozoan for COI sequence diversity, measured their morphological differentiation and performed mating trials to assess reproductive isolation and delineate biological species.

The bryozoan *Celleporella hyalina* (L.), order Cheilostomata, family Hippothoidae, has an almost cosmopolitan distribution in cold and temperate shallow subtidal waters, where it colonizes macro-algal fronds and, to a lesser extent, molluscan shells, rocks and other hard surfaces including plastic refuse. Despite its broad distribution and relatively short larval dispersal (Cancino

& Hughes 1988), *C. hyalina* is presently recognized as a single species. Although described by Linnaeus over 200 years ago, and frequently studied by traditional taxonomic means, high levels of phenotypic plasticity have precluded any consistent taxonomic subdivision of this species. *Celleporella hyalina* is a simultaneous hermaphrodite, but colonies are usually self-incompatible (Hoare & Hughes 2001; Hughes *et al.* 2001). As mate choice depends only on gamete recognition, breeding compatibility can be assessed objectively in the laboratory from larval output and offspring viability (Hughes *et al.* 2002). Moreover, because colonies grow by modular iteration, genetically identical cuttings (ramets) can readily be obtained and propagated indefinitely, facilitating rigorous design of cross-mating experiments (see details in the electronic supplementary material). *Celleporella hyalina*, therefore, provides an especially amenable system with which to investigate the power of DNA barcoding in recovering reproductively isolated taxa (biological species).

Our results on DNA barcoding, mating compatibility and morphological variation between geographical isolates demonstrate prolific cryptic speciation in *C. hyalina*. We show that the major genetic groups recovered by DNA barcoding are distinct biological species that have diverged little by standard morphological measures.

2. MATERIAL AND METHODS

(a) Collection of colonies

Our sampling programme aimed to represent the global distribution of potentially suitable habitat, but it was partially constrained by accessibility and logistics (see figure 2c and the electronic supplementary material). Thus, we failed to sample known localities for *C. hyalina* in the West Pacific, Siberian and Canadian Arctic Ocean, Central American East Pacific, Galápagos Archipelago and Moroccan Atlantic. Regions in the Southern Hemisphere that were sampled but failed to yield *C. hyalina* (Antipodes, Antarctica, Sub-Antarctic islands, South Africa) are not reported hereafter. At least 20 colonies were collected per site from the intertidal or shallow subtidal and kept damp and cool in thermos flasks during transportation.

(b) Production of ramets

Colonies collected from the field were cultured under geographical segregation in the laboratory, where they released larvae over several weeks. Larvae were allowed to settle on acetate sheeting, individually isolated by excision while still attached to underlying acetate and glued to a new master sheet. After growing to a size convenient for propagation, the new colonies were cut into several pieces. Cuttings (ramets) were glued to new acetate and grown until ready for further propagation or use in mating trials (for details of propagation and culturing see Manríquez *et al.* 2001). Clones were established from four larvae randomly selected from more than 100 released, minimizing the probability of obtaining sib or half-sib lineages. The ramets to be used in breeding experiments were kept only in the presence of clone-mates to ensure absence of stored allosperm (Hughes *et al.* 2002).

(c) DNA extraction

Total DNA was extracted from a fragment of colony (5–10 zooids), an ancestrula or a single embryo using Chelex resin (Instagene, Biorad). Owing to the occurrence of a variety of epibionts, the use of single embryos obtained from ripe

colonies was favoured for DNA extractions from wild colonies. DNA was extracted from both the specimens fixed in the wild and, in some cases, also from laboratory-grown colonies, totalling 5–10 colonies per site. A total of 176 colonies were extracted.

(d) PCR amplification and sequencing

PCRs were performed in 10 μ l total volume containing 2 μ l template DNA, 1.5 mM MgCl₂, 200 μ M dNTPs, 2.5 pmol of each primer, 16 mM (NH₄)₂SO₄, 67 mM Tris–HCl (pH 8.8 at 25°C), 0.01% Tween 20 buffer and 0.125 U Taq DNA polymerase. Amplification conditions were 3 min at 93°C (15 s at 92°C, 20 s at 45–55°C and 60 s at 70°C) \times 40 and 3 min at 72°C. A 710 bp region of the COI gene was amplified with primers LCO1490 and HCO2198 (Folmer *et al.* 1994) or purpose-designed *C. hyalina*-specific primers COICH-F (5'-ATTTTATGTTTGGCTTATGAGCTGG-3') and COICH-R (5'-CAAAAATCAAAATAAATGTTGGTA-3'). After trimming the sequences and collapsing the identical haplotypes, the COI dataset consisted of 576 bp for 99 taxa, including outgroups. Primers EFC169 (5'-GGCCATCGTGATTTTCATCAAGAACATGAT-3') and EFC628R (5'-ACNGTDCCAATACCWCCDATYTTGTA-3') were designed from alignments of published animal sequences of the nuclear gene elongation factor-1-alpha (EF-1 α). These were used to amplify intron 2 and part of exons 2 and 3 from 24 individual bryozoan colonies representing the diversity of mtDNA lineages. The PCR products were purified using Wizard spin columns (Promega) and 5 μ l purified PCR product ligated into pGEM-Teasy (Promega) cloning vectors. Transformation was carried out by standard techniques and the presence of plasmid inserts in the resulting colonies was confirmed by PCR using M13 primers. One microlitre of colony PCR product was used to inoculate a cycle sequencing reaction. Cy5-labelled primers were used to sequence both the strands of PCR products using the Thermo Sequenase Primer Cycle Sequencing kit (Amersham Pharmacia Biotech) on an ALFexpress (Amersham Pharmacia Biotech) automated sequencer. The mtDNA dataset could be easily aligned by eye with no insertion or deletion events. However, only the EF-1 α exon sequences (452 bp) were analysed as the intron varied in size and could not be confidently aligned. DNA sequences are available in GenBank (accession numbers DQ999599–DQ999797 for COI and EF-1 α respectively).

