Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 44 (2007) 240–254
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Species boundaries and phylogenetic relationships within the green algal genus *Codium* (Bryopsidales) based on plastid DNA sequences

Heroen Verbruggen [a,*], Frederik Leliaert [a], Christine A. Maggs [b], Satoshi Shimada [c], Tom Schils [a], Jim Provan [b], David Booth [b], Sue Murphy [b], Olivier De Clerck [a], Diane S. Littler [d], Mark M. Littler [d], Eric Coppejans [a]

[a] *Phycology Research Group and Center for Molecular Phylogenetics and Evolution, Ghent University, Krijgslaan 281 (S8), B-9000 Gent, Belgium*
[b] *School of Biological Sciences, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL, UK*
[c] *Center for Advanced Science and Technology, Hokkaido University, Sapporo 060-0810, Japan*
[d] *US National Herbarium, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA*

Received 26 July 2006; revised 6 December 2006; accepted 10 January 2007
Available online 31 January 2007

Abstract

Despite the potential model role of the green algal genus *Codium* for studies of marine speciation and evolution, there have been difficulties with species delimitation and a molecular phylogenetic framework was lacking. In the present study, 74 evolutionarily significant units (ESUs) are delimited using 227 *rbc*L exon 1 sequences obtained from specimens collected throughout the genus' range. Several morpho-species were shown to be poorly defined, with some clearly in need of lumping and others containing pseudo-cryptic diversity. A phylogenetic hypothesis of 72 *Codium* ESUs is inferred from *rbc*L exon 1 and *rps*3–*rpl*16 sequence data using a conventional nucleotide substitution model (GTR + Γ + I), a codon position model and a covariotide (covarion) model, and the fit of a multitude of substitution models and alignment partitioning strategies to the sequence data is reported. Molecular clock tree rooting was carried out because outgroup rooting was probably affected by phylogenetic bias. Several aspects of the evolution of morphological features of *Codium* are discussed and the inferred phylogenetic hypothesis is used as a framework to study the biogeography of the genus, both at a global scale and within the Indian Ocean.

Keywords: Benthic marine algae; *Codium*; Marine biogeography; Molecular clock rooting; Morphological evolution; Outgroup rooting; Phylogenetic bias; Phylogeny; *rbc*L; Species delimitation; Taxonomy

1. Introduction

Within the marine green algae, there are few genera that can be used as a model for studies of speciation history, evolution and biogeography. The genus *Codium* constitutes an ideal example because it is distributed through much of the world's seas, shows a wide variety of forms and occurs in various habitats. It contains approximately 150 species. The form of the algal body (thallus) is the most apparent and variable attribute. *Codium* thalli can spread out over hard surfaces as mats, form spheres or grow upright, either unbranched and finger-like, or branched, with cylindrical or flattened branches (Figs. 1A–E). Anatomically, a *Codium* thallus is composed of a single, giant, branched tubular cell containing multiple nuclei, the branches commonly being called siphons. The center of the thallus (the medulla) consists of an entangled mesh of siphons, whereas in the surrounding cortex, the siphons are closely adjoined and swollen into utricles (Fig. 1F). The utricles occur in a wide array of forms, varying in size, shape and composition (Figs. 1G–I), with gametangia and/or hairs borne along their sides (Figs. 1G–I). *Codium* is found in marine habitats ranging from rocky coasts exposed to full wave-forces to

* Corresponding author. Fax: +32 9 264 8599.
E-mail address: heroen.verbruggen@ugent.be (H. Verbruggen).

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.01.009

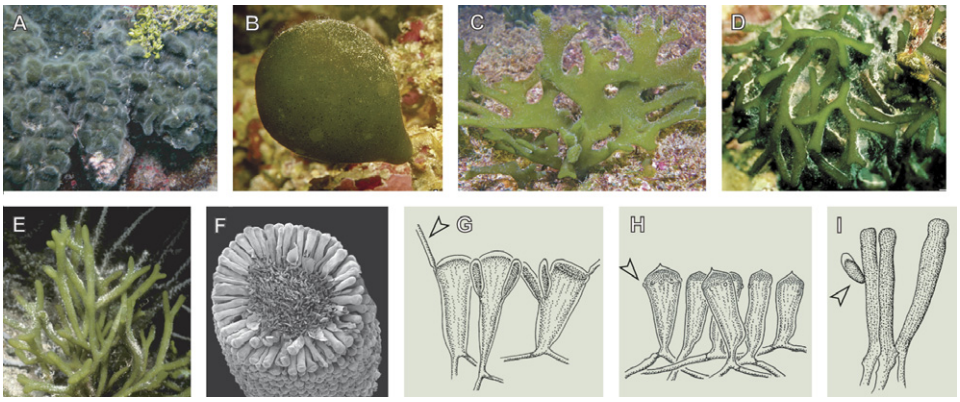

Fig. 1. Morphological diversity of *Codium*. (A) Mat-forming thallus. (B) Spherical thallus. (C) Erect thallus with flattened branches. (D) Branched thallus with a sprawling habit. (E) Erect thallus with cylindrical branches. (F) Cross-section through cylindrical branch showing the central medulla composed of a disorganized mesh of siphons, surrounded by a cortex composed of a uniform layer of utricles. (G) Club-shaped utricles with utricle hairs (arrow) and gametangia. (H) Club-shaped utricles with a pointed tip (mucron) and hair scars (arrow). (I) Cylindrical utricles with a gametangium (arrow).

calm lagoons, from intertidal habitats to deep reefs, from arctic to tropical waters and from eutrophic estuaries to nutrient-depleted coral reefs.

For the last two decades, *Codium* has been in the public and scientific spotlight because of the invasive, bloom-forming nature of certain species. *Codium fragile* subspecies *tomentosoides* is the most invasive seaweed in the world, being unintentionally spread around the globe with cultured shellfish (Trowbridge, 1998; Nyberg and Wallentinus, 2005). Another species, *C. isthmocladum*, forms harmful blooms on South Florida reefs in conjunction with increased eutrophication (Lapointe et al., 2005a,b). Both species can damage shellfish beds and perturb native communities and massive amounts of rotting thalli can smother shores. On a more positive note, *Codium* species are used as food for cultured abalone, are consumed by humans, and are a source of bioactive compounds among which are potential anti-cancer agents and antibiotics.

Codium has served as a model organism for studies of algal physiology and ecophysiology, heavy metal accumulation and bioactive compounds (Trowbridge, 1998). Its potential model role in studies of evolution and speciation has been much less explored. Nonetheless, *Codium* has been the subject of several systematic and biogeographic studies (e.g. Schmidt, 1923; Lucas, 1935; Silva, 1951, 1959, 1960, 1962), resulting in a classification of 2 subgenera and 5 sections based primarily on thallus habit (Appendix 1). Distinction between species within the sections is achieved through utricle anatomy and more subtle differences in thallus habit.

Morphological species delimitation tends to be problematic within the algae: many cases of erroneous species boundaries and cryptic species diversity are being disclosed by application of molecular phylogenetic methods and exploration of different species concepts (e.g. Famà et al., 2002; Kooistra, 2002; van der Strate et al., 2002; Zuccarello and West, 2003). As a consequence, pleas for molecular species delimitation are beginning to crop up in the phycological literature (Saunders and Lehmkuhl, 2005; Verbruggen et al., 2005a,b). *Codium* is no exception as far as problematic species delimitation is concerned. To our knowledge, no crossing studies have been carried out, so that the biological species concept has not yet been explored in this genus. Furthermore, specimens can be morphologically intermediate or show imperfect resemblance to described species. Consequently, there is little compelling evidence for the current species boundaries in *Codium*.

Despite the fact that *Codium* is a model organism for a spectrum of physiological and ecological studies, it lacks a comprehensive and objective phylogenetic framework. The earliest evolutionary hypotheses were based on morphological characters. Schmidt (1923) hypothesized that globular and erect habits have evolved from primitive mat-forming ancestors. These views have been maintained and corroborated by most morpho-taxonomists throughout the 20th century. Additionally, Silva (1954) posited a phylogenetic hypothesis based on anatomical characters. Shimada et al. (2004) published the first molecular phylogenetic study focusing specifically on *Codium*. They sequenced the first exon of the large RuBisCo subunit (*rbc*L) of a considerable number of specimens belonging to 17 Japanese species and concluded that this marker was suitable for distinguishing between species and that mat-forming and erect species (representing the two traditionally recognized subgenera) were not reciprocally monophyletic.

Codium, although widely distributed, has its largest species diversity in the subtropical regions, with several cases of disjunct distributions of individual or morphologically similar species, and thus serves as a model to investigate biogeographic affinities. One of the most intriguing biogeographic patterns in the marine realm is the apparent affinity of the algal floras of distant subtropical regions (Arabian Sea, SE Africa, SW Australia and Japan). Biogeographic links between these regions, which feature rich algal floras and high endemism (Phillips, 2001; Schils and Wilson, 2006; Bolton et al., 2004), have been described (e.g. Joosten and Van den Hoek, 1986; Lüning, 1990; Norris and Aken,

1985; Schils and Coppejans, 2003; Wynne, 2000, 2004). Aside from the overall similarity of these regions' algal floras in terms of diversity and biomass, several species are common to all or some of them while absent from intervening tropical locations. Similarly, the distinct regions feature morphologically similar congenerics that are absent from the tropical seas separating them. Two possible explanations for the affinities between the algal floras of SW Australia and SE Africa have been delineated: (1) a common origin of the floras along the Cretaceous coast of Gondwanaland which became separated in a series of tectonic events (Hommersand, 1986); (2) Dispersal of species through the low latitudes of the Indian Ocean during Pliocene or Pleistocene periods of global cooling, which could also account for their occurrence in the Arabian Sea and Japan (Hommersand, 1986; Lüning, 1990). Alternatively, the apparent resemblance could be an artifact caused by convergent evolution as a response to similar environmental selection regimes. Silva (1962) also suggested a link between the *Codium* floras of Japan and the temperate Pacific coasts of N America (California and Baja), the North Pacific gyre acting as a dispersal vector (see also De Clerck et al., 2006; Hommersand, 1971; Lane et al., 2006).

The first goal of the present study is to achieve delimitation of evolutionarily significant units (ESUs) using DNA sequence data and compare the resultant compartmentalization with current taxonomic viewpoints. The second goal is to expand the current phylogenetic framework and interpret the results in light of the morphological evolutionary and biogeographic hypotheses described above.

