

# CHAPTER 4

## BIOGEOGRAPHY AND THERMAL NICHE EVOLUTION OF WARM-TEMPERATE MARINE MACROALGAE IN THE WESTERN INDO-PACIFIC OCEAN

Lennert Tyberghein<sup>1</sup>, Heroen Verbruggen<sup>2</sup>, Olivier De Clerck<sup>1</sup>

<sup>1</sup> Phycology Research Group, Biology Department, Ghent University, Krijgslaan 281 S8, 9000 Ghent, Belgium

<sup>2</sup> School of Botany, University of Melbourne, Victoria, 3071 Australia

Manuscript in prep.

# ABSTRACT

Here, an empirical approach is used to gain understanding in the biogeography and evolution of warm-temperate marine floras. The focus lies on the marine green algal *Codium decorticatum* species complex. We use climatic niche modeling and divergence time estimation to explore the historical biogeography and evolution of thermal variables within the complex. The climatic models were projected to the Last Glacial Maximum and the Pliocene to examine whether the Indian Ocean served as a thermal barrier for dispersal. Our results suggest that during the late Miocene – early Pliocene species within the *C. decorticatum* complex dispersed via episodic dispersal events. After geographic isolation, species have adapted to local conditions and their thermal niches have diverged from their closely related species.

# INTRODUCTION

Knowledge of how species' climatic niches evolve can help us understand how they are distributed. Analyzing distributions of closely related taxa between geographically separated regions can yield important information about evolutionary lability of these climatic relationships or, alternatively, about their conservatism (Qian & Ricklefs, 2004). Both the tendency for related taxa to retain (e.g. Kozak & Wiens, 2006) or shift (e.g. Graham *et al.*, 2004) their ancestral ecological niche has been confirmed in the last decade and both mechanisms are assumed to promote diversification in different ways. The allopatric speciation model has dominated theories of speciation for decades. Also in the marine environment diversification as a result of long-term geographic isolation is regarded as the most prevalent mode of speciation, even though barriers to gene flow are often less visible (Palumbi, 1994). The emphasis, however, has always resided with the purely geographical setting while relatively little attention has been given to the possible influence of environmental and climatic factors that favor isolation between populations of marine organisms.

Among marine organisms, macroalgae are good models for evolutionary, macroecological and biogeographical research because of their broad distribution, variability in ecological dimensions and their sessile nature making them easier to collect (Bolton *et al.*, 2004; Verbruggen *et al.*, 2009b). One of the most mesmerizing patterns of macroalgal biodiversity is the apparent affinity of the marine floras of distant subtropical warm temperate regions in the Western Indo-Pacific, namely the Arabian Sea, SE Africa and SW Australia (Norris & Aken, 1985; Joosten & Van Den Hoek, 1986; Lüning, 1990). These regions harbor a rich species diversity with many endemic species (Phillips, 2001; Bolton *et al.*, 2004; Schils & Wilson, 2006) and similarities between them have often been depicted (Wynne, 2000; Schils & Coppejans, 2003; Verbruggen *et al.*, 2007; Bolton, 2010) and genetically confirmed (Pillmann *et al.*, 1997; Hommersand & Fredericq, 2003; Verbruggen *et al.*, 2007; De Clerck *et al.*, 2008). The marine floras of these regions also show similarities with the warm temperate flora of Japan (Børgesen, 1934; Hommersand, 1986; Wynne, 2000). Aside from the overall similarity of these regions' algal floras in terms of species diversity, the distinct regions feature morphologically similar congeners that are absent from the tropical seas separating them (Verbruggen *et al.*, 2007).

Several plausible explanations for the affinities between these distant subtropical algal floras have been proposed. To elucidate such affinities between S Africa and Australia Norris and Aken (1985) suggested that common taxa represent vicariant relics of a continuous distribution along the Cretaceous coast of Gondwanaland, distributions that were eventually fragmented by the northward migration of Africa, Australia and the Indian subcontinent. Hommersand (1986; 2007) proposed that descendents of lineages that originally evolved in warm-temperate SW Australia migrated through the Indian Ocean via the North Equatorial current as far as the east coast of S Africa and, occasionally, beyond the southern tip of Africa into the Atlantic Ocean. He postulated that such migrations took place during Miocene-Pliocene periods during which global sea surface temperatures decreased markedly. An alternative to this dispersal based scenario invokes a common ancestor that was distributed at low latitudes throughout the Indian Ocean during cool periods, with subsequent isolation in relatively small geographic areas taking place during warmer climatic periods (Verbruggen *et al.*, 2005).

Despite the existence of various biogeographic hypotheses, few studies have addressed the association between algal lineage diversification and its environmental setting in an explicit timeframe (but see Verbruggen *et al.*, 2009b). Phylogenetic analyses with relaxed molecular clock methods provide a means of tracing the evolution of a lineage in a temporal framework. Furthermore, biogeographical research and studies of speciation have rarely quantified

environmental factors that might have caused sister species to be allopatric (Kozak *et al.*, 2008). Yet several tools are available to estimate ecological niches of species and predict their responses to environmental change, so called ecological niche modeling (ENM). Integration of niche modeling techniques with time-calibrated phylogenies can yield valuable information about niche shifts and the evolution of environmental preferences (Yesson & Culham, 2006b; Yesson & Culham, 2006a). These advances could provide new understanding of biogeographical patterns and niche dynamics.

Here, we present an empirical approach to gain understanding in the biogeography and macroecological evolution of warm-temperate marine floras. More specifically, we focus on the green algal genus *Codium*, a widely distributed group that has its largest species diversity in warm-temperate regions. Our main goal is to understand the biogeography and niche evolution of a *Codium* species complex distributed in warm-temperate regions and to interpret observed patterns in consideration of the biogeographical hypotheses mentioned above. To achieve our goals, we combine phylogenetic, distributional and environmental data. First, we define a temporal framework for lineages within this complex. Second, we determine whether the thermal niches of these closely related entities have diverged or if they show a high degree of overlap. Third, we check if ecological niche models can be used to inform us about dispersal barriers and biogeographical patterns in geological time.



# MATERIALS AND METHODS

## STUDY SPECIES

We focus on the *Codium decorticatum* complex, a clade containing five morphologically similar species present in different warm-temperate regions (Verbruggen *et al.*, 2007). *Codium cylindricum* (Holmes, 1896) is a Pacific species that occurs along the coasts of Japan and on both sides of the Baja California peninsula, Mexico. In the latter region, the species is known as *Codium amplivesiculatum* (Setchell & Gardner, 1924; Norris, 2010) but our genetic data show that it is conspecific with *Codium cylindricum*. *Codium decorticatum* (Howe, 1911) has a very broad geographical area, ranging from the North Atlantic (Bermuda, Azores and North Africa) to the Caribbean and along the south coast of South Africa and into the Indian Ocean (Madagascar). The remaining three species have not been formally described and are here referred to as *C. duthieae* 1, 2 and 3 (Verbruggen *et al.*, 2007). The geographical area of *Codium duthieae* 1 is restricted to the Eastern coast of South Africa ranging just to the South of Mozambique. *Codium duthieae* 2 solely occurs on the southern Australian and Tasmanian coast. *Codium duthieae* 3 occurs in the Arabian Sea around the coasts of Oman, Yemen, India and Pakistan and was also recently collected in Madagascar.

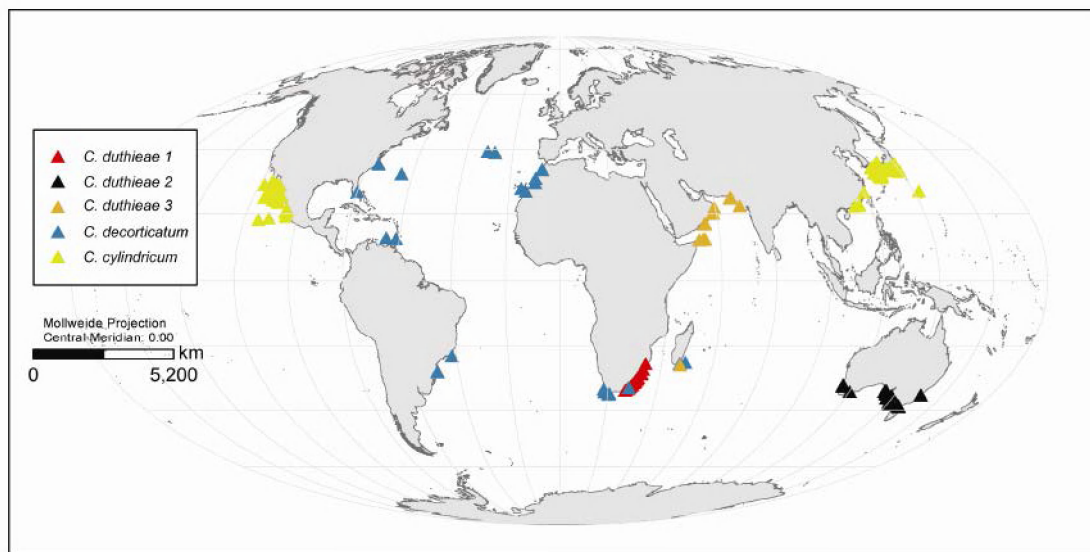


Figure 1: Species records of all *Codium decorticatum* complex species.

## PHYLOGENETIC INFERENCE AND DIVERGENCE TIME ESTIMATION

Three plastid markers (*rbcL* exon1, *tufA* and UCP6) were used to reconstruct phylogenetic relationships for 21 *Codium* species. For the *Codium decorticatum* complex and five other closely related species four extra markers were added (*rbcL* exon2, *atpA*, *atpA* intron and UCP7) (Table S1). Using the Bayesian information criterion (BIC), a suitable partition strategy and suitable models of sequence evolution were selected. The guide tree used during the entire procedure was obtained

by a maximum-likelihood (ML) analysis of the unpartitioned concatenated alignment using a JC +  $\Gamma$ 8 model (Guindon & Gascuel, 2003). All likelihood optimizations and BIC calculations were carried out with Treefinder (Jobb *et al.*, 2004). The preferential model of sequence evolution and partitioning strategy (lowest BIC) was a GTR + G4 for which all genes were partitioned in three codon positions. This model was applied in all subsequent phylogenetic analyses.

ML analyses were performed using RAxML v. 7.2.8 (Stamatakis, 2006). Statistical branch support was assessed via bootstrapping with 100 replicates. Bayesian phylogenetic inference (BI) was carried out with MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003). Two parallel runs, each consisting of four incrementally heated chains were run for 2 million generations, sampling every 2,000th generation. Convergence of log-likelihoods and parameter values was assessed in Tracer v1.5 (Rambaut & Drummond, 2007). An appropriate burn-in sample of trees (5%) was removed before constructing the majority rule consensus tree.

To obtain an ultrametric tree, a Bayesian phylogenetic analysis was conducted in BEAST v1.7.2 (Drummond *et al.*, 2012) under a GTR+G4 model with an uncorrelated lognormal (UCLN) relaxed molecular clock model and using a Yule tree prior. Three Markov Chain Monte Carlo (MCMC) analyses were run for 10 million generations, sampling every 10,000th generation. The output was diagnosed for convergence using Tracer v1.5, and summary statistics and trees were generated using the last 9 million generations with TreeAnnotator v1.5.3 (Rambaut & Drummond, 2007). The tree was calibrated by assuming an age of 153 (SD 35) million years (Myr) for the crown node of the genus (Verbruggen *et al.*, 2009a).