(e) Phylogenetic analyses

Phylogenetic analyses on both genes were carried out using maximum-likelihood, neighbour-joining and Bayesian inference as implemented in PHYML v. 2.4 (Guindon & Gasquet 2003), PAUP* v. 4.0b10 (Swofford 1998) and MRBAYES v. 3.0b4 (Ronquist & Huelsenbeck 2003). MODELTEST (Posada & Crandall 1998) was used to select the maximum-likelihood model of evolution. The model selected was GTR+I+G for the COI dataset and K80+G for the EF-1 α dataset. PAUP* was used to generate neighbour-joining trees using maximum-likelihood distances. Maximum-likelihood per cent bootstraps for 1000 pseudoreplicates were generated with PHYML v. 2.4. For the Bayesian inference, the Markov chain Monte Carlo search was run with four chains (one cold and three heated) for 2 000 000 generations with trees being sampled every 100 generations (the first 500 trees were discarded as 'burn-in'). A majority-rule consensus of all the trees sampled was created using PAUP*.

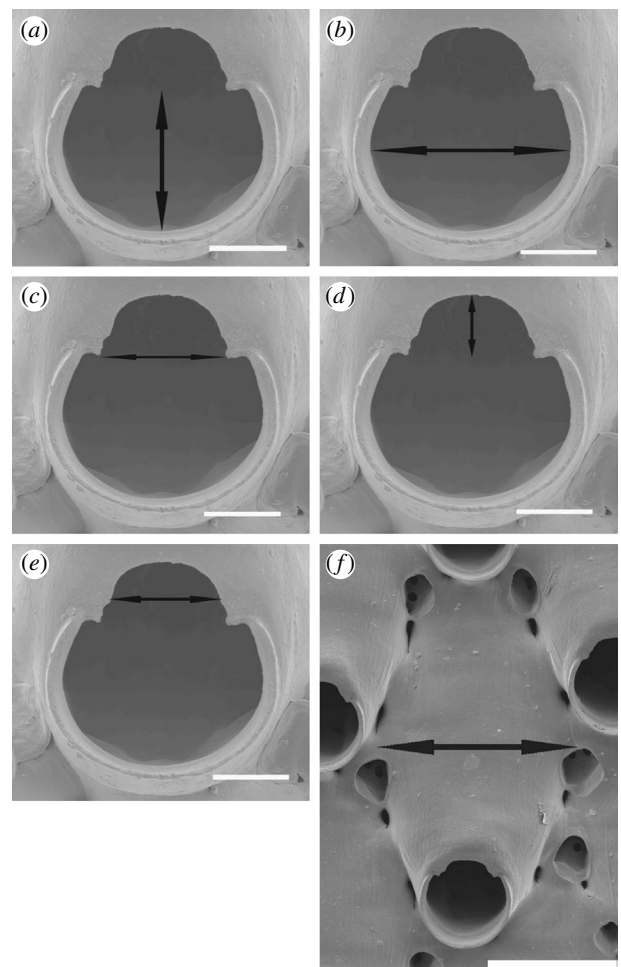


Figure 1. Morphological measurements taken from *Celleporella hyalina* zooids. Scanning electron micrographs taken from an autozooid in a *C. hyalina* colony collected from Dublin Harbour, Ireland. (a) Depth of orifice, (b) width of orifice, (c) distance between condyles, (d) depth of sinus, (e) width of sinus at 0.5 depth and (f) width of autozooid (length of zooid is normal to width); this micrograph also shows six adventitious pore chambers surrounding the autozooid. Scale bars, (a–e) 50 μ m and (f) 200 μ m.

(f) Morphological analysis

Classical taxonomy of the genus *Celleporella* focuses on the shape of both the autozooid and its orifice, here summarized by six linear measurements (figure 1), and on meristic characters including the number of adventitious pore chambers (budding points) surrounding the autozooid (figure 1f), which was also measured. Morphometry was based on the cloned F₁ progeny of native colonies (mentioned earlier), using ramets of standard age that had been grown under uniform conditions of temperature, salinity, photoperiod and food supply, thus avoiding potential variability associated with colonial development or environmental plasticity. Single measurements of each character were taken per autozooid, using 10 autozooids per colony and 3–10 colonies per population. Linear characters were expressed as residuals about their regressions on autozooid length, hence removing size effects on shape. Canonical variate analysis of character means per colony was performed using mtDNA lineage as the *a priori* grouping variable. Prior to inclusion in the analysis, each character was subject to a one-way analysis of variance to ensure that there was a significant variation among localities. Statistical analyses were performed using SPSS v. 12.0.