2. Materials and methods

2.1. Sampling and morphology

We examined *Codium* collections covering most of the geographical range. The highest diversity of *Codium* species is found in transitional floras of subtropical and warm-temperate regions (Arabian Sea, Japan, South Africa, southern Australia and southern California—Baja) and to a certain extent our sampling efforts reflect this bias. Collections were preserved in silica gel or 95% ethanol for DNA analysis. Vouchers for morphological and anatomical analysis were pressed or wet preserved (95% ethanol or 5% formalin-seawater). Specimens were identified using local taxonomic treatises when possible (Burrows, 1991; Chihara, 1975; Dellow, 1952; Kraft, 2000; Nizamuddin, 2001; Pedroche et al., 2002; Silva, 1951, 1959, 1960; Silva and Womersley, 1956; Taylor, 1960; Van den heede and Coppejans, 1996; Yoshida, 1998), or on the basis of a close match to descriptions of specimens from elsewhere. Identifications are presented in Appendix 1. Eight external morphological and 11 anatomical characters were scored for each species in order to aid identifications and map morphological traits onto phylogenetic trees (see Appendix 2 for an exhaustive list).

2.2. DNA sequencing and alignments

DNA extraction followed a CTAB protocol modified from Doyle and Doyle (1987) or used the Qiagen DNeasy Plant Mini-preps (Qiagen Ltd., Crawley, UK). Two plastid markers were amplified in PCRs and directly sequenced. The first *rbc*L exon was amplified according to Shimada et al. (2004), with different primers for certain specimens (pos. 12–34, forward: 5′-AACTGAAACTAAAGCAGGT GCAG-3′; pos. 799–778, reverse: 5′-GCATRATAATAGG TACGCCRAA-3′). The *rps*3–*rpl*16 region (UCP6) was amplified according to Provan et al. (2004). PCR products were purified with the ExoSAP-IT kit (USB Europe GmBH, Staufen, Germany), and sequenced with an ABI Prism 3100 automated sequencer (Applied Biosystems, Foster City, CA) using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and the above-mentioned PCR primers and/or internal primers for *rbc*L (pos. 331–353, forward: 5′-GGWTCKGTTACWAATTTA TTTAC-3′; pos. 522–500, reverse: 5′-AATAGTACARCC TAATARTGGAC-3′). Some sequencing was outsourced to Macrogen (Seoul, Korea). In total, 227 *rbc*L exon 1 and 119 *rps*3–*rpl*16 sequences were generated and submitted to GenBank (Appendix 1). The 227 *rbc*L sequences include those previously reported by Shimada et al. (2004).

The *rbc*L sequences were all of equal length (735 bases); their alignment was straightforward and unambiguous. The coding regions of *rps*3 and *rpl*16 sequences could be readily aligned. Towards the 3′ terminus of *rps*3, sequences were considerably more variable and featured several codon indels. Some sequences featured a spacer between *rps*3 and *rpl*16, whereas in others *rps*3 and *rpl*16 showed overlap. The length of the sequences ranged from 354 to 404 bases. The indel-containing terminal part of *rps*3 and the spacer region were removed from the alignment, yielding an unambiguous alignment of 345 coding nucleotides. Two alignments were created. The first, which will be referred to as the ESU delimitation alignment, contained 227 *rbc*L exon 1 sequences. The second alignment, referred to as the concatenated alignment, consisted of concatenated *rbc*L exon 1 and *rps*3–*rpl*16 sequences of 72 *Codium* ESUs. Both alignments can be obtained from TreeBase and phycoweb.net.

2.3. Delimitation of ESUs using molecular data

The ESU delimitation alignment was subjected to Neighbor Joining (NJ) bootstrapping analysis in MEGA 3.1 (Kumar et al., 2004). The specifications of the analysis can be found in Appendix 3. In the bootstrap consensus tree, we looked for clusters of sequences (1) containing little intra-cluster sequence divergence, (2) receiving very high bootstrap support and (3) sitting on long branches. One specimen of each of these clusters, which we refer to as evolutionarily significant units (ESUs; Moritz, 1994), was used to construct the concatenated alignment, except for two

ESUs for which we were unable to obtain an *rps*3–*rpl*16 sequence.

2.4. Exploration of phylogenetic data

The amount of phylogenetic signal versus noise in the concatenated alignment (ingroup only) was assessed using two methods. First, the g_1 statistic, a measure of the skewness of tree length distribution, was calculated (Hillis and Huelsenbeck, 1992). The length of 1000 random trees was calculated using PAUP 4.0b10. Strongly left-skewed distributions ($g_1 < 0$) indicate that relatively few solutions exist near the shortest, optimal tree, implying significant phylogenetic structure in the data, whereas unskewed distributions ($g_1 = 0$) are typical for random datasets lacking phylogenetic structure. The g_1 value of the length distribution of the random trees was calculated under ML for the alignment as a whole and for each codon position separately, using GTR + Γ + I models with the parameters converged upon by Bayesian phylogenetic analyses (settings as below). The obtained g_1 statistics were compared to the threshold values in Hillis and Huelsenbeck (1992). Second, the I_{ss} statistic, a measure of substitution saturation in molecular phylogenetic datasets, was calculated for the dataset as a whole and for each of the codon positions separately. I_{ss} is derived from the amount of entropy in the data and needs to be compared to critical values for which simulation studies showed decreased accuracy (Xia et al., 2003). The DAMBE software (Xia and Xie, 2001) was used to calculate I_{ss} values and compare them against critical I_{ss} values for symmetric and asymmetric topologies (Xia et al., 2003). Since critical I_{ss} values depend on the number of taxa and the sequence length and hence are dataset-specific and impractical to tabulate, DAMBE samples one thousand random subsets of 4, 8, 16 and 32 sequences from the alignment and calculates I_{ss} for the subsets.

Comparison of substitution rates and base frequencies of the different genes and codon positions can aid in choosing appropriate models for phylogenetic inference. For example, large base frequency differences between genes would indicate partitioning the alignment accordingly and uncoupling the model's base frequency parameters between partitions. Site-specific substitution rates of the *rbc*L, *rps*3 and *rpl*16 genes were calculated under a Jukes–Cantor model using the HyPhy package (Pond et al., 2005). The reference topology was obtained by Bayesian analysis (MrBayes 3.1.2; Ronquist and Huelsenbeck, 2003) of the concatenated alignment using a GTR + Γ + I model, a single run of four chains, standard priors and two million generations of which the first million was discarded as burn-in.

2.5. Substitution model fitting and molecular phylogenetic analyses

The fit of different nucleotide substitution models to the concatenated alignment was examined as follows. First, a tree was inferred from the alignment using the GTR + Γ + I substitution model as specified above. This tree was used as the reference topology against which 61 different models were tested. These models included some conventional nucleotide substitution models and models in which the substitution rates and/or model parameters were uncoupled across codon positions and/or genes (e.g. Shapiro et al., 2006). The tested models are listed in Section 3. The likelihood of the tree was calculated under the different models using PAML (Yang, 1997). The Akaike Information Criterion (AIC), which penalizes complex models, was used to compare the fit of different models. Since the length of our alignment was relatively small to estimate all parameter values of highly complex models, the second order AIC_c, which includes an additional penalty for model complexity, was calculated in addition to AIC (Posada and Buckley, 2004). The fit of a covariotide model (allowing rate variation through time; Huelsenbeck, 2002) was compared to that of other models using the Bayes factor because the covariotide option is not available in PAML. The Bayes factor, calculated as the ratio between the marginal likelihoods of two competing models, can be used to evaluate how well the models approximate the processes generating the data (Huelsenbeck et al., 2004; Posada and Buckley, 2004). The Bayes factor is not a statistical test but cut-off values have been published to aid in their interpretation (Kass and Raftery, 1995; Nylander et al., 2004).

Phylogenetic inferences for the genus *Codium* were made from the concatenated alignment using Bayesian methods (MrBayes 3.1.2). Three analyses were performed. First, the unpartitioned dataset was analyzed using a single general time-reversible model with rate variation across sites and a proportion of invariable sites. This analysis is referred to as the GTR + Γ + I analysis. Second, the dataset was divided into two partitions, corresponding to the first plus second and the third codon positions, and GTR + Γ + I models were applied to each of the partitions. Rates and all model parameters were uncoupled between the partitions. This analysis is referred to as the codon position analysis. Third, the codon position analysis was carried out with the covariotide option, allowing substitution rate variation across lineages. This analysis is referred to as the covariotide analysis. All analyses were run for five million generations, with two parallel runs of four chains each, the default priors of MrBayes 3.1.2, and trees and parameter estimates saved every 1000 generations. Convergence of parameter estimates was checked by plotting them against the generation number. Summary statistics and trees were generated using the last three million generations, well beyond the point at which convergence of parameter estimates had taken place.

The evolution of morphological characters and geographic origin was traced along the tree using maximum parsimony in the Mesquite software package (Maddison and Maddison, 2006). In determining geographic ranges of the ESUs, only specimens from this study were used.

2.6. Tree rooting

The root of the *Codium* phylogenetic tree was inferred using two alternative methods. First, the root position was inferred using the molecular clock. The rationale behind this approach is that, if evolution is clock-like, the root of the tree is to be found along its oldest branch, at exactly the same distance from each terminal taxon. Molecular clock rooting was achieved by analyzing the concatenated alignment in MrBayes 3.1.2 using a GTR + Γ + I model constrained by the assumption of a strict (uniform) molecular clock (analysis specifications in Appendix 4). Second, the more commonly used outgroup comparison method was applied to infer the root position. Sequences of a *Bryopsis* species (a sister genus of *Codium*; Lam and Zechman, 2006) and an *Ostreobium* species (a more distantly related bryopsidalean genus) were added to the concatenated alignment. The alignment was analyzed with each of the outgroup sequences separately and together using GTR + Γ + I models as specified in Appendix 5.

For reasons explained below, our principal phylogenetic analyses (Section 2.5) were carried out with ingroup sequences only and manually rooted along the branch inferred to be the oldest using the molecular clock rooting method.

3. Results

3.1. Species delimitation and taxonomic considerations

The ESU delimitation alignment contained 227 sequences and was 741 bases in length, although many sequences were shorter due to missing parts at either terminus (average sequence length 701 bases). In the NJ bootstrap phylogeny inferred from this alignment, 74 ESUs preceded by a relatively long branch, having high bootstrap support and low intra-cluster sequence divergence, could be demarcated (Appendix 3).