## NICHE MODELING

We used a correlative ecological niche modeling approach to estimate the climatic niche of species within the *Codium decorticatum* complex using species occurrences and climatic data. We also projected the inferred niche onto past climatic circumstances (Last Glacial Maximum and Pliocene) in order to visualize potential climatic barriers.

### Occurrence data

Species occurrence records are primarily obtained from recently collected specimens for which the identification has been confirmed with DNA barcodes. These records are complemented with herbarium specimens and verified literature data that were carefully re-examined. For specimens that were not geotagged in the field, georeferencing was conducted using Google Earth, only including records that could be unambiguously assigned to a precise locality. ArcGIS (ESRI, Redlands, California, USA) was used to map the coordinates for visual verification. Our final dataset of unique georeferenced records ranged between 24 (*C. duthieae* 1) and 119 (*C. cylindricum*) unique localities per species (Table S2).

### Climate data

Climate data were derived for three different time periods. For present-day conditions, we used Bio-ORACLE, a marine environmental dataset for global-scale applications, to derive sea surface temperature (SST) data (minimum, maximum and mean) (Tyberghein *et al.*, 2012). Climate data for the Last Glacial Maximum (LGM) were obtained from model-interpolated GLAMAP data (Paul & Schäfer-Neth, 2003). The global monthly LGM SST data originally had a spatial resolution of one



degree and were interpolated (bilinear) to match the 9km resolution of Bio-ORACLE. SST data for the Pliocene was derived from the Pliocene Research, Interpretation and Synoptic Mapping Project (PRISM3) (Dowsett *et al.*, 2009; Dowsett *et al.*, 2010) at a spatial resolution of 2 degrees (left at default spatial resolution). This dataset documents the global climate in the mid-Pliocene warm period using various proxies (e.g., foraminifera, diatoms, ostracods, pollen and plant macrofossil data). To be consistent with the Bio-ORACLE data set, we built the same types of grids (minimum, maximum, mean and SST range) based on the monthly data for both past periods (LGM & Pliocene).

Because the 4 temperature-related layers we use are correlated, principal component analysis (PCA) was conducted in ArcGIS to reduce data dimensionality and generate a set of new independent variables from the original unstandardized Bio-ORACLE data. The first two PCA axes accounted for more than 99% of the underlying environmental variation. The factor loadings from this analysis were used to construct PC axes for both the LGM and Pliocene time period. These PC axes were used for subsequent niche modeling.

## Modeling technique

We reconstructed niche models using MaxEnt3.3.3k (Phillips *et al.*, 2006; Phillips & Dudik, 2008). This machine learning algorithm yields largely better results than other presence-only niche modeling methods (Elith *et al.*, 2006) and performs well even when the number of occurrence records available for modeling is small (e.g. Wisz *et al.*, 2008). MaxEnt allows for model performance evaluation by calculating the area under the ROC (receiver operating characteristic) curve (AUC), based on training and test data, which represent the ability of the model to distinguish presence data from background data (Phillips *et al.*, 2006). All analyses were replicated 10 times using a subsampling approach with 30% of localities for model testing. Background points were randomly chosen based on an equal area map in order not to bias towards the higher latitudes. All other settings were left at their defaults.

MaxEnt was also used to predict areas of suitable thermal conditions during the LGM and Pliocene by creating the models on the species' present thermal distribution range and projecting them onto past sea surface temperature grids.

## NICHE COMPARISONS

We compared the thermal niches of all five members of the *Codium decorticatum* complex using several methods. First, we examined the level of divergence in environmental space. We conducted a PCA on all species within the complex to assess the overall level of environmental divergence but also quantified how pairs of sister taxa overlapped in environmental space. Second, we calculated overlap of the inferred ecological niche models, which are a representation of the species' environmental affinities in geographic space.

A PCA was conducted with all five species of the *Codium decorticatum* complex. Data for all occurrence localities were extracted from the original Bio-ORACLE temperature layers. We also assessed niche overlap for each species pair from separate PCAs. Statistical significance of the possible niche shifts between species was evaluated using multivariate analyses of variance (MANOVAs) of the first and second principal component, a method fully detailed by Graham *et al.* (2004). Statistical analyses were conducted using R (R Development Core Team, 2008).

As a complement to the PCA analyses, we also used computed overlap statistics using niche models. The predicted thermal niche of all five species was compared using indices for niche overlap D and I, as proposed by Warren et al. (2008) implemented in the package ENMTools 1.3 (Warren *et al.*, 2010). The significance of both indices, ranging from 0 (no overlap) to 1 (identical niche models), was evaluated with two randomization tests (detailed in Warren *et al.*, 2008). First, we tested for 'niche equivalency' by comparing the predicted niches to a null distribution (100 pseudoreplicates) that was generated by randomly resampling the pooled sample of occurrence points for both species with the condition that sample sizes for both species were kept constant. This procedure tests whether the observed niche similarity differs significantly from that expected under the null hypothesis that the two species are ecologically equivalent.

## ANCESTRAL NICHE RECONSTRUCTIONS

To examine thermal niche evolution in the *Codium decorticatum* complex, we reconstructed ancestral states for all SST variables. We used Bayesian MCMC to construct the posterior distribution for the states at internal tree nodes as implemented in the Phytools R package (Revell, 2011). We used extracted temperature data from sampling localities and add the five closest *Codium* taxa in order to get a more comprehensive picture of the ancestral niche of the *C. decorticatum* clade.

In addition, we applied an approach that allowed taking intraspecific variability of thermal preferences into account (Evans *et al.*, 2009). Niche models created by MaxEnt are combined with individual SST layers to produce a probability distribution for each variable (Predicted Niche Occupancy (PNO)). This distribution of thermal tolerances gets sampled for data to reconstruct. We also took phylogenetic uncertainty into account by repeating this analysis over 1000 samples from the posterior distribution of chronograms. The estimation of ancestral climatic tolerances sensu Evans et al. (2009) was conducted in the R package Phyloclim (<http://www.christophheibl.de/Rpackages/>).



# RESULTS

## PHYLOGENY AND DIVERGENCE TIME ESTIMATION

The resulting phylogenetic trees were generally congruent with previous hypotheses for this group (Verbruggen *et al.*, 2007). Thirteen of the 19 nodes in the tree were recovered with Bayesian posterior probabilities greater than 0.95 (Figure S1). Phylogenetic relationships remained poorly resolved among the three *Codium duthieae* species despite the addition of four extra plastid markers.

The root age used this study was based on previous molecular dating of siphonous green algae (Verbruggen *et al.*, 2009a). Using this root age, the estimated crown group age for the *Codium decorticatum* complex is about 13.4 Myr (24.5-4.1 Myr 95% highest posterior density [HPD]). *Codium decorticatum* diverged during the Late Miocene at about 7.5 Myr ago while the three *Codium duthieae* lineages diverged in the Early Pliocene (~ 5.3-3.6 Myr ago).

## ECOLOGICAL NICHE ANALYSES

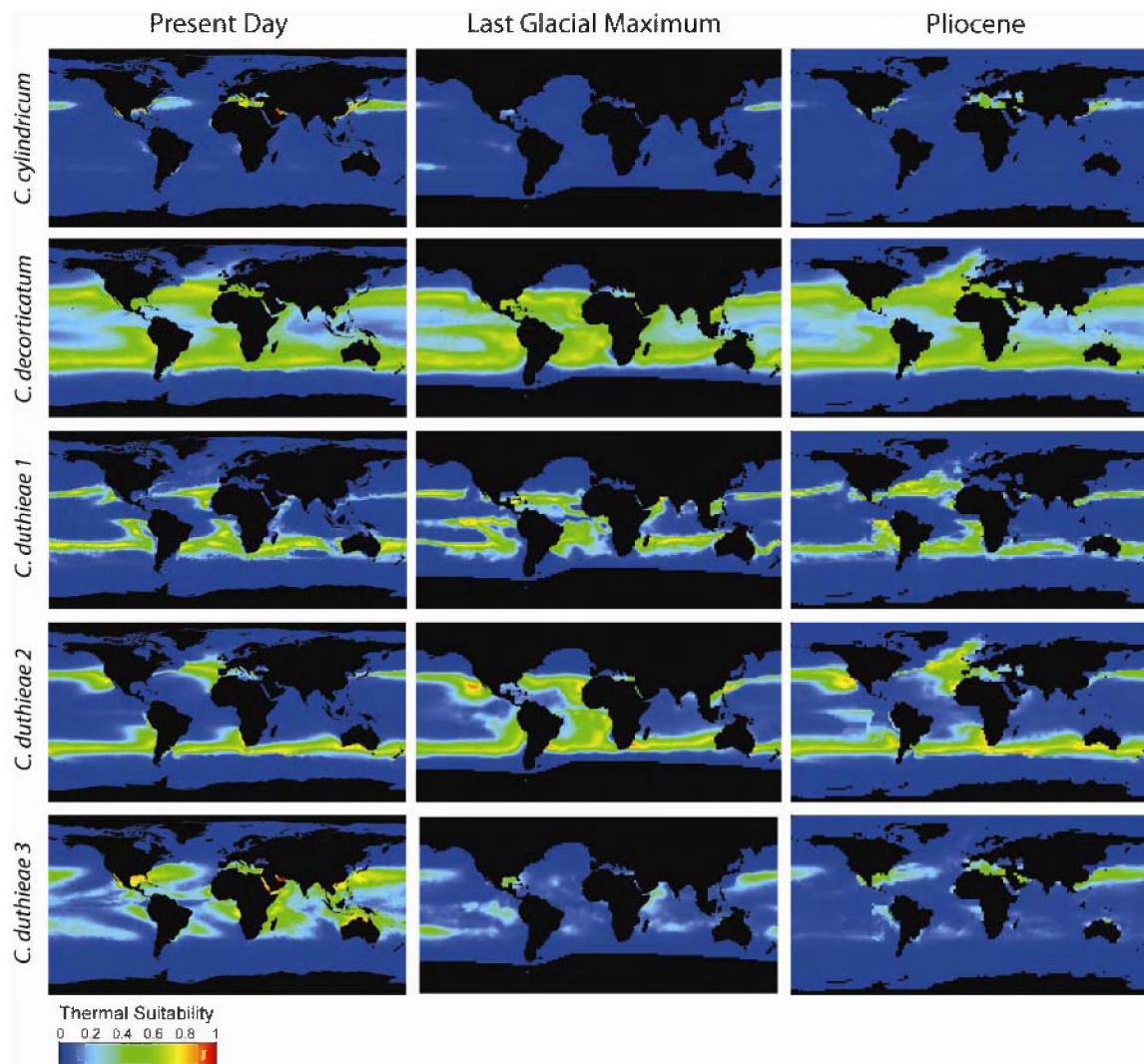
The ecological niche models of all species within the *Codium decorticatum* complex had mean test AUC values ranging from 0.833 to 0.974 (Table 1, Figure 2) indicating reliable model performances (Phillips *et al.*, 2006). Overall, maps of predicted species distribution showed broad suitability outside their realized ranges. These areas show similar temperature conditions indicating the potential distribution of the species. One exception is the niche model for *Codium cylindricum*, which shows a less extensive extrapolation towards other areas.