Table 1. Mating trials. (Names indicate the geographical origin of isolates and/or genetic lineage in molecular phylogeny (figure 2). ✓, successful brooding of embryos and release of viable larvae; X, absence of brooded embryos; NP, potential crosses not performed; *, one genotype brooded embryos but failed to release larvae. Control experiments occupy the diagonal, where the first symbol refers to the intrapopulation crosses indicating active reproduction and the number of offspring resulting from self-fertilization is given after the forward slash. Each mating trial used four colonies per locality except Oban and Woods Hole, where only two colonies could be used per trial. Larvae produced by Woods Hole test colonies (bottom row) are presumed to have resulted from self-fertilization.)

acting female	acting male					
	Amlwch	Maine	California	Alaska	Oban	Woods Hole
Amlwch (Northeast Atlantic)	✓/*	X*	X*	X*	X	X
Maine (Northwest Atlantic)	X	✓/0	X	X	X	X
California	X	X	✓/0	X	X	X
Alaska (Northeast Pacific)	X	X	X	✓/0	X	NP
Oban (Celtic)	X	X	X	X	✓/0	NP
Woods Hole	✓	✓	✓	NP	NP	✓/2

(g) Breeding experiments

Our purpose was to assess reproductive isolation among a geographically representative sample of clades. The protracted mating protocol (below) meant that only six clades could be tested in all the paired combinations within the available 3 years. The clades chosen for mating were among the most fecund under our culture conditions, with certain others proving difficult to bring into reproductive condition (see §4). Four replicates were staged for each pair of populations tested (table 1). In each experiment, a pair of virgin ramets was placed in a 300 ml glass jar and checked weekly for the presence of embryos. Ramets found to be brooding were cultured in isolation for a further two to three months and monitored for larval production. To verify reproductive competence, a control ramet of each genotype was paired with another genotype from its own lineage and monitored for larval output. To monitor self-fertilization, a second control ramet of each genotype was cultured in isolation and monitored for larval output (table 1).

3. RESULTS

(a) Mitochondrial DNA phylogeny

A fragment of 576 bp of the mitochondrial COI gene was sequenced for 176 colonies from 33 sites (see the electronic supplementary material) encompassing the known geographical range of *C. hyalina*. After removing 77 identical haplotypes, 99 distinct haplotypes formed the mtDNA database. There were 293 invariable (monomorphic) and 283 variable (polymorphic) sites, with 255 of them being parsimony informative. Out of the 136 segregating sites, 10 involved replacement changes (three of them were found only in the outgroup).

All the phylogenetic reconstructions of the COI alignment agreed in the topology of the main branches and showed numerous deep and strongly supported lineages (figure 2a), with a range of (uncorrected) sequence divergence of 8–20%.

The exact method of analysis of barcode data (e.g. ratio of intra- to inter-group diversity or MOTU thresholds) determines to a large degree the number of groups recovered, and the underlying nature of the genetic diversity sampled can also play a large role in determining the number of discrete groups detected (Floyd *et al.* 2002; Hebert *et al.* 2004b; Meyer & Paulay 2005). Here, we have taken a conservative approach where we do not try to

determine the exact number of organismal groups, but rather identify major lineages supported strongly by the phylogenies (70% or more maximum-likelihood bootstrap support) and that correspond to distinct geographical regions in order to test reproductive compatibility. This conservative approach almost certainly underestimates the number of species groups, but does identify 10 deep lineages that are likely to be consistently recovered by most approaches.

The different lineages had discrete ranges and mainly allopatric or parapatric distributions (figure 2a,c); therefore, we assigned them geographical names accordingly. We found a 'Northwest Atlantic' lineage in Maine, Nova Scotia and Newfoundland; an 'Arctic I' lineage in Svalbard and the Barents Sea; an 'Arctic II' lineage in Svalbard; a genetically very diverse 'Northeast Pacific' lineage in Alaska, the Pacific Coast of Canada and Svalbard; a monophyletic lineage including a 'North Polar' lineage, with two sublineages in Fennoscandia and the Arctic; a 'South Polar' lineage in the Chilean Magellanic region; a 'Northeast Atlantic' lineage from Spain to the UK and Ireland to Iceland; a 'Celtic' lineage with three sublineages in Scotland and Northwest Ireland; a highly subdivided 'South American' lineage in Chile and Argentina; and its sister group, a 'Californian' lineage, on the Californian coast. The COI dataset analysis does not permit us to resolve the deep phylogenetic relationships of all the lineages, although the sister status of some of the groups is well supported by bootstrap. The Celtic and Northeast Atlantic lineages form a well-supported clade, as well as the Arctic I and Northwest Atlantic, and the Californian and 'South American' clades. When we define lineages according to deep genetic divergence and allopatric distribution, the Atlantic Ocean contains eight lineages and the Pacific Ocean three lineages, the Northern Hemisphere contains eight lineages and the Southern Hemisphere two lineages. The Woods Hole isolate appeared as a separate lineage from the rest.

(b) Nuclear DNA phylogeny

We obtained EF-1 α sequence from 24 colonies representing the *C. hyalina* mtDNA COI lineages and four colonies representing outgroups (*Celleporella angusta* and *Celleporella carolinensis*). Amplified fragments of the two exons were used in the analysis totalling 452 bp for 28 sequences including outgroups. The trees retrieved by the three

phylogenetic methods used were remarkably similar and substantially concordant with the COI analysis (figure 2). The Northwest Atlantic, Northeast Pacific, North Polar, Californian and South American lineages were deep and well supported by bootstrap. Celtic and Northeast Atlantic lineages appeared as sister taxa with high bootstrap support. Sequences from all the samples recognized as *C. hyalina* form a strongly supported monophyletic clade except for the isolate from Woods Hole, which appears as a sister group to the rest of *C. hyalina* lineages.

(c) Reproductive isolation

We tested reproductive compatibility in 26 pairwise combinations of populations representing six of the geographical lineages (table 1). Complete reproductive isolation was found in 23 tests, while three tests involving the Woods Hole colonies acting as female gave inconclusive results, as these colonies proved capable of self-fertilization.