In many cases, morphological identifications did not correspond to ESUs. In some cases, a single morphological species (e.g. *C. geppiorum*) was represented in several ESUs. In most of these cases, subtle morphological differences existed between these ESUs. In other cases, several morphological species clustered within a single ESU. The closely related species *C. acuminatum* and *C. arabicum* (Silva, 1959) could be conspecific and *C. inerme* sequences are recovered among *C. fragile* sequences (see also Shimada et al., 2004). Both examples indicate that the presence of a mucron, the diagnostic character used within these species pairs, may not always be trustworthy. Sequences attributed to several Arabian Sea species (cf. Nizamuddin, 2001) often fell within a single ESU. The ESU named *Codium duthieae* 3 contained specimens conforming to *C. fastigiatum*, *C. duthieae* and *C. decorticatum*. *Codium* cf. *latum* 2 included specimens attributed to no less than ten morphological species: *C. bartlettii*, *C. bilobum*, *C. boergesenii*, *C. fimbriatum*, *C. flabellatum*, *C. gerloffii*, *C. indicum* sensu Nizamuddin, *C. latum*, *C. pseudolatum* and *C. shameelii*. Furthermore, we found considerable morphological overlap between these ten morphological species in our collections of *C.* cf. *latum* 2. Lastly, we included some specimens that probably represent species new to science.

3.2. Exploration of the phylogenetic data

The length distribution of random trees, calculated against the concatenated alignment, was considerably left-skewed ($g_1 = -0.99$), indicating that the concatenated alignment is significantly more structured than random data. The same is true for the first plus second and third codon positions separately ($g_1 = -0.94$ and $g_1 = -0.71$, respectively). The I_{ss} statistics were significantly smaller than the critical values for the alignment as a whole and the first plus second and third codon positions separately ($p < 0.001$ in all cases), indicating that substitution saturation is not an issue in our dataset.

The base frequencies and substitution rates of the different genes and codon positions, calculated against a phylogeny obtained from Bayesian analysis using a GTR + Γ + I substitution model showed that neither base frequencies nor substitution rates differ much between genes (Fig. 2). However, there are large differences between codon positions. Third codon positions have very high AT content (84–89%) whereas first and second codon positions have more balanced base frequencies (52–60% AT content; Fig. 2B). Rates at third codon positions are 5.5–18 times as high as at first and second codon positions (Fig. 2A).

As a general rule, more complex (parameter-rich) nucleotide substitution models fit the data better. Results obtained with the first and second order AIC were nearly identical and we have presented only the first order AIC (Fig. 3). Partitioning into genes does not contribute much to the fit. Uncoupling rates and model parameters among codon positions, on the other hand, seems crucial to obtaining a good fit. The difference between an AAB or ABC configuration of codon position uncoupling did not have a large impact on the fit, implying that the principal contrast is between the third and first two codon positions. Allowing the rates to vary across sites (+Γ) increased model fit considerably. The fit of a covariotide model was evaluated using Bayes factors. The Bayes factors were calculated as the ratio of the model likelihoods obtained from the three main Bayesian analyses (GTR + Γ + I, codon position and covariotide analyses—see below). The fit of the covariotide model was much better than that of the codon position model ($BF = e^{59.76}$) and the GTR + Γ + I model ($BF = e^{428.76}$). The calculation of the Bayes factor in another context is detailed in Appendix 4.

3.3. Molecular phylogenetic analyses

The observations of substitution rate and base frequency variation across codon positions and the fit of the

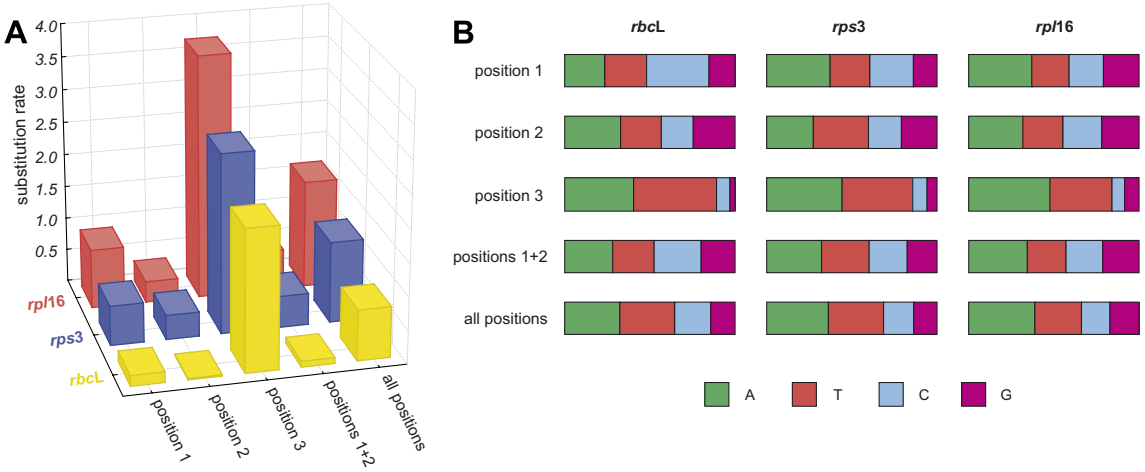

Fig. 2. Substitution rates (A) and base composition (B) of different genes and codon positions. Rates and composition mainly differ between codon positions, much less between genes. First and second codon positions have similar characteristics, which are well represented when they are joined (positions 1 + 2). Joining all codon positions, however, yields average characteristics that deviate from those of all individual codon positions.

partitioning →	unpartitioned	genes - 3		codon positions - AAB - 2		codon positions - ABC - 3		genes & codon positions - AAB - 6		genes & codon positions - ABC - 9	
uncoupled parameters →		rates	rates + model par	rates	rates + model par	rates	rates + model par	rates	rates + model par	rates	rates + model par
JC	-12756 0	-12651 2		-11738 1		-11697 2		-11568 5		-11519 8	
JC + Γ	-11024 1	-11018 3	-11003 5	-10873 2	-10778 3	-10840 3	-10764 5	-10801 6	-10731 7	-10739 9	-10693 11
HKY	-12767 4	-12660 6	-12633 14	-11641 5	-11685 9	-11602 6	-11650 14	-11490 9	-11473 29	-11420 12	-11410 44
HKY + Γ	-10936 5	-10931 7	-10894 17	-10773 6	-10767 11	-10741 7	-10762 17	-10675 10	-10670 35	-10636 13	-10621 53
GTR	-12708 8	-12603 10	-12538 26	-11500 9	-11147 17	-11458 10	-11089 26	-11353 13	-10919 53	-11282 16	-10819 80
GTR + Γ	-10824 9	-10823 11	-10761 29	-10656 10	-10346 19	-10626 11	-10326 29	-10567 14	-10195 59	-10520 17	-10133 89

AIC score: 20500 / 23000 / 25500

Fig. 3. Fit of different substitution models to the phylogenetic data. For each model tested, the log-likelihood (big print), number of parameters (small print) and Akaike Information Criterion (AIC) score (color code) are given. Model fit increases with decreasing AIC scores (increasingly red color). More complex models (with more parameters) fit the data best. Partitioning the data into genes does not differ much from the unpartitioned situation. Partitioning into codon positions causes a considerable increase in model fit. Partitioning codon positions into an AAB or ABC configuration hardly affects the model fit. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

different base substitution models to the data led us to choose three different combinations of data partitions and substitution models, which were used to infer the phylogeny of *Codium* species. In addition to using the common GTR + Γ + I model, the data were partitioned into first plus second and third codon positions, and analyzed with separate rates and GTR + Γ + I parameters for each partition (codon position analysis). The partitioned alignment was also subjected to the codon position model with the option to allow rate variation across the tree (covariotide model). The three analyses converged onto virtually identical topologies, differing only in certain node support values and a few alternative ramifications in regions with very low support. The phylogram obtained from the covariotide analysis is shown in Fig. 4, those from the other analyses in Appendix 6.

The same three major lineages (A, B and C) were recovered in all analyses. Lineage A consisted of two early branching lineages (grade A1) and a strongly supported clade A2. Lineage B was divided into two clades (B1 and B2) that received strong support. Lineage C comprised a grade of early branching species (C1) among which relationships were not well resolved in all analyses and a strongly supported clade (C2) containing almost half of the species in our study.

3.4. Tree rooting

Two alternative methods were used to root the *Codium* phylogenetic tree. First, the root position was inferred by constraining a phylogenetic analysis with clock-like evolutionary rates. This analysis, presented in detail in Appendix 4, resulted in the root position shown in Fig. 4, between lineage A and lineages B + C. The outgroup analyses resulted in another root position (Appendix 5). The analyses with only *Bryopsis* and *Ostreobium* plus *Bryopsis* placed the root on the branch leading to *C. megalophysum* in clade B2. In the analysis with only *Ostreobium*, the root was placed within clade B1, along the branch leading to the remainder of species after *C. papenfussii* branched off. Branches leading to the outgroups were very long. This is also illustrated by the intra- and intergeneric sequence divergences: whereas the largest pairwise uncorrected distance between *Codium* species was 14%, intergeneric comparisons between *Codium* and the outgroups were at least 16% for *Bryopsis* and 21% for *Ostreobium*. For reasons discussed below, we doubt the results obtained with outgroup rooting and have used the root position obtained with the molecular clock method in further analyses (mapping of morphology and geography).