**Table 1: Mean test AUC values (area under the receiver operation characteristics curve) of niche models per species, Schoener's index D and Hellinger's index I of niche overlap, and assessment of niche equivalency via randomization tests.**

| Species X              | Species Y              | AUC (±SD) X   | AUC (±SD) Y   | D        | I        |
|------------------------|------------------------|---------------|---------------|----------|----------|
| <i>C. duthieae 1</i>   | <i>C. cylindricum</i>  | 0.974 (0.008) | 0.968 (0.011) | 0.198*** | 0.445*** |
| <i>C. duthieae 1</i>   | <i>C. duthieae 2</i>   | 0.974 (0.008) | 0.944 (0.006) | 0.340*** | 0.632*** |
| <i>C. duthieae 1</i>   | <i>C. duthieae 3</i>   | 0.974 (0.008) | 0.950 (0.026) | 0.282*** | 0.574*** |
| <i>C. duthieae 1</i>   | <i>C. decorticatum</i> | 0.974 (0.008) | 0.833 (0.029) | 0.528*** | 0.806*** |
| <i>C. duthieae 2</i>   | <i>C. cylindricum</i>  | 0.944 (0.006) | 0.968 (0.011) | 0.207*** | 0.467*** |
| <i>C. duthieae 2</i>   | <i>C. duthieae 3</i>   | 0.944 (0.006) | 0.950 (0.026) | 0.056*** | 0.190*** |
| <i>C. duthieae 2</i>   | <i>C. decorticatum</i> | 0.944 (0.006) | 0.833 (0.029) | 0.569*** | 0.840*** |
| <i>C. duthieae 3</i>   | <i>C. cylindricum</i>  | 0.950 (0.026) | 0.968 (0.011) | 0.318*** | 0.593*** |
| <i>C. duthieae 3</i>   | <i>C. decorticatum</i> | 0.950 (0.026) | 0.833 (0.029) | 0.342*** | 0.615*** |
| <i>C. decorticatum</i> | <i>C. cylindricum</i>  | 0.833 (0.029) | 0.968 (0.011) | 0.401*** | 0.724*** |

Significance values (95% level) of niche overlap are indicated by asterisks; \*\*\*P < 0.001

Paleo-ENMs for the Pliocene show similar predictions as for present day models (Figure 2). The projection for *C. duthieae* 3 predicts remarkably less suitable habitat in the Pliocene. For *C. duthieae* 1, *C. duthieae* 2 and *C. decorticutum*, predictions onto LGM circumstances show similar geographical patterns, however it is clearly visible that predictions narrow towards the tropics. Projections for *C. duthieae* 3 and *C. cylindricum* yielded less predicted suitable area.



**Figure 2: Predicted thermal niche models of *Codium* species inferred from temperature layers and occurrence points. The left column depicts present day habitat suitability. The second and third column represent projections onto past temperature circumstances, respectively the Last Glacial Maximum (~18,000 kyr) and the Pliocene (Mid Piacenzian 3.29 – 2.97 myr). Predicted habitat suitability is indicated with colors along a gradient; warmer colors indicating areas with better predicted conditions.**



Thermal profiles show that all five *Codium* species primarily occur in warm temperate environments (Table 2). One species, *C. duthieae* 2 experiences somewhat colder temperatures in S Australia, whereas the Arabian *C. duthieae* 3 mainly occurs in warmer waters. Environments of the Pacific species *C. cylindricum* that occur both along the coasts of Japan and Baja California, are characterized by temperatures ranging from 20.7-32.2 °C in the warmest month to 17.4-25.1 °C in the coldest month. Predicted niche occupancy (PNO) plots (Figure 3) indicate a broad range of suitable temperatures for *Codium decorticatum*, a species occurring in both tropical and temperate areas. The PNO profile for *Codium cylindricum* clearly shows the broad temperature range of this species.

**Table 2:** Thermal data for the *Codium decorticatum* clade. Sea Surface temperature data in degrees Celsius, standard deviation between parentheses.

| Species                    | maximum sst  | mean sst     | minimum sst  | sst range    |
|----------------------------|--------------|--------------|--------------|--------------|
| <i>Codium cylindricum</i>  | 28.27 (2.25) | 22.31 (2.52) | 17.21 (3.15) | 11.05 (2.81) |
| <i>Codium decorticatum</i> | 24.99 (3.37) | 21.86 (3.37) | 19.17 (3.55) | 5.82 (1.80)  |
| <i>Codium duthieae</i> 1   | 25.62 (1.52) | 22.86 (1.38) | 20.49 (1.31) | 5.13 (0.38)  |
| <i>Codium duthieae</i> 2   | 20.92 (2.02) | 17.86 (2.03) | 15.01 (2.25) | 5.91 (1.67)  |
| <i>Codium duthieae</i> 3   | 28.89 (0.77) | 26.21 (0.51) | 22.65 (0.97) | 6.24 (1.17)  |

**Table 3:** Results of MANOVA. This analysis, using the PC scores as dependent variables and species as categorical variables, shows differences in the thermal space of the five species using the first two principal component scores. The multivariate analysis represents an overall statistic to determine whether separation in environmental space was significant while the univariate analysis is a test to determine which PCA axis accounts for significance in the overall test.

| Species comparison                             | multivariate |  | univariate  |             |
|--|--------------|--|-------------|-------------|
|  | F stat       |  | F stat PC1  | F stat PC2  |
| <i>C. duthieae</i> 1 - <i>C. cylindricum</i>   | 146***       |  | 15.724***   | 83.952***   |
| <i>C. duthieae</i> 1 - <i>C. duthieae</i> 2    | 68***        |  | 108.15***   | 0.2383 (Ns) |
| <i>C. duthieae</i> 1 - <i>C. duthieae</i> 3    | 50***        |  | 104.58***   | 0.0282 (Ns) |
| <i>C. duthieae</i> 1 - <i>C. decorticatum</i>  | 50**         |  | 9.1588**    | 2.3048 (Ns) |
| <i>C. duthieae</i> 2 - <i>C. cylindricum</i>   | 165***       |  | 156.39***   | 48.832***   |
| <i>C. duthieae</i> 2 - <i>C. duthieae</i> 3    | 69***        |  | 330.03***   | 0.3282 (Ns) |
| <i>C. duthieae</i> 2 - <i>C. decorticatum</i>  | 69***        |  | 24.652***   | 1.3546 (Ns) |
| <i>C. duthieae</i> 3 - <i>C. cylindricum</i>   | 147***       |  | 68.991***   | 10.005**    |
| <i>C. duthieae</i> 3 - <i>C. decorticatum</i>  | 51***        |  | 62.509***   | 0.1032 (Ns) |
| <i>C. decorticatum</i> - <i>C. cylindricum</i> | 147***       |  | 0.8887 (Ns) | 98.325***   |

Significance values are indicated by asterisks; \*\*\*P < 0.001, \*\*P < 0.01, \*P < 0.05, Ns: not significant.



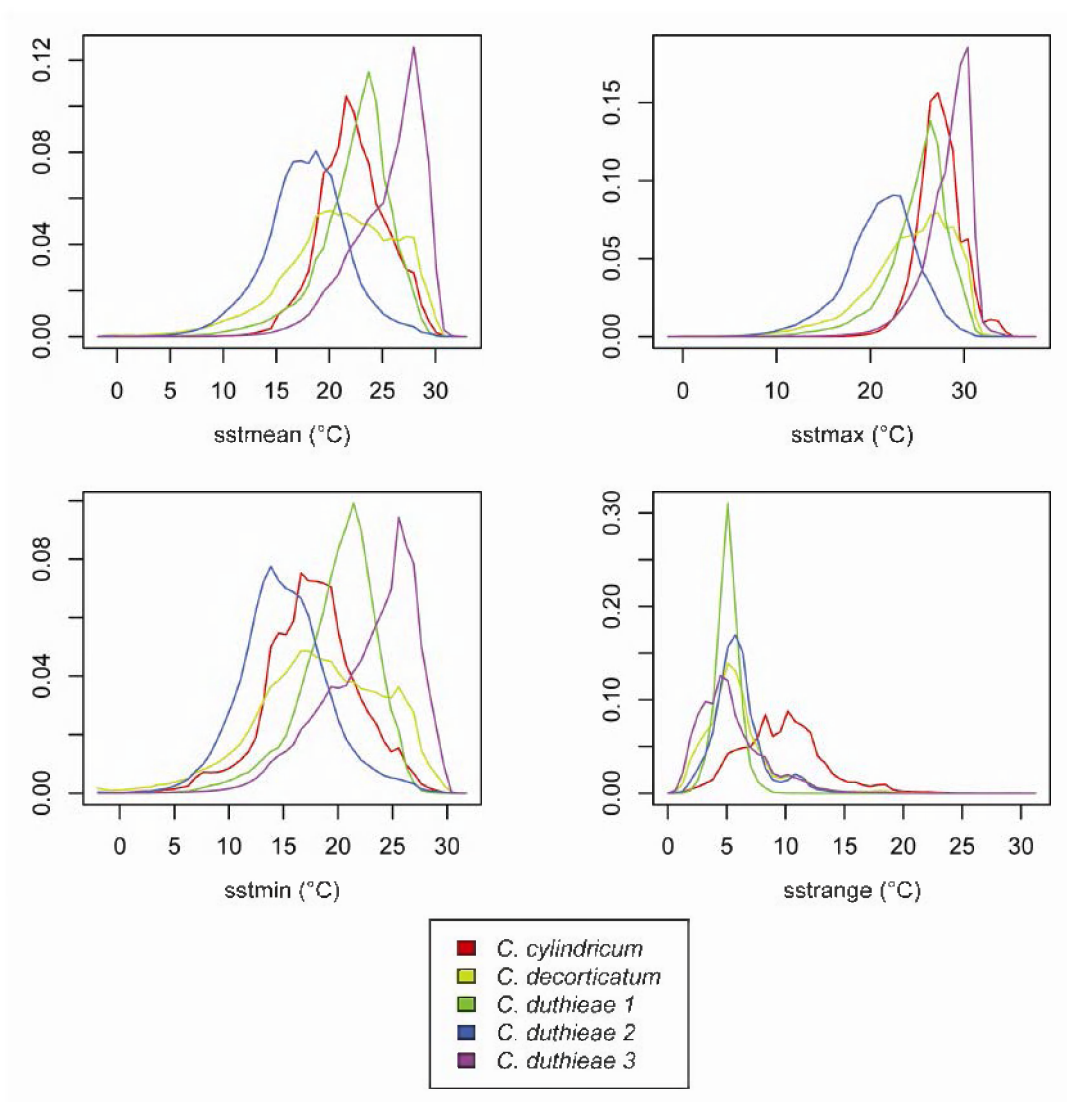
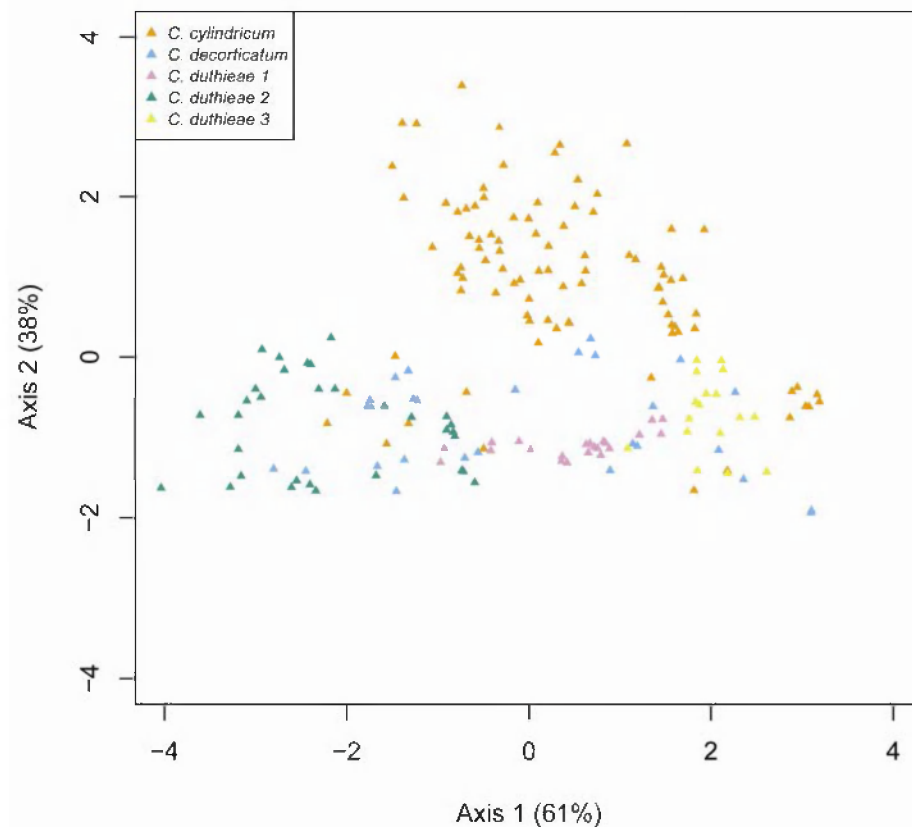