(d) Morphological analyses

Canonical variate analysis of operational taxonomic units (OTUs) defined *a priori* from mtDNA divergence produced five interpretable axes (CV1–CV5), which respectively accounted for 58.4, 25.5, 7.4, 5.2 and 3.0% of the data variance and represented the number of pore chambers, sinus depth, sinus size, autozooid shape and aperture size (orifice+sinus). Colony scores on paired axes produced coherent, but overlapping clusters corresponding to OTUs (figure 3). CV1 distinguished California and South America (more pore chambers) from the other lineages, CV2 separated North Polar (shallow sinus) from Woods Hole (deep sinus), CV3 separated Northwest Atlantic (larger sinus) from Celtic (smaller sinus), CV4 separated Northeast Pacific, California and Northwest Atlantic (narrower autozooid) from South America (wider autozooid) and CV5 separated South America (smaller aperture) from California (larger aperture). All the paired comparisons of centroids (Mahalanobis distances) were statistically significant after Bonferroni correction, but the probability of correct assignment of colonies to OTUs varied as follows: Northeast Atlantic 86.7% (90 colonies); North Polar 100% (36 colonies); Northeast Pacific 83.3% (18 colonies); California 100% (10 colonies); Woods Hole 90% (10 colonies); Northwest Atlantic 90% (10 colonies); Celtic 75% (four colonies); and South America 100% (six colonies). Morphological differentiation of OTUs therefore was statistically demonstrable, but with varying degree of uncertainty owing to the continuous variation and overlap in measured characters.

4. DISCUSSION

We have shown that the taxon previously known as *C. hyalina* comprised a large number of deep genetic lineages supported by mtDNA and nuclear DNA analysis. These lineages, retrieved by DNA barcoding, are difficult or impossible to distinguish using morphological methods alone, yet, judging from the sample tested, they are reproductively isolated and represent distinct biological species.

Given the wide range of genetic distances found between clusters, we used a preliminary approach that grouped deep mtDNA lineages according to their geographical distributions. This initial discrimination tool defined a conservative number of taxa. The range of

inter-lineage COI sequence divergences (8–20%) is of the same scale or greater than that typically encountered between sister species in other taxa (Avise & Johns 1999; Avise 2000; Hebert *et al.* 2003a). The topology and the level of divergence obtained from the analysis of the nuclear gene (EF-1 α) were largely concordant with the mtDNA tree (figure 2b). The molecular evidence therefore strongly suggests that these geographical lineages have had long independent evolutionary histories.

Complete reproductive isolation was found among geographical isolates representing selected mtDNA lineages. DNA-recovered lineages therefore would be incapable of interbreeding if events had brought them into secondary contact, indicating that *C. hyalina* comprised a large number of biological species, adding to the growing number of cryptic species complexes to be found in the oceans (Levitan *et al.* 2004).

Haplotypes belonging to known biogeographically distinct lineages were recorded in sympatry from East Svalbard (two lineages), Barents Sea (two lineages), Scotland (two lineages), Shetland (two lineages), Western Ireland (two lineages) and Nova Scotia (two lineages). On the other hand, certain other lineages appear to be too differentially adapted to their local environments for potential sympatry. Several isolates from the high Arctic with a mean summer inshore sea temperature of 6°C could not be grown to sexual maturity in our standardized laboratory environment of 15°C, while isolates from yet other regions grew at quite different rates. Owing to such apparent local adaptation, failure to recognize these cryptic species would invalidate any environmental study that treated all as a single unit, as well as overlooking an additional dimension to diversity of community structure.

The deep levels of genetic substructuring (figure 2) suggest that most major lineages of *C. hyalina* retrieved by the COI analysis contain more than one species. On smaller geographical scales, such as the Northeast Atlantic, we detected phylogeographical structuring of groups differing by *ca* 2% (A. Gómez, R. N. Hughes, P. J. Wright, G. R. Carvalho & D. H. Lunt, unpublished data). Although genetic divergence is sensitive to numerous factors such as population size, number of generations and reproductive mode (Avise 2000), our data were similar in both pattern and divergence level to diverse taxa, which were structured during the Pleistocene glacial cycles (Hewitt 1996; Avise 2000; Wares & Cunningham 2001). Indeed, the results of breeding experiments within the Northeast Atlantic lineage (P. J. Wright & R. N. Hughes, unpublished data) support further species consistent with such phylogeographical units. Such a pattern of substructuring is found in several other regional lineages. One of the most striking patterns occurs on the South Pacific coast, where the South American lineage comprised six to nine parapatric, genetic clusters, one of them corresponding to the previously described species *Celleporella retiformis* (Moyano 1986). Preliminary crosses, involving populations representing some of the South American clusters, indicate reproductive isolation (table 2). If these intra-lineage groups represent what most biologists might consider to be species, it follows that marine biodiversity would be even more extensively underestimated than we have presented here.

Although our results demonstrate some congruence between genetic/biological groups and morphological