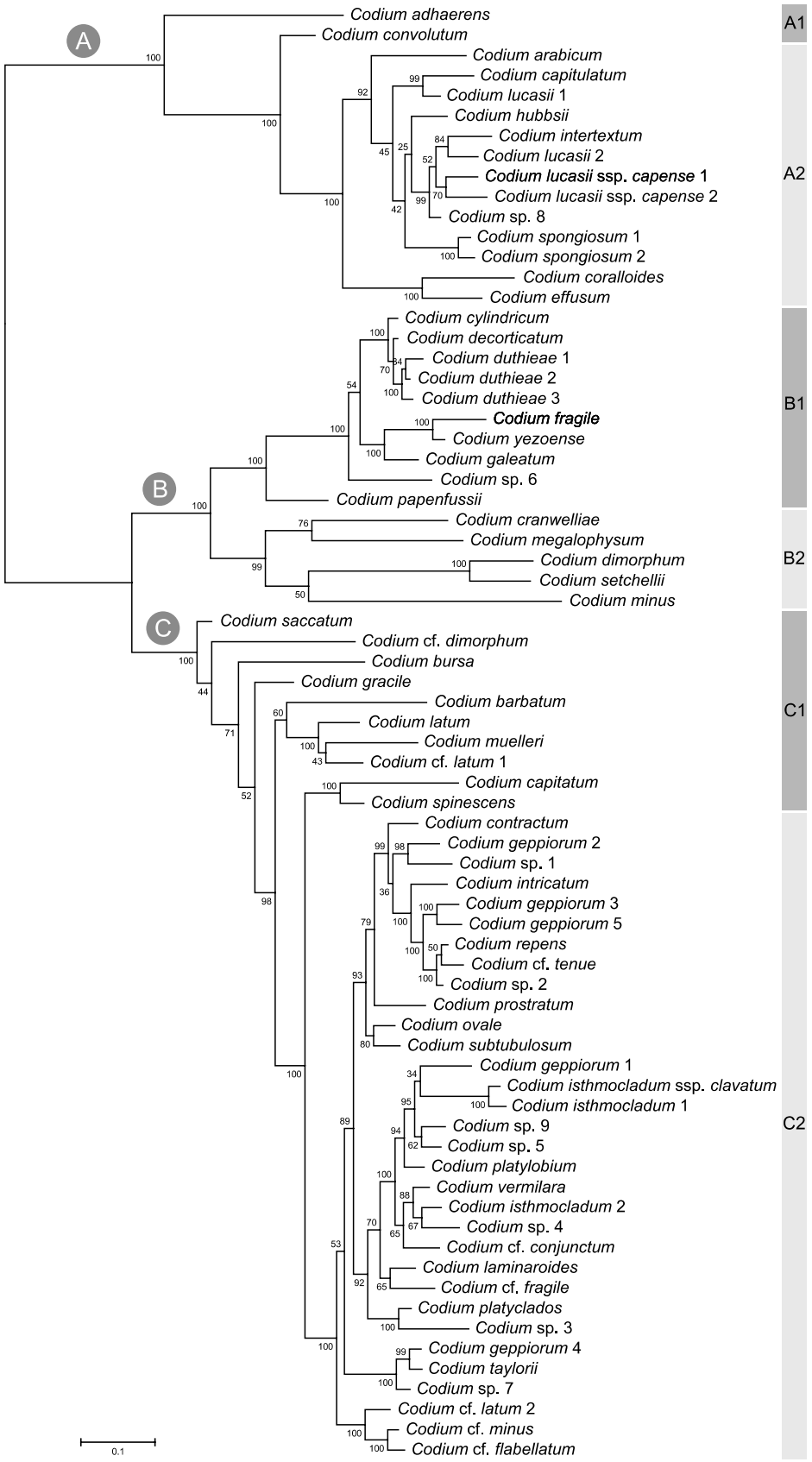

Fig. 4. Phylogenetic hypothesis of *Codium* species inferred from concatenated plastid genes. The tree is the majority rule consensus tree resulting from a Bayesian analysis of five million generations, using a covariotide model in which the alignment was partitioned into first plus second and third codon positions and rates and GTR + Γ + I model parameters were uncoupled among partitions. Values at the nodes represent posterior node probabilities; the scale is in number of substitutions per site. The tree was manually rooted along its oldest branch.

3.5. Mapping morphology and geography

The parsimony reconstruction of the evolution of a number of external morphological characters along the phylogram (Fig. 5) shows that general thallus architecture is clearly correlated with the diversification of the genus (Fig. 5A). Whereas clade A consists entirely of mat-forming species, the early diverging lineages of clade B feature spherical thalli. Clade B1 also features a distinct, monophyletic lineage with erect species. *Codium dimorphum* and *C. setchellii* deviate from the remainder of the clade in being mat-forming. Clade C has a few early branching spherical and mat-forming species, but the bulk of its species are branched, either erect or sprawling. The erect thallus habit seems to be the ancestral situation from which the sprawling habit has evolved several times independently. Spherical thalli have evolved from branched ones on two occasions. Looking at branched species in more detail, one can see that the distribution of branch broadening is less clear-cut (Fig. 5B). In clade B1, a lineage with branches that are markedly broadened below ramifications (the *C. decorticatum* morphology) may have originated from a grade of species with cylindrical branches. From here onwards, we will refer to this clade with broadened branches below the ramifications, comprising *C. cylindricum*, *C. decorticatum* and three ESUs identified as *C. duthieae*, as the *decorticatum* clade. It is important to note that this morphology is not restricted to clade B; it has evolved independently in *C. subtubulosum* of clade C2. Clade C consists of a series of derivations of a thallus with cylindrical branches. Entirely flattened thalli have evolved several times independently and changes between entirely cylindrical branches and branches that are slightly broader than thick below nodes or throughout seem to have been plentiful. Branch diameter changes frequently along the topology, especially within clade C (Fig. 5C). It must be noted that many nodes in clade C receive mediocre support and the actual number of changes may be slightly less than suggested in the figures. Clade B1 is characterized by thick branches, reducing significantly only in the *C. fragile-yezoense* lineage.

Some anatomical characters are traced along the phylogram in Fig. 6. With the notable exception of *C. spongiosum* and *C. coralloides*, species of clade A predominantly have narrow utricles (Fig. 6A). Clade B is characterized by large, sometimes enormous (*C. megalophysum* and *C. papenfussii*) utricles, and in clade C utricles of intermediate size dominate. *Codium dimorphum* and *C. setchellii* have markedly narrower utricles than the remainder of clade B2. Whereas composite utricles dominate clade A, species of clades B and C predominantly have simple utricles (Fig. 6B). Mucrons (pointed appendages on top of the utricles) and umbos (inwardly pointing appendages) have arisen several times independently (Fig. 6C) and are not always a consistent feature within species (e.g. *C. inerme* and *C. acuminatum*; see Section 3.1).

When interpreted against the geographic origin of each ESU, the topology does not reveal many overall patterns

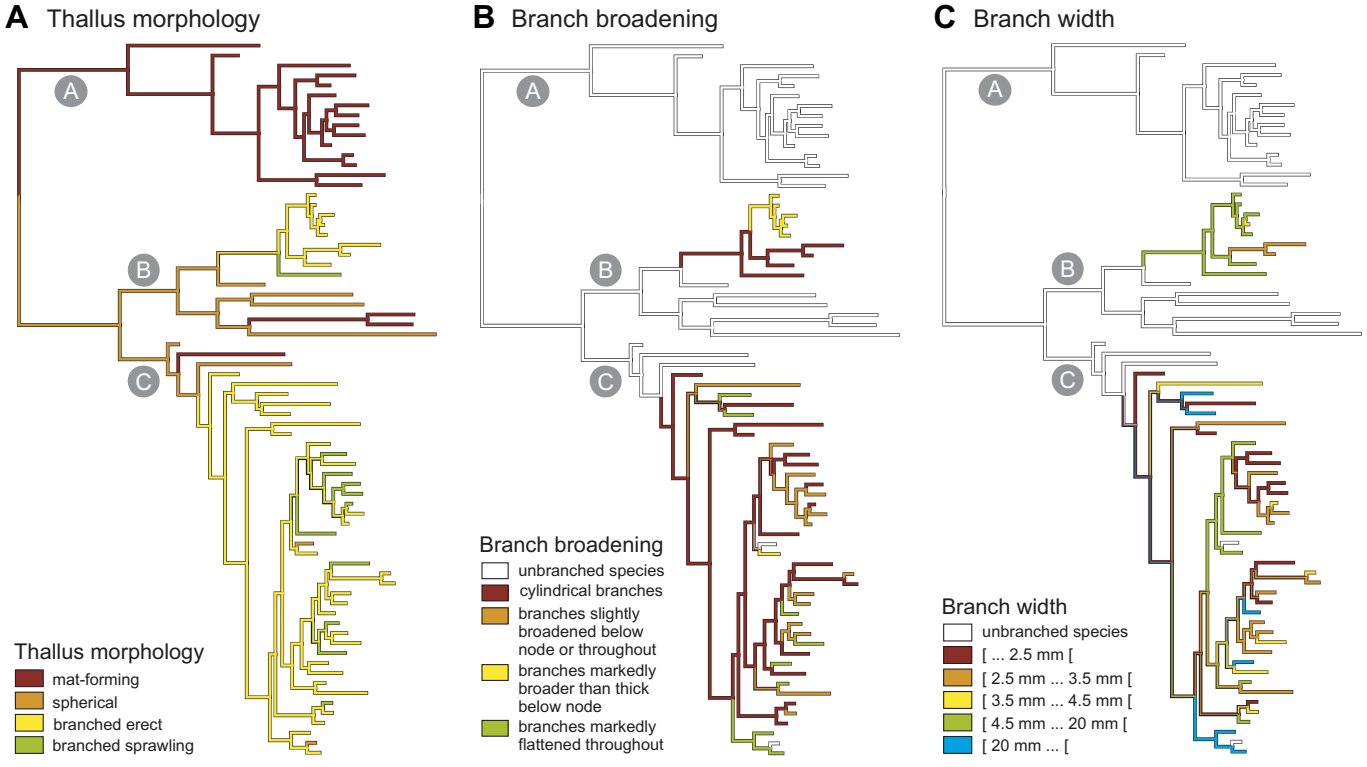

Fig. 5. Evolution of external morphological characters mapped onto the phylogenetic tree. Ancestral traits were reconstructed using maximum parsimony.