Figure 3: Predicted niche occupancy of the 5 taxa in the *Codium decorticatum* complex. These plots represent unit area histograms of suitability, illustrating a species' (predicted) occupancy of each sst variable. Pictures were developed using the R package Phylolclim (<http://www.christophheibl.de/Rpackages>).

or the overall PCA on all five species within the *Codium decorticatum* clade, the first and second principal components explained nearly all variation (PC1 = 61.5% and PC2 = 38.2%) among the extracted data from the four Bio-ORACLE layers (Figure 4, Table S4). Mean, maximum and minimum sea surface temperature loaded on the first axis, whereas SST range was responsible for the factor loading on the second axis. MANOVAs were calculated on both PC1 and PC2 on all species pairs within the *C. decorticatum* clade (see Table S4 for detailed results of PCAs). Overall separation in environmental space was statistically significant for all species pairs (Table 3). Post-hoc analyses indicated that PC1 was responsible for this separation in all cases except for *C. decorticatum* / *C. cylindricum*. Divergence along PC2 was only statistically significant in species comparisons with *C. cylindricum*. This result is expected and corroborates the observation that *C. cylindricum* grows in areas with substantial seasonal temperature differences.



**Figure 4:** Principal components analysis depicting the position of the 5 focal *Codium* species in thermal niche space, with each data point representing a unique species locality.

Overlap of ecological niche models resulted in D values ranging from 0.06 to 0.57 and I values ranging from 0.19 to 0.84 (Table 1). Values of 0 symbolize no overlap whereas values of 1 indicate identical niche models. We find the highest niche overlap between *C. decorticatum* and *C. duthieae* 2 and between *C. decorticatum* and *C. duthieae* 1. Niche overlap between *C. duthieae* 2 and *C. duthieae* 3 is noticeably very low.

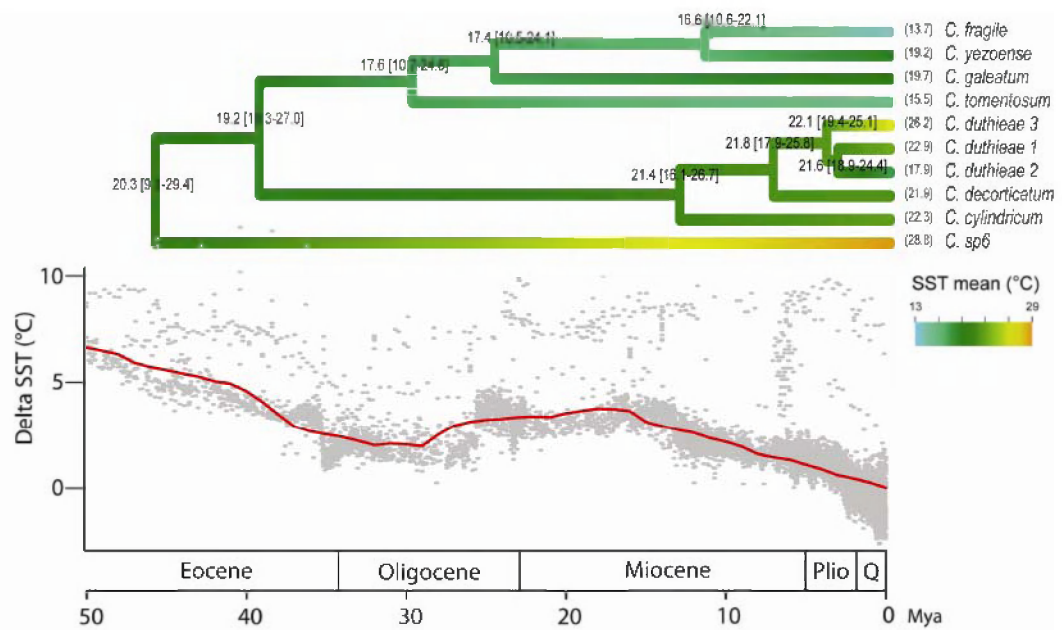


Figure 5: Calibrated phylogenetic tree of 10 *Codium* species including the *C. decorticans* complex. Ancestral state estimations for mean sea surface temperature are plotted on the tree. The mean ancestral state of 1,000,000 Bayesian generations is given in the branch nodes and indicated by a color gradient (95% confidence interval between square brackets). At the bottom, the ocean temperature curve (in red) is based on a high-resolution deep-sea oxygen isotope record (grey dots;  $\delta^{18}\text{O}$ ; Veizer *et al.*, 1999). Delta SST represents the temperature difference between past and current temperatures. A 10-Ma geological time-scale indicates the timeframe, spanning epochs since 50 Ma (Plio, Pliocene; Q, Quaternary).

A mean sea surface temperature affinity of 21.4°C (95% HPD: 16.1-26.7) was inferred for the basal split of the *Codium decorticans* complex (Figure 5). *Codium duthieae* 3 evolves towards warmer waters, while the opposite is true for *Codium duthieae* 2. Similar results were obtained in the analysis of the history of niche occupancy for separate temperature variables (Figure 6).



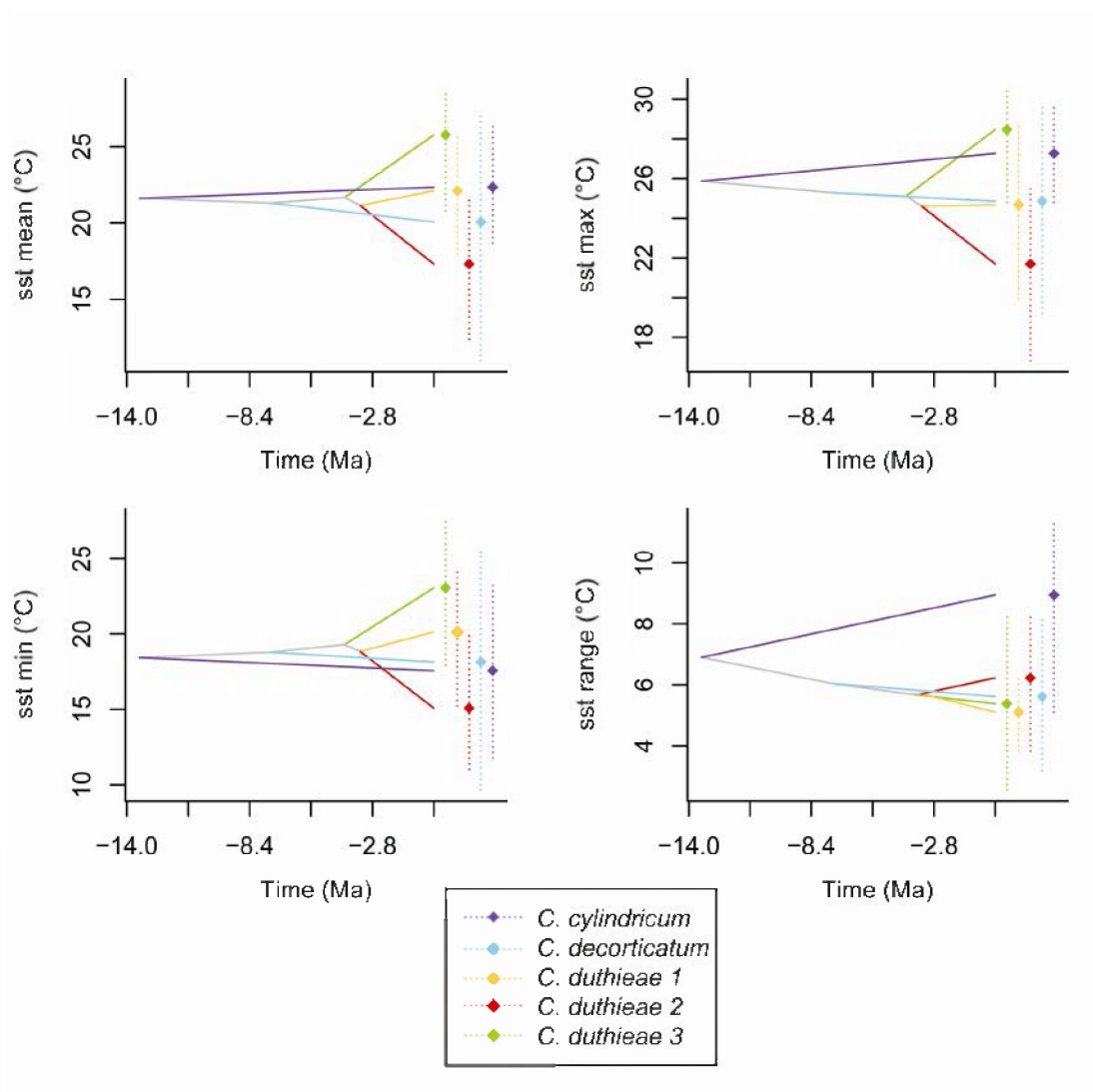


Figure 6: Ancestral reconstructions of thermal niches. The x-axis represents time. The 80% central density of thermal tolerance for each extant species is indicated by a vertical dashed line, the mean by a diamond symbol. The graphs were developed using the R package Phyloclim (<http://www.christopheibl.de/Rpackages>).

# DISCUSSION

The approach presented here explores the ability to study macroecological and biogeographical patterns using a combination of phylogenetic, geographic and environmental data. The temporal framework of diversification in combination with the results from thermal niche studies allow us to discuss observed patterns of subtropical to temperate algal floral affinities in light of previously proposed hypotheses.

## BIOGEOGRAPHY

According to our phylogenetic results, the *Codium decorticatum* complex diversified from the Mid-Miocene (13.4 mya) onwards with the most recent inferred split dated in the Pliocene (3.4 mya). Considering the distributions of *Codium* species in this study, it is unlikely that tectonic events affected the diversification of the complex. It was hypothesized that disjunct warm temperate species ranges in the Western Indo-Pacific could represent vicariant relics of a previously continuous distribution along the Cretaceous coast of Gondwanaland at a time when the connection between the coastlines of Africa and Australia were more or less continuous (Norris & Aken, 1985; Hommersand, 1986). Our results refute this hypothesis for the *C. decorticatum* complex, as species divergence times are much too recent to account for a Cretaceous common ancestor.