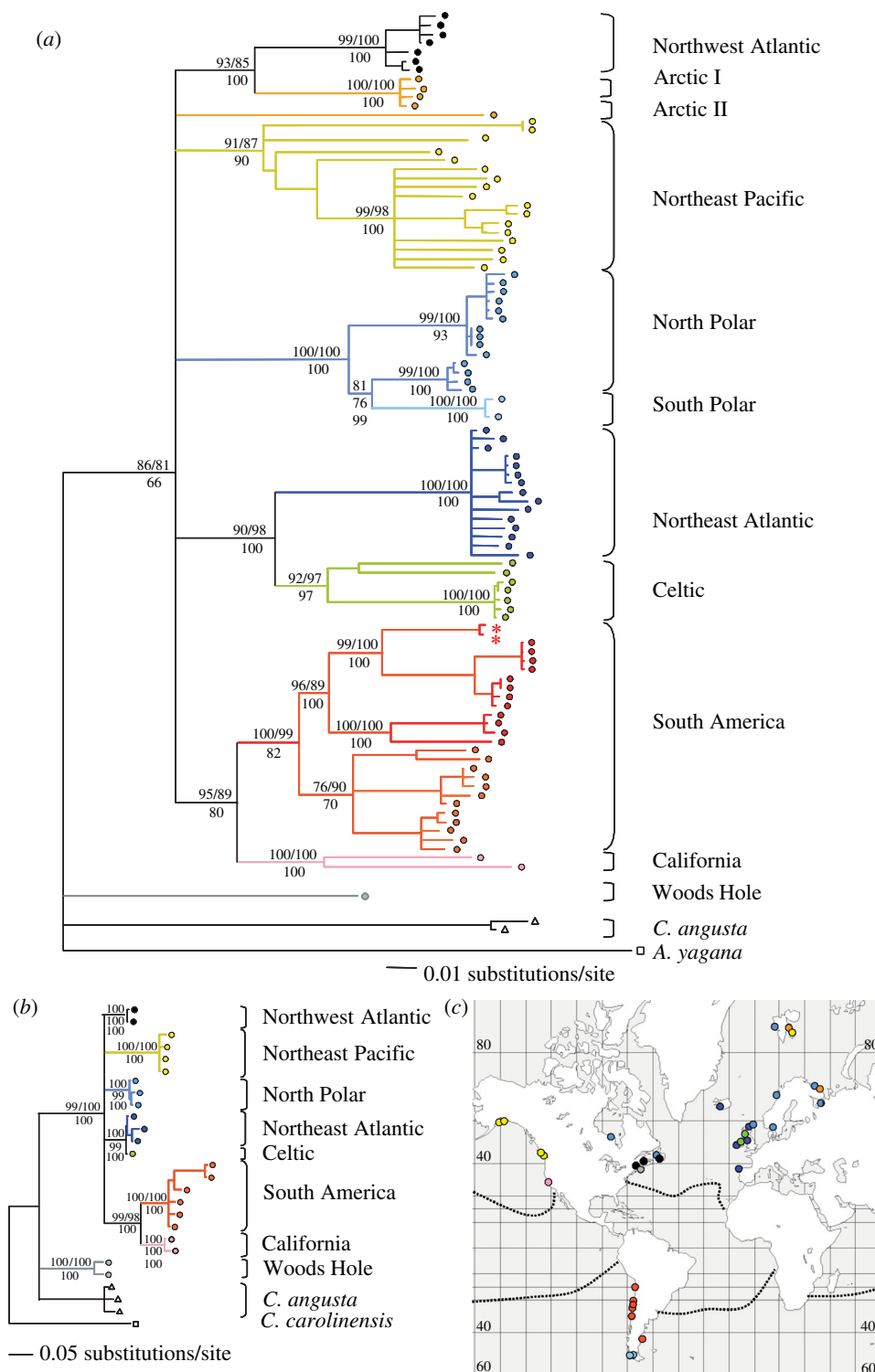


Figure 2. Phylogeny and geographical distribution of the major *C. hyalina* lineages. (a) Maximum-likelihood tree of COI haplotype data. (b) Maximum-likelihood tree of EF-1 α haplotype data. In both trees, nodes having less than 70% bootstrap support were collapsed. Per cent bootstrap support values for maximum-likelihood, neighbour-joining (1000 pseudoreplicates each) and posterior probabilities for Bayesian inference (2 000 000 generations) are shown on selected nodes. Individuals traditionally described as *C. hyalina* are marked with circles coloured according to the geographical region listed to the right. *Celleporella retiformis* is shown with asterisks. Outgroups are *C. angusta* (triangles) and *Astrothoa yagana* (square) in the COI tree and *C. carolinensis* (square) in the EF-1 α tree. (c) Map of the sample locations included in the genetic analysis. Coloured circles indicate the major lineage according to the phylogenetic analysis. Dotted lines indicate the limits of the temperate oceans (20°C isotherm).

differentiation in *C. hyalina*, identification from morphology alone is uncertain; therefore, it is an impractical means of assessing the biodiversity of such taxa. Similar analyses of the cheilostome genus *Stylopoma* (Jackson & Cheetham 1990, 1994) revealed three new,

subtly distinct morphospecies that had been overlooked by classical taxonomy. Such results are predictable, since although biological speciation may be indirectly associated with visible characters shaped by drift or local adaptation (Cheetham *et al.* 1993), morphological differentiation is

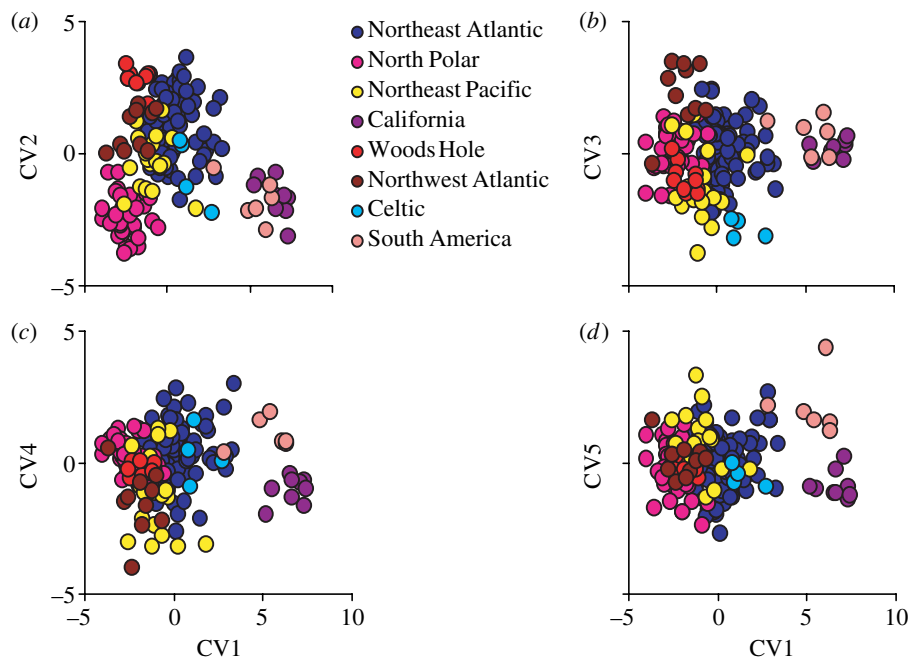


Figure 3. Ordination of *C. hyalina* colony morphology on paired axes from canonical variate analysis using mtDNA lineage as the *a priori* criterion for group membership. CV1–CV5 are the first five axes generated by canonical variate analysis of morphological data.