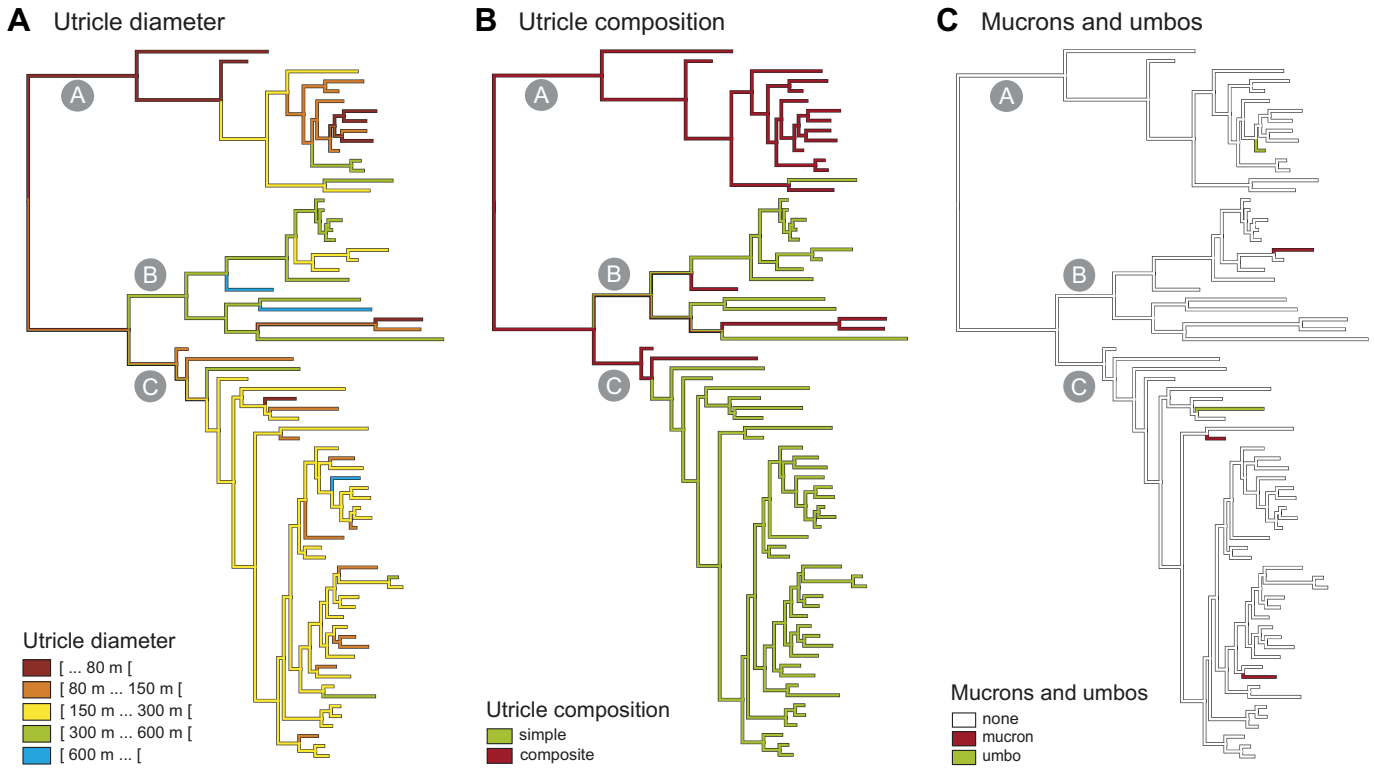

Fig. 6. Evolution of anatomical characters mapped onto the phylogenetic tree. Ancestral traits were reconstructed using maximum parsimony.

(Fig. 7). Nonetheless, in general, Atlantic species or clades have emerged from predominantly Indo-Pacific clades on several occasions. We verified a number of previously reported biogeographic hypotheses against our topology; this will be detailed in Section 4.4.

4. Discussion

4.1. Taxonomic challenge

Identifying *Codium* specimens using only morphological characters can be extremely challenging. Even though a few distinctive species clearly stand out from the rest, most collections are very hard to identify. Whereas the species of *Codium* along the coasts of North America, Europe, South Africa and southern Australia are well characterized, specimens collected elsewhere are often difficult or impossible to identify. It is usually easy to place specimens in the morphological framework on which the sectional subdivision is based. Within sections, however, there are many species to choose from, and some specimens match aspects of multiple descriptions, or possess characters of two or more species yet do not conform to any of these species in all aspects.

In our opinion, accurate identification can currently be achieved best by comparing specimens' DNA sequences. The *rbc*L exon 1 can be sequenced easily and compared to our sequence dataset. Judging from our 227 sequences, *rbc*L exon 1 facilitates accurate identification because in most parts of the tree, sequences cluster in groups with low intra-cluster and high inter-cluster divergences. Among-cluster divergences are lower in clade C2 and increased sampling may obscure ESU boundaries in this region of the tree.

The morphological diversity within ESUs varies strongly. One extreme case is *C.* cf. *latum* 2, an ESU containing a wide spectrum of flattened *Codium* morphologies from the Arabian Sea, most of which were previously considered to be different species (Nizamuddin, 2001). At the other extreme, specimens identified as *C. geppiorum* were resolved into five distinct ESUs. Silva (1962) has already noted that the anatomical variability of *C. geppiorum* from reef to reef is perplexing. A particularly noteworthy observation is that the general morphology of the invasive species *C. fragile* is not unique to this species, making the use of DNA data to identify the invasive strain indispensable (see Stam et al., 2006 for an example in the genus *Caulerpa*). Our morphological survey revealed subtle differences among the ESUs in most cryptic species pairs or complexes, suggesting that in-depth morphological and molecular surveys could result in morphological characterization of the ESUs. Pseudo-cryptic diversity is common in algae—many studies have recognized multiple entities within morphological species that could be identified using post hoc morphological examination (e.g. De Clerck et al., 2005; Saunders and Lehmkuhl, 2005). Although not a guarantee of success, juxtaposition of congruent morphometric and molecular datasets seems to be particularly useful for pinpointing morphological boundaries between pseudo-cryptic species (De Senerpont-Domis et al., 2003;

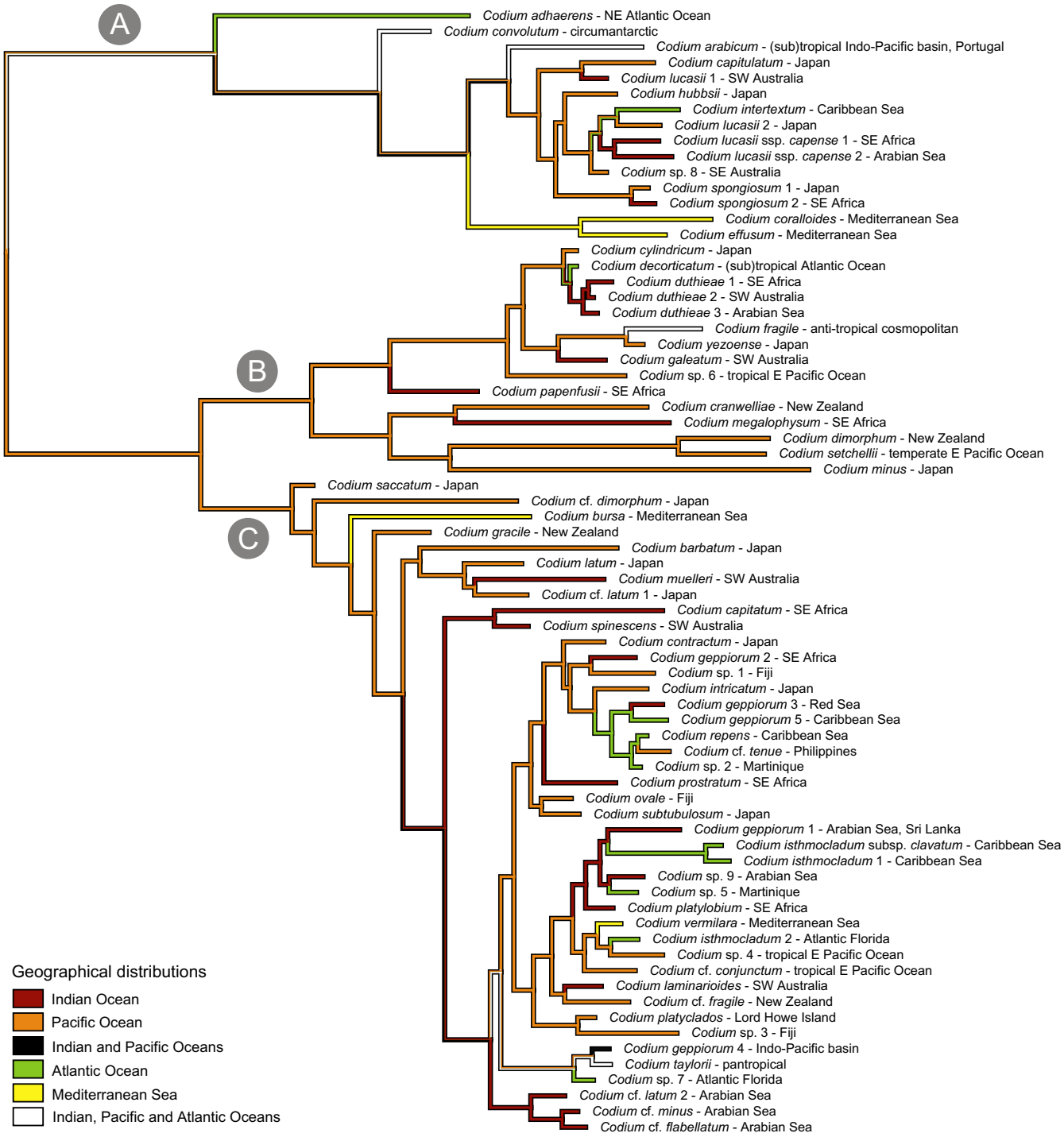

Fig. 7. Geographical distributions mapped onto the phylogenetic tree.

Verbruggen et al., 2005a). A morphometric modus operandi has been developed for *Codium* but has not been applied to taxonomic questions on a broad scale (Hubbard and Garbary, 2002).

Considering our data, species diversity in *Codium* needs a thorough re-examination. We believe that the only successful approach to the development of a sounder taxonomy would be to carry out broad-scale, regional surveys of *Codium* species using molecular tools to identify specimens and recognize additional ESUs, supplemented with morphological observations allowing the description of the regional morphological variability of the ESUs in question. This approach would also allow the type specimens of currently recognized species to be fitted into the proposed taxonomic system, ideally by sequencing a short stretch of their *rbc*L gene or by critically comparing them to ESUs using those morphological features that are diagnostic characters for the ESUs.

In any attempt to upgrade a taxonomic framework, it is important to reflect on species boundaries. Cross-fertility is difficult to assess in *Codium*—to our knowledge crossing experiments have never been carried out. On the other hand, it turns out to be fairly straightforward to identify ESUs using molecular data. The ESUs we identified can be considered species under the phylogenetic species concept and may or may not conform to other species definitions. Considering the fact that most, if not all, of these ESUs show at least some morphological differences, there is a fair chance that they are distinct species. Nonetheless, the species status of ESUs could be disputed in some parts of the tree. For example, in the *decorticatum* clade (see also Goff et al., 1992), different ESUs were assigned to distinct clades of specimens with different geographical origins, even though the branches towards them were fairly short. Another option would have been to group the whole clade in a single ESU. Similarly, in clade C2, branches towards some ESUs were rather short. One should therefore interpret our ESUs as representatives of various stages in the speciation process, from recently diverged populations to clear-cut biological species.