Instead, a dispersal-based hypothesis, originally proposed by Hommersand (1986; 2007), is far more likely to account for the current distributions of all species within the *C. decorticatum* complex. Hommersand suggested that species which originally evolved in warm-temperate environments of SW Australia have descendents that migrated through the Indian Ocean via the North Equatorial current as far as the east coast of S Africa and, exceptionally beyond, into the Atlantic Ocean. Likewise, species originated in warm-temperate climates in East Asia could migrate through the tropics via the Arabian Sea into S Africa. *Codium duthieae* 2 (SW Australia), *C. duthieae* 3 (Arabian Sea), *C. duthieae* 1 (SE Africa), *C. decorticatum* (Atlantic Ocean) and *C. cylindricum* (Japan) seem to support this hypothesis of biogeographic affinities. Species intolerant of tropical temperatures may have dispersed across equatorial waters at shallow depths during periods of climatic cooling (Darwin, 1859; Hubbs, 1952; Lindberg, 1991). It was assumed that such migrations took place during Miocene-Pliocene periods in which the temperature decreased markedly. The divergence times inferred in the present study corroborate this time setting.

During the Pliocene (~5-2 Myr ago), when the *C. duthieae* species diverged, global sea surface temperatures were approximately 3°C warmer than at present (Ravelo *et al.*, 2004). Our model approach permitted to predict areas of suitable thermal conditions during this period. We realize that this projection represents only a snapshot (3.29 – 2.97 myr ago) in geological time; however, it clearly indicates that the Indian Ocean was too warm to allow for a continuous range of the common ancestor of the *C. duthieae* clade. Model projections onto the LGM (~18 kyr ago), the period in the Earth's Cenozoic history when ice sheets reached their maximum extent, show that even during these cool periods the low latitudes of the Indian Ocean were not predicted as suitable. Hence, the current distribution pattern among *C. duthieae* species can only have arisen by means of long-distance dispersal. Only for *Codium decorticatum*, models indicate a broad range across low latitudes during glacial maxima, which is still reflected in the species broad geographic distribution which spans tropical to temperate regions in the Atlantic Ocean.



Dispersal among marine organisms is considered to be a function of larval behavior, life history strategies, drifting capabilities, oceanographic features, habitat availability and environmental limits (Bellwood *et al.*, 1998; Cowen *et al.*, 2000; Gillespie *et al.*, 2012). In *Codium*, mechanisms for dispersal include motile flagellated reproductive cells (Ramus, 1972) and thallus fragmentation (Fralick & Mathieson, 1972). Occasional long-distance dispersal is possible as detached thalli can float due to internal oxygen entrapment (Dromgoole, 1982). This, in combination with the ability to withstand temporary desiccation (Schaffelke & Deane, 2005), makes successful long distance dispersal possible. The observed distribution pattern, whereby 3 entities are nearly exclusively confined to warm temperate pockets in the Indian Ocean despite their divergence approximately 3.4 mya, serves as evidence for the exceptional nature of such events. Likewise, the presence of *C. duthieae* 3 in southern Madagascar provides evidence, be it circumstantial, that *Codium* species are capable of occasionally dispersing over long distances and bridging large areas of unsuitable environmental conditions. At such low frequency, dispersal events are so rare that they can allow for genetic differentiation (bottlenecks and Founder effects) and eventually even allopatric speciation (Thiel & Haye, 2006).

As to the directionality of dispersal events, our phylogeny remains largely inconclusive. The short branch lengths separating the *C. duthieae* entities and lack of support keep us from putting forth a specific hypothesis. Furthermore, hypotheses about directionality can be seriously confounded by extinction events (Crisp *et al.*, 2011), the importance of which is difficult to estimate based on a phylogeny of extant taxa (e.g. Rabosky, 2009; Rabosky, 2010; Morlon *et al.*, 2011; Stadler, 2011). Hommersand (1986) favored dispersal from Australia through the northern part of the Indian Ocean to South Africa by way of the islands of Indonesia and the North Equatorial Current. We should also consider there was also a possibility for dispersal from South Africa to Australia, for example via the South Indian Current (southern part of the Indian Ocean gyre). The fact that most of the warm, salty waters of the Agulhas current, the largest ocean jet current in the Southern Hemisphere, feeds back into the subtropical gyre (Zahn, 2009), lends credibility to this theory. Alternatively, migration by means of the Antarctic circumpolar current (west wind drift) could also be possible. A phylogeographic study on a widespread coastal sea star (*Patiriella exigua*) revealed colonization of Australia from southern Africa via the west wind drift (Waters & Roy, 2004). Such dispersal pattern is not unique among marine organisms and has since gained more evidence (Waters, 2008; Fraser *et al.*, 2009; Zemlak *et al.*, 2009; Macaya, 2010; Nikula *et al.*, 2010). It seems that most taxa for which evidence of long distance dispersal by means of the Antarctic Circumpolar Current has been presented have affinities for cold temperate rather than warm temperate waters. It is questionable whether a warm-temperate *Codium* species could reach the Antarctic Circumpolar Current and as such be dispersed via this route. Furthermore, species would have to endure suboptimal temperatures for a very long time in order to reach the coast of Australia.

## NICHE EVOLUTION

The climatic analyses and ENM results for the *C. decorticatum* clade yield some insights in the evolution of the thermal niche. All five species are associated with temperate to tropical conditions, but there is considerable variation in their thermal characteristics. *Codium cylindricum* experiences temperatures ranging from 17 to 28°C, nearly encompassing the tolerances of the other four species. All species occupy significantly different thermal niches based on multiple lines of evidence. We observed a common trend of environmental divergence among sister species. This was reflected in the MANOVA and PCA results that showed clear separation of sister species in ecological space. For all species comparisons, except for those with *C. cylindricum*, the first principal axis (temperature) was responsible for this separation. PCAs involving *C. cylindricum* showed that species were very similar along this first axis, while there were more differences along the second



axis (seasonality). This species occurs in both Mexico and Japan, temperate regions experiencing high temperature ranges. Here, one could argue that *C. cylindricum* is biased towards a certain season. Large differences in seasonal temperatures cause some organisms to be limited by winter cold and other by summer heat (Adey & Steneck, 2001). However, by comparing sampling dates and monthly temperature values, we confirm the presence of this species on the Japanese and northern Mexican coasts all year around and thus the species appears to be able to persist across a wide temperature range. This effect is also reflected in the thermal niche model, which is dominated by seasonality (the second PCA axis) and hence shows only limited amounts of suitable habitat at a global scale compared to other species in the complex (Figure 2, present-day). Although this model represents the correct distribution of the species, we cannot rule out that *C. cylindricum* has the potential to survive in areas with a smaller seasonal temperature range.

By incorporating ecological niche models, climatic and geographic distributions could be estimated beyond the sample locations. Overlap tests based on these models showed that not a single species' niche was found to be equivalent to another, regardless of the measure of similarity used (Table 1). The highest values of niche equivalency were presented by comparing *C. decorticans* to the other species. These high values are due to the fact that *C. decorticans* is characterized by a broad environmental niche that overlaps with all other species. The lowest niche overlap was observed between *C. duthieae* 2 from Australia and *C. duthieae* 3 from the Arabian Sea. These values are corroborated by the results of the PCA (Figure 4), which separates both species completely on the first axis. The potential niches of both species are also virtually mutually exclusive in geographical space (Figure 2). When comparing equivalency of niches, however, it is important to take scale effects into account (Pearman *et al.*, 2008). Here, we must consider scale effect along a thermal axis. One can imagine species being close to one another with respect to the entire range of temperatures observed on earth, but different from each other within the "temperate" subset of temperatures. Thus, despite the significant differences observed, we can consider that all species have retained their affinities for warm-temperate environmental conditions.

Thermal niche models for *Codium duthieae* 1, *C. duthieae* 2 and *C. decorticans* predict good suitability along the South American coast. Interestingly, Setchell (1937) described "*Codium fernandezianum*" from the Juan Fernandez Islands, an archipelago off the coast of Chile, stating 'It may be questioned whether the Juan Fernandez plants should all be placed under the same species and it may also be questioned, in case it may be decided that they are all only variants of one species, whether to segregate them from *Codium decorticans*.' Setchell's uncertainty as to exactly how to call the Chilean entity is not surprising given the morphological similarity between species in the *C. decorticans* complex, but it does indicate that a species belonging to the complex may be present in the Juan Fernandez Islands. Whether it is a proper species (Setchell decided to call it *C. fernandezianum*) or that it consists of a population of *Codium duthieae* 2 that migrated along the Antarctic circumpolar current, or whether yet another explanation is needed, remains to be determined.

## CONCLUSION

Using a combination of a calibrated phylogeny, ecological niche models and projections of those models into the past, we have shed light on the biogeography and macroevolutionary pattern of a warm-temperate complex of seaweed species. Diversification of the *Codium decorticatum* complex could be placed in the late Miocene – early Pliocene. This timeframe in combination with the thermal and geographic distribution of the complex favor a dispersal hypothesis. Because niche projections give evidence for a tropical barrier impeding regular connection between current geographic ranges, we hypothesize that rare transoceanic dispersal events are likely to be responsible for this pattern. Low frequency dispersal would allow for genetic differentiation and founder speciation as species adapted to their local thermal environment.

## *ACKNOWLEDGEMENTS*

We thank Matthias Leroy and Katrien Van Nieuwenhuyze for their exploratory work on this research topic and Satoshi Shimada, Taiju Kitayama, Line Le Gall, Barrett Brooks, Vaibhav Mantri and Ingrid Lin for providing specimen records. Sofie D'hondt is gratefully acknowledged for her persisting laboratory work. LT was funded by the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT, 83721). HV is a Future Fellow of the Australian Research Council (project FT110100585). Additional financial support was provided by the King Leopold III Fund for Nature Exploration and Conservation.

## *AUTHOR CONTRIBUTIONS*

LT and HV conceived and designed the study. LT, HV and ODC constructed the phylogenetic alignments. HV carried out the phylogenetic analyses. LT prepared the species and environmental data sets, and performed all niche related analyses. LT wrote the manuscript and HV and ODC gave comments.