irrelevant to, and therefore unrefined by, a mating system based entirely on cellular recognition (Knowlton 1993).

The correlation between molecular groups and biological species is important if morphology does not also covary in a discrete, easily quantifiable fashion, since DNA barcoding may be the only option for species identification, other than performing time-consuming breeding experiments. Through DNA sequence analysis, we have identified deeply divergent lineages, a substantial sample of which are the confirmed biological species, but this has not been mirrored by salient morphological differentiation within the *C. hyalina* complex (figure 3). Furthermore, canalization of phenotypic plasticity by uniform experimental conditions is likely to have underestimated the extent of morphological overlap evident in the wild. Some groups that have previously been described as cryptic species have subsequently been found to be morphologically distinguishable and renamed as ‘pseudo-cryptic’ species (Knowlton 1993; Sáez *et al.* 2003). Despite rearing the *C. hyalina* lineages in a standardized laboratory environment, taking detailed scanning electron micrograph measurements of important characters in hippothoid taxonomy and applying multivariate analyses, we could not discriminate unequivocally the morphology of the major lineages from one another. Therefore, it is clear that morphological analysis alone has limited utility for species identification in this complex and that DNA barcoding provides a robust and biologically significant addition.

Although it is evident from figure 2 that there has been a considerable divergence in allopatry, at least two long-distance (trans-oceanic) dispersal events are supported by our data. The first event involved colonization of Svalbard by genotypes from coastal Alaska. The fact that Alaska and Svalbard shared an identical haplotype indicates that this dispersal event must have been recent and possibly anthropogenic. The other event is older and more interesting, as it gives evidence for transtropical exchange

Table 2. Mating trials between Chilean *C. hyalina* populations within the South American lineage. (Symbols as in table 1. Two replicate mating trials were performed within the colonies from Ramuntcho. The South American sublineages are indicated by roman numerals.)

acting female	acting male		
	Antofagasta	Ramuntcho	Las Cruces
Antofagasta (I)	✓/0	X	NP
Ramuntcho (I)	✓/3	✓/0	X
Las Cruces (II)	NP	X	✓/0

in antitropical taxa (Darling *et al.* 2000). This transtropical colonization episode involves interchange between the North Atlantic and the Magellanic region evidenced by the geographical distribution of two sister taxa, namely a sublineage of the North Polar lineage and the South Polar lineage. The fact that the South Polar lineage is nested within a diverse North Polar lineage suggests that the colonization occurred from North to South. The split between the South Polar and its sister clade (with 6.4% uncorrected average sequence divergence) suggests that it predates any possible anthropogenic transport. For other systems, it has been proposed that the closure of the Panama strait (Central American Seaway), approximately 3 million years ago, triggered a number of episodes of North–South transequatorial migrations (Lindberg 1991; Vermeij 1992; Hilbish *et al.* 2000).

We have shown by DNA barcoding that the nominally cosmopolitan bryozoan *C. hyalina* comprised numerous deep genetic lineages distributed largely allopatrically in the world’s oceans. These lineages are reproductively incompatible and represent biological species or groups of species. Biodiversity has been grossly underestimated in this model system, and if the major lineages are shown to contain further species (as our data suggest), this may be

by one or two orders of magnitude. We have further shown that these lineages cannot confidently be distinguished by morphological analysis alone even when environmental plasticity is reduced by standardized laboratory culture, and therefore they are cryptic species.

In conclusion, DNA barcoding clearly identifies biologically meaningful groups in the *C. hyalina* complex. Although most nominal species are more restricted in distribution than *C. hyalina*, some 15% of cheilostome bryozoans found in British waters are known to have geographical ranges of at least 30 000 km (Watts *et al.* 1998), indicating that the present results may have much wider applicability. Moreover, like sponges, corals and other sessile invertebrates, bryozoans such as *C. hyalina* enrich communities that dominate subtidal firm substrata globally (Jackson 1977), thereby underpinning benthic community structure. Failure to recognize cryptic speciation among sessile benthos therefore may seriously underestimate marine biodiversity (Knowlton 1993, 2000), as well as impeding attempts to predict the response of marine benthos to environmental change.

This project was funded by the Natural Environment Research Council (UK). We thank many colleagues who helped in sampling collection, laboratory work and discussion of the manuscript, including: J. César-Aldariz, D. Barnes, A. Clarke, M. Dick, H. Hughes, J. López-Gappa, O. Reverter-Gil, V. Gontar, C. Griffiths, A. Grisichenko, N. Aldo-Landoni, N. Knowlton, P. Kuklinski, S. Mawatari, C. Mitchell, S. Morley, B. Okamura, A. Ostrovski, J. Pearse, C. Pook, J. Porter, E. Fernández-Pulpeiro, J. Ryland, B. Scheibling, N. Shunatova, A. Waschenbach, A. Navarrete, D. Weetman, G. Turner and R. Thorpe.