4.2. Tree rooting

Rooting our trees was arduous. Using two genera (one closely and another more distantly related) or either of these separately as outgroups, the root was always recovered within clade B, most often within clade B2, which is composed of a number of taxa sitting on long branches. Placing the root within clade B resulted in trees with highly unequal root-to-leaf distances, leading us to believe that the root position obtained with the outgroup method is a product of phylogenetic bias. Therefore, the trees we present result from analysis of ingroup sequences and are manually rooted at the root position inferred using an analysis under a uniform molecular clock model (GTR + Γ + I; Appendix 4). The early branching position of the *C. minus* clade in the outgroup-rooted phylogenetic tree by Shimada et al. (2004) is most likely an artifact of phylogenetic bias.

Outgroup rooting introduces a significantly more distantly related sequence in phylogenetic analyses, making them prone to long branch attraction and other forms of phylogenetic bias. It has been documented that outgroups can be mistakenly inferred on a long ingroup branch as a consequence of long branch attraction and that inclusion of outgroup sequences can yield erroneous ingroup topologies (e.g. Holland et al., 2003). In our outgroup rooting experiments, the root was placed in clade B2, which is characterized by long-branch taxa. It has been shown that, when random sequences are used as outgroups, they preferably root the tree at a long branch, often a terminal one (Graham et al., 2002; Wheeler, 1990). This was likely the case in our analyses, too. Phylogenetic methods can be positively misled by incorrect assumptions about the model of evolution (Chang, 1996) and by parameters varying across lineages, such as evolutionary rates (Fares et al., 2006; Omilian and Taylor, 2001), base composition (Conant and Lewis, 2001; Rosenberg and Kumar, 2003) and the number of sites that are free to vary (Lockhart et al., 1998). Considering the disparate root position obtained with molecular clock and outgroup methods, the placement of the root on the *Codium* phylogenetic tree should be examined in more detail. Aside from examining the confounding factors listed above, such an examination should explore different rooting methods, test a variety of outgroup taxa, use markers that evolve at different rates and attempt to improve sampling to break up the long branches in the B2 clade.

4.3. Morphological evolution

It is a long-standing belief that all *Codium* morphologies evolved from mat-forming ancestors (Schmidt, 1923). Globose thalli were thought to have originated from mat-forming ones by bulging upward and erect thalli by longitudinal outgrowth. If we may assume that the root position inferred using the molecular clock is correct, our data largely confirm these hypotheses. The maximum parsimony-reconstructed character evolution shown in Fig. 5 clearly indicates that the mat-forming and spherical thallus morphologies are the most primitive ones. The character state at the root is ambiguous; it could be either mat-forming or spherical.

The evolution of *Codium* is characterized by relatively few important morphological shifts. Branched forms, which make up the bulk of the species, have evolved twice independently. In addition, there have been two independent 'reversals' from branched to spherical morphologies (*C. ovale* and *C.* cf. *minus*). Within the clades containing branched species, variation on the basic pattern has evolved considerably more commonly. Sprawling species are scattered across the predominantly erect clade C. Marked broadening of branches below ramifications (the *C. decorticatum* morphology) has evolved twice independently. More subtly broadened and cylindrical branches alternate throughout clade C. Entirely flattened branches have evolved multiple times independently, at least six or seven times in the taxon sample here analyzed.

The small number of fundamental shifts in thallus morphology (between mat-forming, spherical and branched) indicates that these basic morphologies have relatively strong historical and genetic determinants. After all, one could imagine a situation in which free niches in a region were occupied by new *Codium* forms through adaptive morpho-ecological shifts, causing convergent evolution. Although the general pattern may not support this hypothesis, it could explain the origin of *C. ovale* and *C.* cf. *minus*, two spherical species in a clade of otherwise erect, branched species. The latter species, occurring in the Arabian Sea, is embedded in a clade of erect species, all from the same region, strongly suggesting that the spherical habit in *C.* cf. *minus* originated by local adaptation.

In contrast to the limited number of fundamental morphological shifts, there have been many evolutionary experiments within the branched species, more particularly within clade C, where the sprawling habit and the entirely flattened morphology have originated multiple times independently. Consequently, section *Elongata*, a clearly delineated group of flattened species, turns out to be an artificial assemblage of species resulting from convergent evolution. Since both subgenera and many of the sections contain species from different places in the phylogenetic tree, a critical evaluation of the generic subdivision is required.

Silva (1954) stressed the phylogenetic importance of anatomical characters. In his view, the composite utricles typically found in mat-forming species represent a primitive state from which simple utricles were derived. Our data confirm that composite utricles are primitive and have given rise to simple utricles in all major lineages. Likewise, primitive utricles are likely to have been small, and bigger utricles evolved in all lineages independently.

Relying on the number and nature of siphons extending from the base of utricles, Silva (1954) suggested that spherical thalli with large utricles were independently derived from mat-forming ancestors with smaller utricles three times; once in *C. bursa* and allies, once in *C. mamillosum* and allies, and once in an undescribed species. Our phylogeny places *C. bursa* in grade C1 and *C. minus*, a species extremely similar to *C. mamillosum* (once considered to be conspecific; Schmidt, 1923), is recovered in clade B2. Although better taxon sampling and more detailed morphological observations are needed to test Silva's hypothesis, we expect it to be supported. In addition to the cases listed by Silva, the spherical thallus habit has evolved at least two more times (*C. ovale* and *C.* cf. *minus*), not from a mat-former but from a branched ancestor. Here too, detailed anatomical analyses should be carried out to find the characters linking it to its natural allies. It is clear that in order to delineate natural groupings within *Codium* and other siphonous algae one must not rely solely on external morphological characters (Verbruggen and Kooistra, 2004).

4.4. Biogeographic considerations

One of the most striking observations in our data was that specimens belonging to a single morphological species often separated into multiple, geographically separated ESUs. This was the case for *C. lucasii* (lineage A), dividing up into four widely geographically separated ESUs, and specimens with a *C. decorticatum*-like morphology (the *decorticatum* clade in lineage B), which were resolved into five geographically separated ESUs. Several other examples are present, but are less conclusive because of limited taxon sampling. Finding multiple ESUs within morphological species is common in algae, and the resulting ESUs are often geographically restricted (e.g. Kooistra et al., 2002; De Clerck et al., 2005; Gurgel et al., 2004). Regional endemism is being disclosed using molecular data for a variety of benthic and sedentary marine organisms (e.g. Carlin et al., 2003; Meyer et al., 2005; Muss et al., 2001), suggesting the importance of regional adaptation and dispersal limitation despite the high dispersal potential brought about by ocean currents (Scheltema, 1968). Surveys of population genetic data showed that macroalgae are among the poorest dispersers of all marine organisms (Kinlan and Gaines, 2003; Kinlan et al., 2005). In *Codium*, regional endemism seems to be particularly high, with only one ESU in our present sample occurring both in the Atlantic and Indo-Pacific (*C. taylorii*). The dispersal stages of *Codium* include motile flagellated cells, which account for local dispersal, and mature thallus fragments, which are responsible for long-distance dispersal (Carlton and Scanlon, 1985; and references therein). Thallus fragments float because of oxygen bubble formation and can withstand fairly long periods of desiccation, increasing chances of successful dispersal by rafting (Schaffelke and Deane, 2005). The question of dispersal is particularly important with respect to *C. fragile* ssp. *tomentosoides*, listed as the most vigorous of all invasive algae (Nyberg and Wallentinus, 2005). This entity of Japanese origin has repeatedly invaded European and American shorelines and its spread has been well documented (Carlton and Scanlon, 1985; Provan et al., 2005).

Considering the phylogeny as a whole, no large vicariance events stood out: each of the three major clades encompasses species from the world's three major oceans. This indicates that any such events acting on *Codium* speciation must have happened after the initial diversification into the three major clades and/or that the imprint of early vicariance is masked by more recent dispersal. A general observation is that Indo-Pacific diversity is greater than Atlantic diversity and that Atlantic species are usually embedded in clades dominated by Indo-Pacific species. This could lead one to believe that the genus originated and diversified in the Tethys Sea and subsequently dispersed into the Atlantic Ocean several times independently. Too few algal genera have been examined in enough detail to come to general conclusions about their historical biogeography. The historical biogeography of *Codium* can however be compared with that of the calcified genus *Halimeda*, a relative with an extensive fossil record and a history of molecular biogeographic studies (Hillis, 2001; Kooistra et al., 1999, 2002; Verbruggen et al., 2005b). *Halimeda* originated and diversified into its major lineages in the Tethys Sea. Each major lineage subsequently underwent a vicariance event causing a split between Atlantic and Indo-Pacific species. These vicariance events were reinforced because *Halimeda* is strictly tropical and subtropical, making the north-south oriented African and American continents impassable barriers between the Atlantic and Indo-Pacific basins. There are no indications for an impact of an Atlantic versus Indo-Pacific vicariance event on the diversification of *Codium*. One could hypothesize that the fact that *Codium* ranges into colder waters makes migration around the southern

tip of Africa and via the Antarctic circumpolar current easier, resulting in species with a global distribution and multiple sister clades across land barriers. In this context, it should be noted that dispersal by means of the Antarctic circumpolar current should impact only on subantarctic to cold temperate species whereas the examples in our phylogeny are mostly tropical or subtropical species. The only circumantarctic species in our analysis is *C. convolutum* (lineage A) of which we have sequenced samples from New Zealand and Tristan da Cunha.

Our *Codium* phylogeny can be used as a framework to test the validity of some previously proposed biogeographic links between subtropical floras (Hommersand, 1986). In the literature, the disjunct distribution of the entirely flattened erect species (*C. latum* and *C.* cf. *latum* 1 in Japan, *C. laminarioides* in SW Australia, *C. platylobium* in SE Africa and *C.* cf. *latum* 2 in the Arabian Sea) was invoked as evidence for a biogeographic link between these regions (Silva, 1962; *C.* cf. *latum* 1 and 2 added by us). Our results leave no doubt that the flattened morphology evolved several times independently and that the biogeographic link is artificial in this case. *Codium lucasii* also features a disjunct distribution in these regions (*C. lucasii* 1 in SW Australia, *C. lucasii* 2 in Japan, *C. lucasii* ssp. *capense* 1 from SE Africa and *C. lucasii* ssp. *capense* 2 from the Arabian Sea). Here, the link between Australia and South Africa originally suggested by Silva (1962) and explored further by Hommersand (1986) is proven to be a result of convergent evolution. Nonetheless, the Japanese, SE African and Arabian Sea populations, together with the Atlantic species *C. intertextum*, do share a relatively recent common origin. The occurrence of *Codium minus* in Japan and the Arabian Sea was also used to invoke biogeographic affinities between these regions (Wynne, 2004). Here again, morphological convergence is the cause of the apparent link.