# REFERENCES

- Adey, W. H. & Steneck, R. S. (2001) Thermogeography over time creates biogeographic regions: a temperature/space/time-integrated model and an abundance-weighted test for benthic marine algae. *Journal of Phycology*, **37**, 677-698.
- Bellwood, D. R., Leis, J. M. & Stobutzki, I. C. (1998) Fishery and Reef Management. *Science*, **279**, 2019-2025.
- Bolton, J. (2010) The biogeography of kelps (Laminariales, Phaeophyceae): a global analysis with new insights from recent advances in molecular phylogenetics. *Helgoland Marine Research*, **64**, 263-279.
- 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. *Marine Biology*, **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**, 72.
- Cowen, R. K., Lwiza, K. M. M., Sponaugle, S., Paris, C. B. & Olson, D. B. (2000) Connectivity of marine populations: Open or closed? *Science*, **287**, 857-859.
- Crisp, M. D., Treweek, S. A. & Cook, L. G. (2011) Hypothesis testing in biogeography. *Trends in Ecology & Evolution*, **26**, 66-72.
- Darwin, C. M. A. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*, edn., London.
- De Clerck, O., Verbruggen, H., Huisman, J. M., Faye, E. J., Leliaert, F., Schils, T. & Coppejans, E. (2008) Systematics and biogeography of the genus *Pseudocodium* (Bryopsidales, Chlorophyta), including the description of *P. natalense* sp. nov. from South Africa. *Phycologia*, **47**, 225-235.
- Dowsett, H., Robinson, M., Haywood, A., Salzmann, U., Hill, D., Sohl, L., Chandler, M., Williams, M., Foley, K. & Stoll, D. (2010) The PRISM3D paleoenvironmental reconstruction. *Stratigraphy*, **7**, 123-139.
- Dowsett, H. J., Robinson, M. M. & Foley, K. M. (2009) Pliocene three-dimensional global ocean temperature reconstruction. *Climate of the Past*, **5**, 769-783.
- Dromgoole, F. I. (1982) The buoyant properties of *Codium*. *Botanica Marina*, **25**, 391-397.
- Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*.
- Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M., Peterson, A. T., Phillips, S. J., Richardson, K., Scachetti-Pereira, R., Schapire, R. E., Soberon, J., Williams, S., Wisz, M. S. & Zimmermann, N. E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151.
- Evans, M. E. K., Smith, S. A., Flynn, R. S. & Donoghue, M. J. (2009) Climate, Niche Evolution, and Diversification of the "Bird-Cage" Evening Primroses (*Oenothera*, Sections *Anogra* and *Kleinia*). *American Naturalist*, **173**, 225-240.
- Fralick, R. A. & Mathieson, A. C. (1972) Winter fragmentation of *Codium fragile* (Suringar) Hariott ssp. *tomentosoides* (van Goor) Silva (Chlorophyceae, Siphonales) in New England. *Phycologia*, **11**, 67-70.
- Fraser, C. I., Nikula, R., Spencer, H. G. & Waters, J. M. (2009) Kelp genes reveal effects of subantarctic sea ice during the Last Glacial Maximum. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 3249-3253.
- Gillespie, R. G., Baldwin, B. G., Waters, J. M., Fraser, C. I., Nikula, R. & Roderick, G. K. (2012) Long-distance dispersal: a framework for hypothesis testing. *Trends in Ecology & Evolution*, **27**, 47-55.
- Graham, C. H., Ron, S. R., Santos, J. C., Schneider, C. J. & Moritz, C. (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, **58**, 1781-1793.
- Guindon, S. & Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696-704.
- Holmes, E. M. (1896) New marine algae from Japan. *Journal of the Linnean Society of London, Botany*, **31**, 248-260.
- Hommersand, M. (2007) Global biogeography and relationships of the Australian marine macroalgae. *Algae of Australia: Introduction* (ed. by P.M. McCarthy & A.E. Orchard), pp 511-542. CSIRO Publishing, Melbourne.
- Hommersand, M. H. (1986) The biogeography of the South African marine red algae: a model. *Botanica Marina*, **24**, 257-270.
- Hommersand, M. H. & Fredericq, S. (2003) Biogeography of the marine red algae of the South African West Coast: a molecular approach. *17th international Seaweed Symposium*, pp 325-336.
- Howe, M. A. (1911) Phycological studies - V. Some marine algae of lower California, Mexico. *Bulletin of the Torrey Botanical Club*, **38**, 489-514.
- Hubbs, C. L. (1952) Antitropical distribution of fishes and other organisms. Symposium on the problems of bipolarity and of pantemperate faunas. *Proceedings of the Seventh Pacific Science Congress (Pacific Science Association)*, **3**, 324-329.

- Jobb, G., Von Haeseler, A. & Strimmer, K. (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evolutionary Biology*, **4**, 18.
- Joosten, A. M. T. & Van Den Hoek, C. (1986) World-wide relationships between red seaweed floras: a multivariate approach. *Botanica Marina*, **29**, 195-214.
- Kozak, K. H., Graham, C. H. & Wiens, J. J. (2008) Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution*, **23**, 141-148.
- Kozak, K. H. & Wiens, J. J. (2006) Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution*, **60**, 2604-2621.
- Lindberg, D. R. (1991) Marine biotic interchange between the northern and southern hemispheres. *Paleobiology*, **17**, 308-324.
- Lüning, K. (1990) *Seaweeds: their environment, biogeography, and ecophysiology*, edn. Wiley & Sons, New York.
- Macaya, E. C. (2010) Phylogeny, connectivity and dispersal patterns of the giant kelp *Macrocystis* (Phaeophyceae). Victoria University of Wellington, Wellington.
- Morlon, H., Parsons, T. L. & Plotkin, J. B. (2011) Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences*.
- Nikula, R., Fraser, C. I., Spencer, H. G. & Waters, J. M. (2010) Circumpolar dispersal by rafting in two subantarctic kelp-dwelling crustaceans. *Marine Ecology Progress Series*, **405**, 221-230.
- Nizamuddin, M. (2001) Genus *Codium* Stackhouse from northern coast of the Arabian Sea. *Pakistan Journal Of Marine Biology*, **7**, 147-232.
- Norris, J. N. (2010) Marine Algae of the Northern Gulf of California: Chlorophyta and Phaeophyceae. *Smithsonian contributions to botany*, **94**, 1-276.
- Norris, R. E. & Aken, M. E. (1985) Marine benthic algae new to South Africa. *Suid-Afrikaanse Tydskrif vir Plantkunde*, **51**, 55-56.
- Palumbi, S. R. (1994) Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, **25**, 547-572.
- Paul, A. & Schäfer-Neth, C. (2003) Modeling the water masses of the Atlantic Ocean at the Last Glacial Maximum. *Paleoceanography*, **18**, 1058.
- Pearman, P. B., Guisan, A., Broennimann, O. & Randin, C. F. (2008) Niche dynamics in space and time. *Trends in Ecology & Evolution*, **23**, 149-158.
- Phillips, J. A. (2001) Marine macroalgal biodiversity hotspots: why is there high species richness and endemism in southern Australian marine benthic flora? *Biodiversity and Conservation*, **10**, 1555-1577.
- Phillips, S. J., Anderson, R. P. & Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.
- Phillips, S. J. & Dudik, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161-175.
- Pillmann, A., Woolcott, G. W., Olsen, J. L., Stam, W. T. & King, R. J. (1997) Inter- and intraspecific genetic variation in *Caulerpa* (Chlorophyta) based on nuclear rDNA ITS sequences. *European Journal of Phycology*, **32**, 379-386.
- Qian, H. & Ricklefs, R. E. (2004) Geographical distribution and ecological conservatism of disjunct genera of vascular plants in eastern Asia and eastern North America. *Journal of Ecology*, **92**, 253-265.
- R Development Core Team. (2008) R: A language and environment for statistical computing.
- Rabosky, D. L. (2009) Heritability of Extinction Rates Links Diversification Patterns in Molecular Phylogenies and Fossils. *Systematic Biology*, **58**, 629-640.
- Rabosky, D. L. (2010) Extinction rates should not be estimated from molecular phylogenies. *Evolution*, **64**, 1816-1824.
- Rambaut, A. & Drummond, A. J. (2007) Tracer.
- Ramus, J. (1972) Differentiation of green alga *Codium fragile*. *American Journal of Botany*, **59**, 478-482.
- Ravelo, A. C., Andreasen, D. H., Lyle, M., Olivarez Lyle, A. & Wara, M. W. (2004) Regional climate shifts caused by gradual global cooling in the Pliocene epoch. *Nature*, **429**, 263-267.
- Revell, L. J. (2011) phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, no-no.
- Ronquist, F. & Huelsenbeck, J. P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.
- Schaffelke, B. & Deane, D. (2005) Desiccation tolerance of the introduced marine green alga *Codium fragile* ssp *tomentosoides* - clues for likely transport vectors? *Biological Invasions*, **7**, 557-565.
- Schils, T. & Coppejans, E. (2003) Phytogeography of upwelling areas in the Arabian Sea. *Journal of Biogeography*, **30**, 1339-1356.
- Schils, T. & Wilson, S. C. (2006) Temperature threshold as a biogeographic barrier in northern Indian Ocean macroalgae. *Journal of Phycology*, **42**, 749-756.
- Setchell, W. A. (1937) The *Codiums* of the Juan Fernández Island. *The natural history of Juan Fernández and Eastern Island*, pp 587-600, Uppsala.
- Setchell, W. A. & Gardner, N. L. (1924) New marine algae from the Gulf of California. *Proceeding of the California Academy of Science*, **12**, 695-949.
- Silva, P. C. (1959) The genus *Codium* (Chlorophyta) in South Africa. *Journal of South African Botany*, **25**, 101-165.



- Stadler, T. (2011) Inferring speciation and extinction processes from extant species data. *Proceedings of the National Academy of Sciences*, **108**, 16145-16146.
- Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688-2690.
- Thiel, M. & Haye, P. A. (2006) The ecology of rafting in the marine environment. III. Biogeographical and Evolutionary Consequences. *Oceanography and Marine Biology: An Annual Review*. (ed. by R.N. Gibson & R.J.A. Atkinson & J.D.M. Gordon), pp 323-429. Taylor & Francis.
- Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F. & De Clerck, O. (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, **21**, 272-281.
- Veizer, J., Ala, D., Azmy, K., Bruckschen, P., Buhl, D., Bruhn, F., Carden, G. A. F., Diener, A., Ebner, S., Godderis, Y., Jasper, T., Korte, C., Pawellek, F., Podlaha, O. G. & Strauss, H. (1999) Sr-87/Sr-86, delta C-13 and delta O-18 evolution of Phanerozoic seawater. *Chemical Geology*, **161**, 59-88.
- Verbruggen, H., Ashworth, M., Loduca, S. T., Vlaeminck, C., Cocquyt, E., Sauvage, T., Zechman, F. W., Littler, D. S., Littler, M. M., Leliaert, F. & De Clerck, O. (2009a) A multi-locus time-calibrated phylogeny of the siphonous green algae. *Molecular Phylogenetics and Evolution*, **50**, 642-653.
- Verbruggen, H., De Clerck, O., Schils, T., Kooistra, W. & Coppejans, E. (2005) Evolution and phylogeography of *Halimeda* section *Halimeda*. *Molecular Phylogenetics and Evolution*, **37**, 789-803.
- Verbruggen, H., Leliaert, F., Maggs, C. A., Shimada, S., Schils, T., Provan, J., Booth, D., Murphy, S., De Clerck, O., Littler, D. S., Littler, M. M. & Coppejans, E. (2007) Species boundaries and phylogenetic relationships within the green algal genus *Codium* (Bryopsidales) based on plastid DNA sequences. *Molecular Phylogenetics and Evolution*, **44**, 240-254.
- Verbruggen, H., Tyberghein, L., Pauly, K., Vlaeminck, C., Van Nieuwenhuyze, K., Kooistra, W., Leliaert, F. & De Clerck, O. (2009b) Macroecology meets macroevolution: evolutionary niche dynamics in the seaweed *Halimeda*. *Global Ecology and Biogeography*, **18**, 393-405.
- Warren, D. L., Glor, R. E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868-2883.
- Warren, D. L., Glor, R. E. & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607-611.
- Waters, J. M. (2008) Driven by the West Wind Drift? A synthesis of southern temperate marine biogeography, with new directions for dispersalism. *Journal of Biogeography*, **35**, 417-427.
- Waters, J. M. & Roy, M. S. (2004) Out of Africa: The Slow Train to Australasia. *Systematic Biology*, **53**, 18-24.
- Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H. & Guisan, A. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763-773.
- Wynne, M. J. (2000) Further connections between the benthic marine algal floras of the northern Arabian Sea and Japan. *Phycological Research*, **48**, 211-220.
- Wynne, M. J. (2001) New records of benthic marine algae from the Sultanate of Oman, northern Arabian sea. III. *Contr. Univ. Mich. Herb.*, **23**, 389-406.
- Yesson, C. & Culham, A. (2006a) Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. *Systematic Biology*, **55**, 785-802.
- Yesson, C. & Culham, A. (2006b) A phyloclimatic study of *Cyclamen*. *BMC Evolutionary Biology*, **6**, 72.
- Zahn, R. (2009) Climate change: Beyond the CO2 connection. *Nature*, **460**, 335-336.
- Zemlak, T. S., Ward, R. D., Connell, A. D., Holmes, B. H. & Hebert, P. D. N. (2009) DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources*, **9**, 237-242.