REFERENCES

- Avise, J. C. 2000 *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard University Press.
- Avise, J. C. & Johns, G. C. 1999 Proposal for a standardized temporal scheme of biological classification for extant species. *Proc. Natl Acad. Sci. USA* **96**, 7358–7363. (doi:10.1073/pnas.96.13.7358)
- Avise, J. C., Nelson, W. S. & Sugita, H. 1994 A speciation history of living fossils—molecular evolutionary patterns in horseshoe crabs. *Evolution* **48**, 1986–2001. (doi:10.2307/2410522)
- Blaxter, M. 2003 Counting angels with DNA. *Nature* **421**, 122–124. (doi:10.1038/421122a)
- 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.
- Cancino, J. M. & Hughes, R. N. 1988 The zooidal polymorphism and astogeny of *Celleporella hyalina* (Bryozoa, Cheilostomata). *J. Zool.* **215**, 167–181.
- Cheetham, A. H., Jackson, J. B. C. & Hayek, L. A. C. 1993 Quantitative genetics of Bryozoan phenotypic evolution. 1. Rate tests for random change versus selection in differentiation of living species. *Evolution* **47**, 1526–1538. (doi:10.2307/2410165)
- Darling, K. F., Wade, C. M., Stewart, I. A., Kroon, D., Dingle, R. & Brown, A. J. L. 2000 Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers. *Nature* **405**, 43–47. (doi:10.1038/35011002)
- Dodson, S. I., Grishanin, A. K., Gross, K. & Wyngaard, G. A. 2003 Morphological analysis of some cryptic species in the *Acanthocyclops vernalis* species complex from North America. *Hydrobiologia* **500**, 131–143. (doi:10.1023/A:1024678018090)
- Ebach, M. C. & Holdrege, C. 2005 DNA barcoding is no substitute for taxonomy. *Nature* **434**, 697. (doi:10.1038/434697b)
- Eyuaem, A. & Blaxter, M. 2003 Comparison of biological, molecular, and morphological methods of species identification in a set of cultured *Panagrolaimus* isolates. *J. Nematol.* **35**, 119–128.
- Floyd, R., Abebe, E., Papert, A. & Blaxter, M. 2002 Molecular barcodes for soil nematode identification. *Mol. Ecol.* **11**, 839–850. (doi:10.1046/j.1365-294X.2002.01485.x)
- 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.
- Givnish, T. J., Systma, K. J. & Sytsma, K. J. 2000 *Molecular evolution and adaptive radiation*. Cambridge, UK: Cambridge University Press.
- Gómez, A., Serra, M., Carvalho, G. R. & Lunt, D. H. 2002 Speciation in ancient cryptic species complexes: evidence from the molecular phylogeny of *Brachionus plicatilis* (Rotifera). *Evolution* **56**, 1431–1444. (doi:10.1554/0014-3820(2002)056[1431:SIACSC]2.0.CO;2)
- Guindon, S. & Gasquel, O. 2003 A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704. (doi:10.1080/10635150390235520)
- Hajibabaei, M., DeWaard, J. R., Ivanova, N. V., Ratnasingham, S., Dooh, R. T., Kirk, S. L., Mackie, P. M. & Hebert, P. D. N. 2005 Critical factors for assembling a high volume of DNA barcodes. *Phil. Trans. R. Soc. B* **360**, 1959–1967. (doi:10.1098/rstb.2005.1727)
- Hebert, P. D. N. 1998 Variable environments and evolutionary diversification in inland waters. In *Advances in molecular ecology* (ed. G. R. Carvalho), pp. 267–290. Amsterdam, The Netherlands: IOS Press.
- Hebert, P. D. N. & Gregory, T. R. 2005 The promise of DNA barcoding for taxonomy. *Syst. Biol.* **54**, 852–859. (doi:10.1080/10635150500354886)
- Hebert, P. D. N., Cywinska, A., Ball, S. L. & DeWaard, J. R. 2003a Biological identifications through DNA barcodes. *Proc. R. Soc. B* **270**, 313–321. (doi:10.1098/rspb.2002.2218)
- Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. 2003b Barcoding animal life: cytochrome *c* oxidase subunit I divergences among closely related species. *Proc. R. Soc. B* **270**, S96–S99. (doi:10.1098/rsbl.2003.0025)
- Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. 2004a Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl Acad. Sci. USA* **101**, 14 812–14 817. (doi:10.1073/pnas.0406166101)
- Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. 2004b Identification of birds through DNA barcodes. *PLoS Biol.* **2**, 1657–1663. (doi:10.1371/journal.pbio.0020312)
- Hewitt, G. M. 1996 Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* **58**, 247–276. (doi:10.1006/bjil.1996.0035)
- Hilbish, T. J., Mullinax, A., Dolven, S. I., Meyer, A., Koehn, R. K. & Rawson, P. D. 2000 Origin of the antitropical distribution pattern in marine mussels (*Mytilus* spp.): routes and timing of transequatorial migration. *Mar. Biol.* **136**, 69–77. (doi:10.1007/s002270050010)