Despite these examples of convergence, there are a few clades that seem to support hypotheses of biogeographic affinities between the Arabian Sea, SE Africa, SW Australia and Japan. First, *C. spongiosum* occurs in SE Africa and Japan. Although indicating a sibling species pair rather than a single species, our sequences support the biogeographic link. It must be noted that *C. spongiosum* is also reported from SW Australia, Mauritius, Hawaii, Brazil and the Caribbean Sea and the link may not hold as other samples are added (Silva, 1959). Second, the SW Australian species *C. muelleri* originated within a strongly supported clade of Japanese species (*C. latum* and *C.* cf. *latum* 1). Third, SE African *C. capitatum* and SW Australian *C. spinescens* cluster in a well-supported clade. Fourth, the *decorticatum* clade also comprises ESUs from these different regions. Japanese *C. cylindricum* branches off first, followed by Atlantic *C. decorticatum*. The remainder of the clade, consisting of *C. duthieae* 1 (SE Africa), *C. duthieae* 2 (SW Australia) and *C. duthieae* 3 (Arabian Sea), receives very high support, reflecting a close relationship between these ESUs. It must be noted, however, that the *C. decorticatum* morphology exists in other areas of the world.

In conclusion, molecular phylogenetic investigations of *Codium* provide support for certain biogeographic links between distant subtropical regions of the Indo-Pacific. Several of the original examples used to formulate the hypotheses (based on morphological consistency) are contradicted by our data and are most likely examples of convergent evolution. Nonetheless, a number of examples—some of which are new—support biogeographic links between Japan and SW Australia and between SE Africa, SW Australia and the Arabian Sea. The affinity between the latter three regions was recently confirmed using molecular data for the genus *Halimeda* (Verbruggen et al., 2005b). Surprisingly, despite extensive indications from floristic data (Børgesen, 1934; Wynne, 2000, 2004) and the occurrence of some extremely similar *Codium* morphologies, our data negate all possible links between the *Codium* floras of the Arabian Sea and Japan. We are of the opinion that the affinities between the Japanese and Arabian Sea marine floras should be investigated using molecular data from a wider array of genera.

Acknowledgments

This research was funded by FWO-Flanders (Grants G.0136.01 and G.0142.05), the Energy, Environment and Sustainable Development program of the European Union (ALIENS project: EVK3-CT-2001-00062), the Esmée Fairbairn Foundation (Marine Aliens project), the Flemish Government (bilateral research grant 01/46), the Smithsonian Marine Station (SMS Contr. No. 684), Harbor Branch Oceanographic Institution and the King Leopold III Fund for Nature Exploration and Conservation. H.V., F.L., O.D.C. and T.S. are indebted to BOF (Ghent University) and FWO-Flanders for post-doctoral fellowship grants. Caroline Vlaeminck, Barrett Brooks, Nadjejda Espinel-Velasco, Ellen Cocquyt, Cathy De Maire, and Christelle Vankerckhove are gratefully acknowledged for carrying out parts of the laboratory and administrative work. We sincerely thank Rob Anderson, Lin Baldock, An Bollen, Christian Boedeker, John Bolton, Barrett Brooks, Francis Bunker, Else Demeulenaere, Roxie Diaz, Stefan Draisma, Jelle Evenepoel, Wilson Freshwater, Daniela Gabriel, Cristine Galanza, Nisse Goldberg, Dennis Hanisak, John Huisman, Courtney and Tom Leigh, Lynne McIvor, Deborah Olandesca, Klaas Pauly, Pieter Provoost, Willem Prud'homme van Reine, Sherry Reed, Jose Rico, Gary Saunders, Kerry Sink, Herre Stegenga, Enrico Tronchin, Cynthia Trowbridge and Joe Zuccarello for collecting specimens or assisting in the field.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2007.01.009.

References

Bolton, J.J., Leliaert, F., De Clerck, O., Anderson, R.J., Stegenga, H., Engledow, H.E., Coppejans, E., 2004. Where is the western limit of the tropical Indian Ocean seaweed flora? An analysis of intertidal seaweed biogeography on the east coast of South Africa. Mar. Biol. 144, 51–59.

Børgesen, F., 1934. Some marine algae from the northern part of the Arabian sea with remarks on their geographical distribution. Kongelige Danske Videnskabernes Selskab, Biologiske Meddelelser 11, 1–72.

Burrows, E.M., 1991. Seaweeds of the British Isles, vol. 2. Chlorophyta, Natural History Museum, London, UK, 238 pp.

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

Carlton, J.T., Scanlon, J.A., 1985. Progression and dispersal of an introduced alga: *Codium fragile* ssp. *tomentosoides* (Chlorophyta) on the Atlantic Coast of North America. Bot. Mar. 28, 155–165.

Chang, J.T., 1996. Inconsistency of evolutionary tree topology reconstruction methods when substitution rates vary across characters. Math. Biosci. 134, 189–215.

Chihara, M., 1975. The Seaweeds of Japan. Gakken Illustrated Nature Encyclopedia. Gakken Co., Tokyo Japan, 292 pp.

Conant, G.C., Lewis, P.O., 2001. Effects of nucleotide composition bias on the success of the parsimony criterion in phylogenetic inference. Mol. Biol. Evol. 18, 1024–1033.

De Clerck, O., Gavio, B., Fredericq, S., Barbara, I., Coppejans, E., 2005. Systematics of *Grateloupia filicina* (Halymeniaceae, Rhodophyta), based on *rbc*L sequence analyses and morphological evidence, including the reinstatement of *G. minima* and the description of *G. capensis* sp. nov. J. Phycol. 41, 391–410.

De Clerck, O., Leliaert, F., Verbruggen, H., Lane, C.E., De Paula, J.C., Payo, D.A., Coppejans, E., 2006. Large subunit RUBISCO and 26S ribosomal DNA sequence analyses call for a revised classification of the Dictyoteae (Dictyotales, Phaeophyceae). J. Phycol. 42, 1271–1288.

Dellow, V., 1952. The genus *Codium* in New Zealand. Part I. Systematics. Trans. R. Soc. N. Z. 80, 119–141.

De Senerpont-Domis, L.N., Fama, P., Bartlett, A.J., Prud'homme van Reine, W.F., Espinosa, C.A., Trono, G.C., 2003. Defining taxon boundaries in members of the morphologically and genetically plastic genus *Caulerpa* (Caulerpales, Chlorophyta). J. Phycol. 39, 1019–1037.

Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.

Famà, P., Wysor, B., Kooistra, W.H.C.F., Zuccarello, G.C., 2002. Molecular phylogeny of the genus *Caulerpa* (Caulerpales, Chlorophyta) inferred from chloroplast *tuf*A gene. J. Phycol. 38, 1040–1050.

Fares, M.A., Byrne, K.P., Wolfe, K.H., 2006. Rate asymmetry after genome duplication causes substantial long-branch attraction artifacts in the phylogeny of *Saccharomyces* species. Mol. Biol. Evol. 23, 245–253.

Goff, L.J., Liddle, L., Silva, P.C., Voytek, M., Coleman, A.W., 1992. Tracing species invasion in *Codium*, a siphonous green alga, using molecular tools. Am. J. Bot. 79, 1279–1285.

Graham, S.W., Olmstead, R.G., Barrett, S.C.H., 2002. Rooting phylogenetic trees with distant outgroups: a case study from the commelinoid monocots. Mol. Biol. Evol. 19, 1769–1781.

Gurgel, C.F.D., Fredericq, S., Norris, J.N., 2004. Phylogeography of *Gracilaria tikvahiae* (Gracilariaceae, Rhodophyta): a study of genetic discontinuity in a continuously distributed species based on molecular evidence. J. Phycol. 40, 748–758.

Hillis, L.W., 2001. The calcareous reef alga *Halimeda* (Chlorophyta, Bryopsidales): a Cretaceous genus that diversified in the Cenozoic. Palaeogeogr. Palaeoclim. Palaeoecol. 166, 89–100.

Hillis, D.M., Huelsenbeck, J.P., 1992. Signal, noise, and reliability in molecular phylogenetic analyses. J. Hered. 83, 189–195.

Holland, B.R., Penny, D., Hendy, M.D., 2003. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock – A simulation study. Syst. Biol. 52, 229–238.

Hommersand, M.H. 1971. Taxonomy and phytogeographic relationships of warm temperate marine algae occurring in Pacific North America and Japan. In: Proc. 7th Int. Seaweed Symp., Sapporo, Japan, pp. 66–71.

Hommersand, M.H., 1986. The biogeography of the South African marine red algae: a model. Bot. Mar. 24, 257–270.

Hubbard, C.B., Garbary, D.J., 2002. Morphological variation of *Codium fragile* (Chlorophyta) in Eastern Canada. Bot. Mar. 45, 476–485.

Huelsenbeck, J.P., 2002. Testing a covariotide model of DNA substitution. Mol. Biol. Evol. 19, 698–707.

Huelsenbeck, J.P., Larget, B., Alfaro, M.E., 2004. Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. Mol. Biol. Evol. 21, 1123–1133.

Joosten, A.M.T., Van den Hoek, C., 1986. World-wide relationships between red seaweed floras: a multivariate approach. Bot. Mar. 29, 195–214.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.

Kinlan, B.P., Gaines, S.D., 2003. Propagule dispersal in marine and terrestrial environments: a community perspective. Ecology 84, 2007–2020.

Kinlan, B.P., Gaines, S.D., Lester, S.E., 2005. Propagule dispersal and the scales of marine community process. Divers. Distrib. 11, 139–148.

Kooistra, W.H.C.F., 2002. Molecular phylogenies of Udoteaceae (Bryopsidales, Chlorophyta) reveal nonmonophyly for *Udotea*, *Penicillus* and *Chlorodesmis*. Phycologia 41, 453–462.