# APPENDIX

**Table S1: Data availability matrix. Graphical representation of the concatenated alignment, showing the availability of sequence data.**

|                          | atpA + intron                                    | RbcL exon1  | RbcL exon2                                       | tufA   | UCP6   | UCP7  |
|--------------------------|--|---|--|--|--|---|
| <i>C. adhaerens</i> 1    |  | GenBank: XX<br>strain: Dorset<br>length: 760 nt         |  | GenBank: XX<br>strain: HEC15691<br>length: 905 nt    | GenBank: EF107854<br>strain: SMG05.35<br>length: 357 nt  |   |
| <i>C. arabicum</i> 1     |  | GenBank: EF107969<br>strain: JH9<br>length: 775 nt      |  | GenBank: XX<br>strain: SGAD0509370<br>length: 911 nt | GenBank: EF107856<br>strain: SD0509370<br>length: 357 nt |   |
| <i>C. bursa</i>          |  | GenBank: EF107970<br>strain: KRK001<br>length: 775 nt   |  | GenBank: XX<br>strain: G.113<br>length: 873 nt       | GenBank: XX<br>strain: G.113<br>length: 358 nt           |   |
| <i>C. cf. dimorphum</i>  |  | GenBank: XX<br>strain: HV1877<br>length: 722 nt         |  | GenBank: XX<br>strain: HV1867<br>length: 911 nt      | GenBank: EF107875<br>strain: C29<br>length: 267 nt       |   |
| <i>C. coralloides</i>    |  | GenBank: EF107977<br>strain: KRK003<br>length: 772 nt   |  | GenBank: XX<br>strain: HV03269<br>length: 750 nt     | GenBank: EF107869<br>strain: KRK003<br>length: 356 nt    |   |
| <i>C. cylindricum</i>    | GenBank: XX<br>strain: HV1730<br>length: 872 nt  | GenBank: XX<br>strain: HV1730<br>length: 750 nt         | GenBank: XX<br>strain: HV1730<br>length: 515 nt  | GenBank: new<br>strain: HV1730<br>length: 890 nt     | GenBank: EF107872<br>strain: C223<br>length: 268 nt      | GenBank: XX<br>strain: HV1730<br>length: 331 nt   |
| <i>C. decorticatum</i>   | GenBank: XX<br>strain: H.0775<br>length: 925 nt  | GenBank: XX<br>strain: H.0775<br>length: 741 nt         | GenBank: XX<br>strain: H.0775<br>length: 516 nt  | GenBank: XX<br>strain: H.0775<br>length: 1026 nt     | GenBank: EF107873<br>strain: CDNC07<br>length: 357 nt    | GenBank: XX<br>strain: H.0775<br>length: 561 nt   |
| <i>C. dimorphum</i>      |  | GenBank: EF107981<br>strain: H.0771<br>length: 760 nt   |  | GenBank: XX<br>strain: H.0771<br>length: 682 nt      | GenBank: EF107874<br>strain: H.0771<br>length: 344 nt    |   |
| <i>C. duthieae</i> 1     | GenBank: XX<br>strain: G.300<br>length: 925 nt   | GenBank: XX<br>strain: G.300<br>length: 750 nt          | GenBank: XX<br>strain: G.300<br>length: 515 nt   | GenBank: XX<br>strain: HEC15348<br>length: 928 nt    | GenBank: EF107877<br>strain: KZN2K4.1<br>length: 357 nt  | GenBank: XX<br>strain: HEC15348<br>length: 563 nt |
| <i>C. duthieae</i> 2     | GenBank: XX<br>strain: G.291<br>length: 925 nt   | GenBank: XX<br>strain: G.291<br>length: 750 nt          | GenBank: XX<br>strain: LT105<br>length: 515 nt   | GenBank: XX<br>strain: G.291<br>length: 920 nt       | GenBank: EF107878<br>strain: H0691<br>length: 381 nt     | GenBank: XX<br>strain: G.291<br>length: 667 nt    |
| <i>C. duthieae</i> 3     | GenBank: XX<br>strain: DHO.003<br>length: 881 nt | GenBank: EF107994<br>strain: DHO.006<br>length: 750 nt  | GenBank: XX<br>strain: DHO.003<br>length: 516 nt | GenBank: XX<br>strain: DHO.006<br>length: 928 nt     | GenBank: EF107879<br>strain: DHO2.002<br>length: 358 nt  | GenBank: XX<br>strain: DHO.006<br>length: 621 nt  |
| <i>C. fragile</i>        | GenBank: XX<br>strain: JAP128<br>length: 932 nt  | GenBank: XX<br>strain: JAP128<br>length: 721 nt         | GenBank: XX<br>strain: JAP128<br>length: 515 nt  | GenBank: XX<br>strain: JAP128<br>length: 928 nt      | GenBank: EF107889<br>strain: C41<br>length: 317 nt       | GenBank: XX<br>strain: JAP128<br>length: 560 nt   |
| <i>C. goeleatum</i>      | GenBank: XX<br>strain: G.111<br>length: 853 nt   | GenBank: XX<br>strain: G.111<br>length: 750 nt          | GenBank: XX<br>strain: G.111<br>length: 515 nt   | GenBank: XX<br>strain: G.111<br>length: 920 nt       | GenBank: EF107892<br>strain: H0794<br>length: 380 nt     | GenBank: XX<br>strain: G.111<br>length: 495 nt    |
| <i>C. isthmocladum</i> 1 |  | GenBank: XX<br>strain: PEUFR48498<br>length: 787 nt     |  | GenBank: XX<br>strain: HV907<br>length: 937 nt       | GenBank: EF107905<br>strain: HV907<br>length: 357 nt     |   |
| <i>C. megalophysum</i>   |  | GenBank: EF108058<br>strain: HEC15349<br>length: 763 nt |  | GenBank: XX<br>strain: G.919<br>length: 801 nt       | GenBank: EF107917<br>strain: KZN2K4.29<br>length: 358 nt |   |
| <i>C. minus</i> 1        |  | GenBank: AB102959<br>strain: C43<br>length: 676 nt      |  | GenBank: XX<br>strain: JAP077<br>length: 911 nt      | GenBank: EF107918<br>strain: C43<br>length: 271 nt       |   |
| <i>C. saccatum</i>       |  | GenBank: XX<br>strain: H.0974<br>length: 741 nt         |  | GenBank: XX<br>strain: HA0096<br>length: 905 nt      | GenBank: XX<br>strain: H.0974<br>length: 358 nt          |   |
| <i>C. sp6</i>            | GenBank: XX<br>strain: LL0470<br>length: 892 nt  | GenBank: XX<br>strain: LL0470<br>length: 725 nt         | GenBank: XX<br>strain: LL0470<br>length: 515 nt  | GenBank: XX<br>strain: LL0470<br>length: 768 nt      | GenBank: EF107951<br>strain: LL0470<br>length: 358 nt    | GenBank: XX<br>strain: LL0470<br>length: 562 nt   |
| <i>C. spinescens</i>     |  | GenBank: EF108075<br>strain: H.0693<br>length: 765 nt   |  | GenBank: XX<br>strain: G.112<br>length: 846 nt       | GenBank: EF107931<br>strain: H.0693<br>length: 356 nt    |   |
| <i>C. tomentosum</i>     | GenBank: XX<br>strain: HV1344<br>length: 863 nt  | GenBank: XX<br>strain: HV1344<br>length: 750 nt         | GenBank: XX<br>strain: HV1344<br>length: 515 nt  | GenBank: XX<br>strain: HV1344<br>length: 982 nt      | GenBank: XX<br>strain: CTOV01<br>length: 268 nt          | GenBank: XX<br>strain: HV1344<br>length: 585 nt   |
| <i>C. yezoense</i>       |  | GenBank: XX<br>strain: HV1855<br>length: 674 nt         |  | GenBank: XX<br>strain: HV1855<br>length: 905 nt      | GenBank: EF107945<br>strain: C53<br>length: 308 nt       |   |

Table S2: Species Occurrence Data: References

| Species                    | Source  | # Specimens | # Unique localities |
|----------------------------|---|-------------|---------------------|
| <i>Codium cylindricum</i>  | GENT Herbarium, SAP Herbarium, TNS Herbarium                    | 133         | 119                 |
| <i>Codium decorticatum</i> | GENT Herbarium, PARIS Herbarium, Smithsonian Institution        | 45          | 27                  |
| <i>Codium duthieae</i> 1   | GENT Herbarium; (Silva, 1959)                                   | 26          | 24                  |
| <i>Codium duthieae</i> 2   | GENT Herbarium; Adelaide Herbarium; Herbarium Western Australia | 69          | 40                  |
| <i>Codium duthieae</i> 3   | GENT Herbarium; (Wynne, 2001); (Nizamuddin, 2001)               | 34          | 24                  |

Table S3: Covariance matrix and principal component analysis on 4 Bio-ORACLE layers (sstmean, sstrange, sstmax, sstmin). Analyses conducted in ArcGIS 10 (ESRI).

| Covariance matrix |          |         |          |         |
|-------------------|----------|---------|----------|---------|
|                   | range    | sstmean | sstmin   | sstmax  |
| sstrange          | 1        | 0.01255 | -0.10419 | 0.17146 |
| sstmean           | 0.01255  | 1       | 0.99289  | 0.98699 |
| sstmin            | -0.10419 | 0.99289 | 1        | 0.96196 |
| sstmax            | 0.17146  | 0.98699 | 0.96196  | 1       |

| PCA      |         |          |          |          |
|----------|---------|----------|----------|----------|
|          | PC1     | PC2      | PC3      | PC4      |
| sstrange | 0.00453 | 0.81532  | -0.04316 | -0.57738 |
| sstmean  | 0.5814  | -0.04661 | -0.81228 | -0.00054 |
| sstmin   | 0.57304 | -0.38893 | 0.43286  | -0.57706 |
| sstmax   | 0.57757 | 0.4264   | 0.38855  | 0.57761  |

Table S4: Principal component analyses on extracted *Codium* occurrence point data from Bio-ORACLE.