- Hoare, K. & Hughes, R. N. 2001 Inbreeding and hermaphroditism in the sessile, brooding bryozoan *Celleporella hyalina*. *Mar. Biol.* **139**, 147–162. (doi:10.1007/s002270100566)
- Hogg, I. D. & Hebert, P. D. N. 2004 Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can. J. Zool. Revue Canadienne De Zoologie* **82**, 749–754. (doi:10.1139/z04-041)
- Hughes, R. N., Wright, P. & Manríquez, P. H. 2001 Predominance of obligate outbreeding in the simultaneous hermaphrodite *Celleporella hyalina* sensu lato. In *XII International bryozoology association conference* (ed. P. N. W. Jackson, C. J. Butler & M. E. S. Jones), pp. 159–162. Dublin, Ireland: A.A. Balkema Publishers.
- Hughes, R. N., Manríquez, P. H. & Bishop, J. D. D. 2002 Female investment is retarded pending reception of allosperm in a hermaphroditic colonial invertebrate. *Proc. Natl Acad. Sci. USA* **99**, 14 884–14 886. (doi:10.1073/pnas.162339699)
- Jackson, J. B. C. 1977 Competition on marine hard substrata: the adaptive significance of solitary and colonial strategies. *Am. Nat.* **111**, 743–768. (doi:10.1086/283203)
- Jackson, J. B. C. & Cheetham, A. H. 1990 Evolutionary significance of morphospecies—a test with cheilostome Bryozoa. *Science* **248**, 579–583.
- Jackson, J. B. C. & Cheetham, A. H. 1994 Phylogeny reconstruction and the tempo of speciation in cheilostome Bryozoa. *Paleobiology* **20**, 407–423.
- Jackson, J. B. C. & Cheetham, A. H. 1999 Tempo and mode of speciation in the sea. *Trends Ecol. Evol.* **14**, 72–77. (doi:10.1016/S0169-5347(98)01504-3)
- Johnson, N. K. & Cicero, C. 2004 New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. *Evolution* **58**, 1122–1130. (doi:10.1554/03-283)
- Knowlton, N. 1993 Sibling species in the sea. *Annu. Rev. Ecol. Syst.* **24**, 189–216. (doi:10.1146/annurev.es.24.110193.001201)
- Knowlton, N. 2000 Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* **420**, 73–90. (doi:10.1023/A:1003933603879)
- Kucera, M. & Darling, K. F. 2002 Cryptic species of planktonic foraminifera: their effect on palaeoceanographic reconstructions. *Phil. Trans. R. Soc. A* **360**, 695–718. (doi:10.1098/rsta.2001.0962)
- Levitan, D. R., Fukami, H., Jara, J., Kline, D., McGovern, T. M., McGhee, K. E., Swanson, C. A. & Knowlton, N. 2004 Mechanisms of reproductive isolation among sympatric broadcast-spawning corals of the *Montastraea annularis* species complex. *Evolution* **58**, 308–323. (doi:10.1554/02-700)
- Lindberg, D. R. 1991 Marine biotic interchange between the Northern and Southern hemispheres. *Paleobiology* **17**, 308–324.
- Lorenz, J. G., Jackson, W. E., Beck, J. C. & Hanner, R. 2005 The problems and promise of DNA barcodes for species diagnosis of primate biomaterials. *Phil. Trans. R. Soc. B* **360**, 1869–1877. (doi:10.1098/rstb.2005.1718)
- Manríquez, P. H., Hughes, R. N. & Bishop, J. D. D. 2001 Age-dependent loss of fertility in water-borne sperm of the bryozoan *Celleporella hyalina*. *Mar. Ecol. Prog. Ser.* **224**, 87–92.
- Mayr, E. 1942 *Systematics and the origin of species*. New York, NY: Columbia University Press.
- Meyer, C. P. & Paulay, G. 2005 DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol.* **3**, e422.
- Moon-van der Staay, S. Y., DeWachter, R. & Vaulot, D. 2001 Oceanic 18S sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409**, 607–610. (doi:10.1038/35054541)
- Moritz, C. & Cicero, C. 2004 DNA barcoding: promise and pitfalls. *PLoS Biol.* **2**, 1529–1531. (doi:10.1371/journal.pbio.0020354)
- Moyano, H. 1986 Bryozoa marinos chilenos VI. Cheilostomata Hippothoidae: south eastern Pacific species. *Boletín de la Sociedad de Biología de Concepción (Chile)* **57**, 89–153.
- Posada, D. & Crandall, K. A. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818. (doi:10.1093/bioinformatics/14.9.817)
- Ronquist, F. & Huelsenbeck, J. P. 2003 MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574. (doi:10.1093/bioinformatics/btg180)
- Sáez, A. G. & Lozano, E. 2005 Body doubles. *Nature* **433**, 111. (doi:10.1038/433111a)
- Sáez, A. G., Probert, I., Geisen, M., Quinn, P., Young, J. R. & Medlin, L. K. 2003 Pseudo-cryptic speciation in coccolithophores. *Proc. Natl Acad. Sci. USA* **100**, 7163–7168. (doi:10.1073/pnas.1132069100)
- Schander, C. & Willassen, E. 2005 What can biological barcoding do for marine biology? *Mar. Biol. Res.* **1**, 79–83.
- Swofford, D. 1998 *PAUP*: phylogenetic analysis using parsimony (* and other methods)*, v. 4. Sunderland, MA: Sinauer Associates.
- Vermeij, G. J. 1992 Trans-equatorial connections between biotas in the temperate Eastern Atlantic. *Mar. Biol.* **112**, 343–348. (doi:10.1007/BF00702481)
- Wares, J. P. & Cunningham, C. W. 2001 Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution* **55**, 2455–2469. (doi:10.1554/0014-3820(2001)055[2455:PAHEOT]2.0.CO;2)
- Watts, P. C., Thorpe, J. P. & Taylor, P. D. 1998 Natural and anthropogenic dispersal mechanisms in the marine environment: a study using cheilostome Bryozoa. *Phil. Trans. R. Soc. B* **353**, 453–464. (doi:10.1098/rstb.1998.0222)
- Will, K. W. & Rubinoff, D. 2004 Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics Int. J. Willi Hennig Soc.* **20**, 47–55. (doi:10.1111/j.1096-0031.2003.00008.x)