Kooistra, W.H.C.F., Caldéron, M., Hillis, L.W., 1999. Development of the extant diversity in *Halimeda* is linked to vicariant events. Hydrobiologia 398, 39–45.

Kooistra, W.H.C.F., Coppejans, E.G.G., Payri, C., 2002. Molecular systematics, historical ecology, and phylogeography of *Halimeda* (Bryopsidales). Mol. Phylogenet. Evol. 24, 121–138.

Kraft, G.T., 2000. Marine and estuarine benthic green algae (Chlorophyta) of Lord Howe Island, South-western Pacific. Aust. Syst. Bot. 13, 509–648.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Lam, D.W., Zechman, F.W., 2006. Phylogenetic analyses of the Bryopsidales (Ulvophyceae, Chlorophyta) based on Rubisco large subunit gene sequences. J. Phycol. 42, 669–678.

Lane, C.E., Mayes, C., Druehl, L., Saunders, G.W., 2006. A multi-gene molecular investigation of the kelp (Laminariales, Phaeophyceae) supports substantial taxonomic re-organization. J. Phycol. 42, 493–512.

Lapointe, B.E., Barile, P.J., Littler, M.M., Littler, D.S., Bedford, B.J., Gasque, C., 2005a. Macroalgal blooms on southeast Florida coral reefs. I. Nutrient stoichiometry of the invasive green alga *Codium isthmocladum* in the wider Caribbean indicates nutrient enrichment. Harmful Algae 4, 1092–1105.

Lapointe, B.E., Barile, P.J., Littler, M.M., Littler, D.S., 2005b. Macroalgal blooms on southeast Florida coral reefs. II. Cross-shelf discrimination of nitrogen sources indicates widespread assimilation of sewage nitrogen. Harmful Algae 4, 1106–1122.

Lockhart, P.J., Steel, M.A., Barbrook, A.C., Huson, D.H., Charleston, M.A., Howe, C., 1998. A covariotide model explains apparent phylogenetic structure of oxygenic photosynthetic lineages. Mol. Biol. Evol. 15, 1183–1188.

Lucas, A.H.S., 1935. The marine algae of Lord Howe Island. Proc. Linnean Soc. N. S. W. 60, 194–232.

Lüning, K., 1990. Seaweeds. Their Environment, Biogeography and Ecophysiology. Wiley Interscience, New York.

Maddison, W.P., Maddison, D.R., 2006. Mesquite: a modular system for evolutionary analysis. Version 1.1. <http://mesquiteproject.org>.

Meyer, C.P., Geller, J.B., Paulay, G., 2005. Fine scale endemism on coral reefs: archipelagic differentiation in turbinid gastropods. Evolution 59, 113–125.

Moritz, C., 1994. Defining evolutionarily significant units for conservation. Trends Ecol. Evol. 9, 373–375.

Muss, A., Robertson, D.R., Stepien, C.A., Wirtz, P., Bowen, B.W., 2001. Phylogeography of *Ophioblennius*: the role of ocean currents and geography in reef fish evolution. Evolution 55, 561–572.

Nizamuddin, M., 2001. Genus *Codium* Stackhouse from northern coast of the Arabian Sea (Pakistan). Pak. J. Mar. Biol. 7, 147–232.

Norris, R.E., Aken, M.E., 1985. Marine benthic algae new to South Africa. S. Afr. J. Bot. 51, 55–56.

Nyberg, C.D., Wallentinus, I., 2005. Can species traits be used to predict marine macroalgal introductions? Biol. Invasions 7, 265–279.

Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.

Omilian, A.R., Taylor, D.J., 2001. Rate acceleration and long-branch attraction in a conserved gene of cryptic daphniid (Crustacea) species. Mol. Biol. Evol. 18, 2201–2212.

Pedroche, F.F., Silva, P.C., Chacana, M., 2002. El género *Codium* (Codiaceae, Chlorophyta) en el Pacífico de México. In: Sentíes-Granados, A., Dreckman, K.M. (Eds.), Monografías Ficológicas 2002. Universidad Autónoma Metropoliana-Iztapalapa, México D.F.

Phillips, J.A., 2001. Marine macroalgal biodiversity hotspots: why is there high species richness and endemism in southern Australian marine benthic flora? Biodivers. Conserv. 10, 1555–1577.

Pond, S.L.K., Frost, S.D.W., Muse, S.V., 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676–679.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

Provan, J., Murphy, S., Maggs, C.A., 2004. Universal plastid primers for Chlorophyta and Rhodophyta. Eur. J. Phycol. 39, 43–50.

Provan, J., Murphy, S., Maggs, C.A., 2005. Tracking the invasive history of the green alga *Codium fragile* ssp. *tomentosoides*. Mol. Ecol. 14, 189–194.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rosenberg, M.S., Kumar, S., 2003. Heterogeneity of nucleotide frequencies among evolutionary lineages and phylogenetic inference. Mol. Biol. Evol. 20, 610–621.

Saunders, G.W., Lehmkuhl, K.V., 2005. Molecular divergence and morphological diversity among four cryptic species of *Plocamium* (Plocamiales, Florideophyceae) in northern Europe. Eur. J. Phycol. 40, 293–312.

Schaffelke, B., Deane, D., 2005. Desiccation tolerance of the introduced marine green alga *Codium fragile* ssp. *tomentosoides* – clues for likely transport vectors? Biol. Invasions 7, 557–565.

Scheltema, R.S., 1968. Dispersal of larvae by equatorial ocean currents and its importance to the zoogeography of shoal-water tropical species. Nature 217, 1159–1162.

Schils, T., Coppejans, E., 2003. Spatial variation in subtidal plant communities around the Socotra Archipelago and their biogeographic affinities within the Indian Ocean. Mar. Ecol. Prog. Ser. 251, 103–114.

Schils, T., Wilson, S.C., 2006. Temperature threshold as a biogeographic barrier in northern Indian Ocean macroalgae. J. Phycol. 42, 749–756.

Schmidt, O.C., 1923. Beiträge zur Kenntnis der Gattung *Codium* Stackh. Bibl. Bot. 91, 1–68.

Shapiro, B., Rambaut, A., Drummond, A.J., 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol. Biol. Evol. 23, 7–9.

Shimada, S., Hiraoka, M., Serisawa, Y., Horiguchi, T., 2004. Phylogenetic studies in the genus *Codium* (Chlorophyta) from Japan. Jpn. J. Phycol. 52, S35–S39.

Silva, P.C., 1951. The genus *Codium* in California with observations on the structure of the walls of the utricles. Univ. Calif. Pub. Bot. 25, 79–114.

Silva, P.C., 1954. Phylogenetic significance of anatomical differences in *Codium*. Proc. Eighth Int. Bot. Congr., 102–103.

Silva, P.C., 1959. The genus *Codium* (Chlorophyta) in South Africa. J. S. Afr. Bot. 25, 101–165.

Silva, P.C., 1960. *Codium* (Chlorophyta) of the tropical western Atlantic. Nova Hedwigia 1, 497–536.

Silva, P.C., 1962. Comparison of algal floristic patterns in the Pacific with those in the Atlantic and Indian Oceans, with special reference to *Codium*. Proc. Ninth Pacific Sci. Congr., 201–216.

Silva, P.C., Womersley, H.B.S., 1956. The genus *Codium* (Chlorophyta) in southern Australia. Aust. J. Bot. 4, 261–289.

Stam, W.T., Olsen, J.L., Zaleski, S.F., Murray, S.N., Brown, K.R., Walters, L.J., 2006. A forensic and phylogenetic survey of *Caulerpa* species (Caulerpales, Chlorophyta) from the Florida coast, local aquarium shops, and e-commerce: establishing a proactive baseline for early detection. J. Phycol. 42, 1113–1124.

Taylor, W.R., 1960. Marine Algae of the Eastern Tropical and Subtropical Coasts of the Americas. The University of Michigan Press, Ann Arbor, Michigan, 870 pp.

Trowbridge, C.D., 1998. Ecology of the green macroalga *Codium fragile* (Suringar) Hariot: invasive and noninvasive subspecies. Oceanogr. Mar. Biol. Ann. Rev. 36, 1–64.

Van den heede, C., Coppejans, E., 1996. The genus *Codium* (Chlorophyta, Codiales) from Kenya, Tanzania (Zanzibar) and the Seychelles. Nova Hedwigia 62, 389–417.

van der Strate, H.J., Boele-Bos, S.A., Olsen, J.L., van de Zande, L., Stam, W.T., 2002. Phylogeographic studies in the tropical seaweed *Cladophoropsis membranacea* (Chlorophyta, Ulvophyceae) reveal a cryptic species complex. J. Phycol. 38, 572–582.

Verbruggen, H., Kooistra, W.H.C.F., 2004. Morphological characterization of lineages within the calcified tropical seaweed genus *Halimeda* (Bryopsidales, Chlorophyta). Eur. J. Phycol. 39, 213–228.

Verbruggen, H., De Clerck, O., Kooistra, W.H.C.F., Coppejans, E., 2005a. Molecular and morphometric data pinpoint species boundaries in *Halimeda* section *Rhipsalis* (Bryopsidales, Chlorophyta). J. Phycol. 41, 606–621.

Verbruggen, H., De Clerck, O., Schils, T., Kooistra, W.H.C.F., Coppejans, E., 2005b. Evolution and phylogeography of *Halimeda* section *Halimeda*. Mol. Phylogenet. Evol. 37, 789–803.

Wheeler, W.C., 1990. Nucleic-acid sequence phylogeny and random outgroups. Cladistics 6, 363–367.

Wynne, M.J., 2000. Further connections between the benthic marine algal floras of the northern Arabian Sea and Japan. Phycol. Res. 48, 211–220.

Wynne, M.J., 2004. The benthic marine algal flora of the Sultanate of Oman and its biogeographical relationships. Jpn. J. Phycol. 52, S133–S136.

Xia, X., Xie, Z., 2001. DAMBE: data analysis in molecular biology and evolution. J. Hered. 92, 371–373.

Xia, X.H., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26, 1–7.

Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.

Yoshida, T., 1998. Marine Algae of Japan. Uchida Rokakuho Publishing, Tokyo, Japan, 1222 pp.

Zuccarello, G.C., West, J.A., 2003. Multiple cryptic species: molecular diversity and reproductive isolation in the *Bostrychia radicans*/*B. moritziana* complex (Rhodomelaceae, Rhodophyta) with focus on North American isolates. J. Phycol. 39, 948–959.