|  |          | PC1    | PC2    | PC3    | PC4    |
|--|----------|--------|--------|--------|--------|
| All five species                             | sstmax   | -0.524 | 0.459  | 0.435  | 0.571  |
|  | sstmean  | -0.635 | 0.006  | -0.772 | 0.000  |
|  | sstmin   | -0.566 | -0.371 | 0.462  | -0.573 |
|  | sstrange | 0.043  | 0.807  | -0.029 | -0.588 |
| <i>C. duthieae</i> 1 vs <i>C. cylindrium</i> | sstmax   | -0.240 | 0.746  | 0.454  | -0.424 |
|  | sstmean  | -0.601 | 0.266  | -0.754 | 0.000  |
|  | sstmin   | -0.630 | -0.131 | 0.456  | 0.615  |
|  | sstrange | 0.430  | 0.596  | -0.132 | 0.665  |
| <i>C. duthieae</i> 1 vs <i>C. duthieae</i> 2 | sstmax   | -0.540 | 0.351  | 0.431  | 0.632  |
|  | sstmean  | -0.572 | 0.117  | -0.812 | 0.000  |
|  | sstmin   | -0.574 | -0.079 | 0.393  | -0.714 |
|  | sstrange | 0.228  | 0.926  | -0.027 | -0.300 |

|   |          |        |        |        |        |
|---|----------|--------|--------|--------|--------|
| <i>C. duthieae</i> 1 vs <i>C. duthieae</i> 3    | sstmax   | -0.564 | 0.029  | 0.380  | 0.733  |
|   | sstmean  | -0.559 | -0.100 | -0.823 | 0.000  |
|   | sstmin   | -0.493 | -0.528 | 0.398  | -0.565 |
|   | sstrange | -0.356 | 0.843  | 0.139  | -0.379 |
| <i>C. duthieae</i> 1 vs <i>C. decorticatum</i>  | sstmax   | 0.538  | 0.406  | 0.440  | 0.594  |
|   | sstmean  | 0.601  | 0.067  | -0.796 | 0.000  |
|   | sstmin   | 0.578  | -0.257 | 0.414  | -0.655 |
|   | sstrange | -0.126 | 0.875  | -0.021 | -0.468 |
| <i>C. duthieae</i> 2 vs <i>C. cylindricum</i>   | sstmax   | 0.583  | 0.307  | 0.418  | 0.625  |
|   | sstmean  | 0.620  | -0.125 | -0.774 | 0.000  |
|   | sstmin   | 0.487  | -0.521 | 0.474  | -0.516 |
|   | sstrange | 0.193  | 0.786  | 0.028  | -0.586 |
| <i>C. duthieae</i> 2 vs <i>C. duthieae</i> 3    | sstmax   | -0.576 | 0.157  | 0.413  | 0.688  |
|   | sstmean  | -0.583 | -0.006 | -0.812 | 0.000  |
|   | sstmin   | -0.572 | -0.196 | 0.412  | -0.682 |
|   | sstrange | -0.026 | 0.968  | 0.012  | -0.249 |
| <i>C. duthieae</i> 2 vs <i>C. decorticatum</i>  | sstmax   | -0.557 | 0.340  | 0.428  | 0.625  |
|   | sstmean  | -0.600 | 0.025  | -0.800 | 0.000  |
|   | sstmin   | -0.572 | -0.276 | 0.421  | -0.648 |
|   | sstrange | 0.052  | 0.899  | -0.011 | -0.436 |
| <i>C. duthieae</i> 3 vs <i>C. cylindricum</i>   | sstmax   | -0.331 | 0.764  | 0.411  | 0.370  |
|   | sstmean  | -0.584 | 0.174  | -0.793 | 0.000  |
|   | sstmin   | -0.589 | -0.140 | 0.403  | -0.686 |
|   | sstrange | 0.450  | 0.605  | -0.198 | -0.626 |
| <i>C. duthieae</i> 3 vs <i>C. decorticatum</i>  | sstmax   | 0.563  | 0.308  | 0.433  | 0.634  |
|   | sstmean  | 0.596  | 0.040  | -0.802 | 0.000  |
|   | sstmin   | 0.571  | -0.268 | 0.411  | -0.658 |
|   | sstrange | -0.048 | 0.912  | 0.009  | -0.407 |
| <i>C. cylindricum</i> vs <i>C. decorticatum</i> | sstmax   | -0.412 | 0.610  | 0.454  | 0.502  |
|   | sstmean  | -0.632 | 0.139  | -0.762 | 0.000  |
|   | sstmin   | -0.608 | -0.261 | 0.457  | -0.595 |
|   | sstrange | 0.246  | 0.735  | -0.070 | -0.628 |



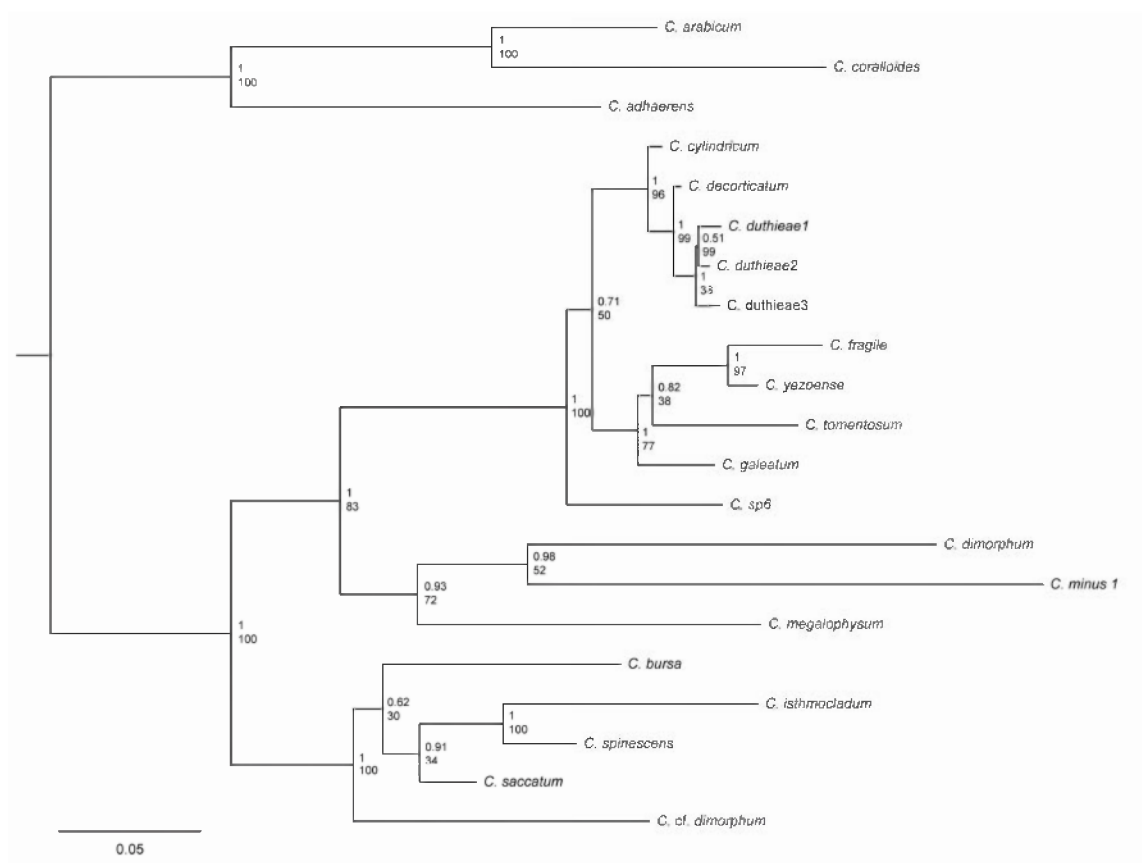


Figure S1: Bayesian phylogenetic tree of 21 *Codium* species. Numbers at nodes indicate posterior probabilities (top) and ML bootstrap values (bottom).

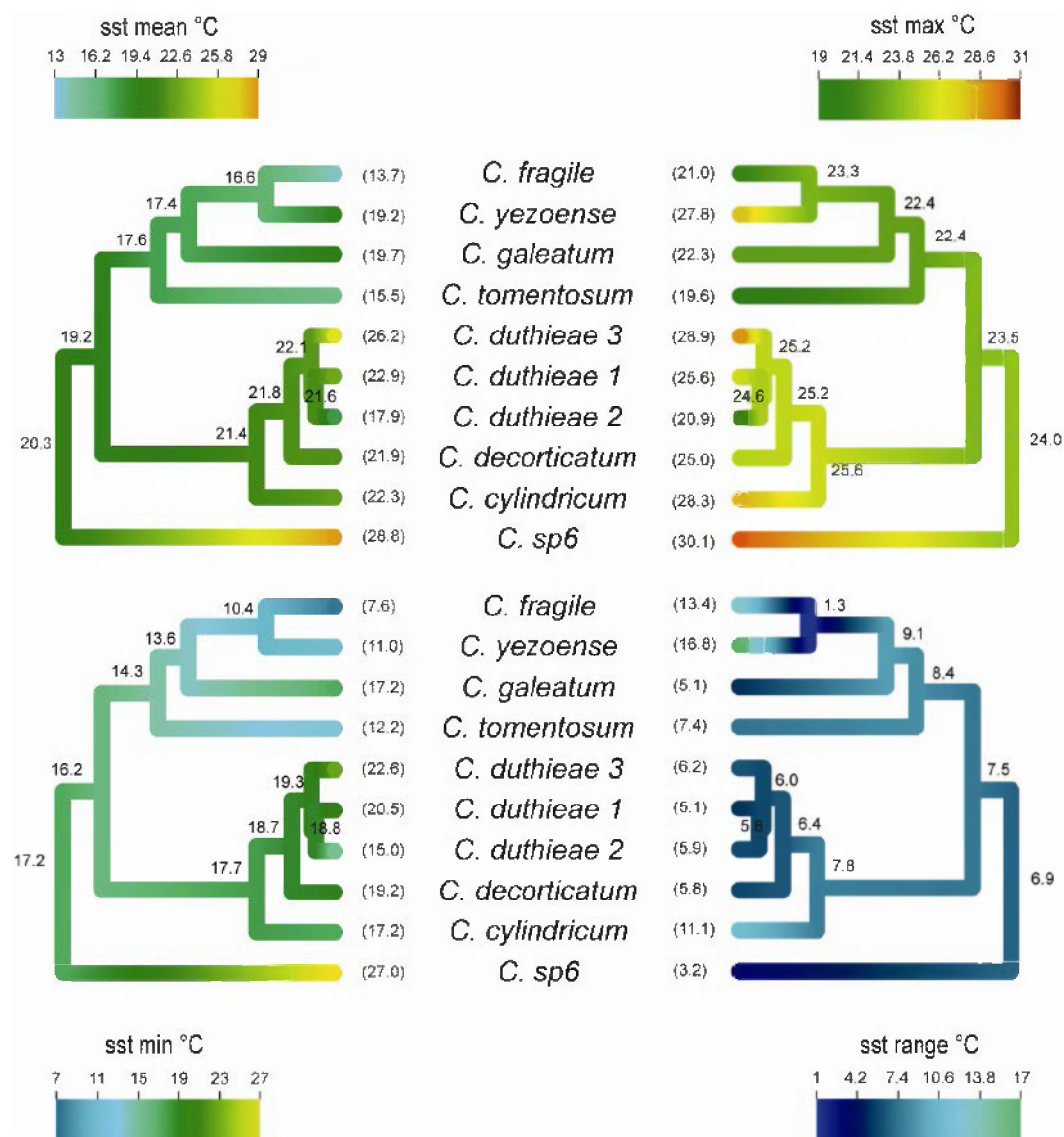


Figure S2: Inferred evolutionary history of niche features in 10 *Codium* species. Ancestral values for sea surface temperature (sstmean, sstmin, sstmax and ssstrange) are plotted along the phylogeny. Numbers plotted at nodes indicate the inferred ancestral values. These values were obtained using a Bayesian MCMC approach (Revell, 2011). The color gradient indicates the temperatures.