BIOGEOGRAPHY AND MACROECOLOGY OF BENTHIC MARINE ALGAE

Thesis submitted by

Ailsa Palmer KERSWELL BA, BSc(Hons) (*U of Qld)*

in December 2006

for the degree of Doctor of Philosophy

in the School of Marine and Tropical Biology

James Cook University

STATEMENT OF ACCESS

I, the undersigned author of this work, understand that James Cook University will make this thesis available for use within the University Library and, via the Australian Digital Theses network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and;

I do not wish to place any further restriction on access to this work

_____ _____

Signature Date

STATEMENT ON SOURCES

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

…….. ……

(Signature) (Date)

STATEMENT OF THE CONTRIBUTION OF OTHERS

Data chapters 3-5 of this thesis include collaborative work undertaken with my supervisor Dr Sean Connolly. While undertaking the collaborations, I was responsible for the project concept and design, data collection, performing analyses, synthesis of results and the preparation of manuscripts for submission to peer-reviewed journals. Sula Blake assisted with the initial data entry and verification of algal nomenclature.

Financial support in the form of stipends (including fees) was obtained as an Australian Postgraduate Award and a James Cook University Doctoral Completion Scholarship. Project funds were kindly provided by PADI Aware Asia Pacific, Phycological Society of America, Sigma Xi, and the JCU Graduate Research Scool. Financial assistance for travel to national and international conferences was provided by the School of Marine Biology and Aquaculture (JCU), JCU Graduate Research Scool, and the Ecological Society of Australia.

James Cook University High Performance Computing (HPC) facilities were used to run analyses presented in Chapter 5 of this thesis. Mizue Hisano provided assistance in accessing this facility.

ACKNOWLEDGEMENTS

Firstly, I would like to extend my most sincere thanks to my supervisors; Sean Connolly, Terry Hughes and Andrew Baird. Thank-you Sean for your unfailing support and interest in this project. I would be a poorer ecologist if it weren't for your guidance in the (sometimes murky) waters of theoretical ecology and the occasional push beyond my comfort zone. Thank-you Terry for advice in the developmental stage of this thesis and the financial support you provided. Thank-you Andrew for your field ecologist perspective, considerable editorial expertise and moral support.

I would also like to extend many thanks to Sula Blake for cheerfully spending hours entering data into the algae database and for chasing up hard to find species lists. I could not have navigated the minefield of GIS without the help of James Maloney and Mia Hoogenboom and I am indebted to Mizue Hisano for her assistance accessing the supercomputer. Many people have kindly read earlier drafts of this thesis (either as chapters or manuscripts), so I thank Matthew Kosnik, Phil Munday, Loic Thibaut, Will Robbins, Niall Connolly and Mark Carr for their extensive and insightful comments.

Thanks to the Theoretical Ecology Lab Group for our lively weekly discussions and ideas about how to analyse my algae data.

I was supported by an Australian Postgraduate Award and a JCU Doctoral Completion Scholarship during my candidature. I would like to extend my thanks to PADI Aware Asia Pacific, the Phycological Society of America, Sigma Xi, and the JCU Graduate Research Scool for funds supporting this research. Financial assistance for travel to national and international

conferences was gratefully received from the School of Marine Biology and Aquaculture (JCU), JCU Graduate Research Scool, and the Ecological Society of Australia.

Lastly, I would like to thank my great support team. In particular Matt, Abbs, Maria D and Jane for your wonderful friendship over the last several years. I can think of no better group of people to go through the ups and downs of a PhD with. Thanks also to my family for your unwavering love and support and for being way more interested in algae than you ever really needed to be. And finally to Will – a huge thank-you for your constant love, support and encouragement and, of course, all those emergency coffees!!

ABSTRACT

Macroecologists strive to understand the distribution and abundance of species over wide spatial scales, long time periods and broad taxonomic categories. The major objective of this thesis was to explored two facets of benthic marine algal diversity, taxonomic richness and community structure, in order to enhance our understanding of the processes that underpin patterns of biodiversity in the marine realm. To do so, I assembled a global database of algal distribution records from the primary literature.

I identified global latitudinal and longitudinal diversity gradients for all genera of benthic marine macroalgae and for species in the Order Bryopsidales. I also quantified the size, location, and overlap of macroalgal geographic ranges to determine how the observed richness patterns are generated. Algal genera exhibit an inverse latitudinal gradient, with biodiversity hotspots in temperate regions, while bryopsidalean species reach peak diversity in the tropics. The geographic distribution of range locations results in distinct clusters of range mid-points. In particular, widespread taxa are centred within tight latitudinal and longitudinal bands in the middle of the Indo-Pacific and Atlantic Oceans while small-ranged taxa are clustered in peripheral locations. I assessed a suite of hypotheses about the causes of marine diversity gradients by comparing algal richness patterns, in combination with the size and location of algal geographic ranges, to the richness and range locations predicted by these hypotheses. The results implicate habitat areas and ocean currents as the most plausible drivers of global marine algae diversity patterns.

Species richness patterns of macroalgae in the order Bryopsidales are strikingly concordant with those of corals and reef fishes throughout the tropical Indo-Pacific Ocean. In order to understand the processes that create and maintain tropical marine diversity gradients, I used the

Bryopsidales to test the generality of a model recently developed for coral and reef fish biodiversity. Model selection was used to select the energy-related variables which best predicted species richness. These were then included along with reef area and an estimate of the mid-domain effect in spatial regression models of species richness. The results confirm the role of geographic domain boundaries as a major predictor of marine species richness patterns across a variety of taxa. They also indicate that the relative importance of environmental variables may differ with the taxa in question, with temperature and nitrate being key predictors of algal richness compared to reef area for corals and fishes. Moreover, even though the best models differ for algae versus corals and fishes, the richness patterns predicted by each model deviate from the observed patterns in a consistent manner. This suggests that additional factors, not included in any of the models, are also important in shaping species richness for multiple tropical taxa in the Indo-Pacific Ocean. I propose several candidate factors that may fulfil this role.

A long running controversy in community ecology concerns the extent to which species interactions influence the structure of assemblages. I examined assemblage structure in marine macroalgal communities at a variety of spatial scales in order to test for the existence of Wilson's (1989) guild proportionality assembly rule and to identify the geographical scales at which this rule operates. In order to overcome limitations of the traditional guild-by-guild tests of Wilson (1989), I developed a new guild proportionality test, which examines communities in the aggregate. The functional group composition of algal assemblages was determined for 120 local assemblages using the global database of marine macroalgae distribution records. Using a hierarchy of models and the newly developed guild proportionality test, I examined patterns of assemblage structure at scales ranging from regional to global. Observed communities were compared to null models, which assumed that species occurred in assemblages independently of

one another (i.e., “random” assemblage structure). These comparisons revealed highly non-random structure in algal assemblages at all scales. Communities were more similar than predicted under a random assembly model within tropical regions and throughout the tropical biome, indicating the existence of guild proportionality within these scales. In contrast, communities were more heterogeneous than predicted in all temperate areas, within oceans and across the globe. These patterns suggest that species interactions homogenize assemblage structure within the tropics, but extrinsic processes such as regional environment and historical contingency play an important role in shaping how assemblages vary within temperate regions and at very broad spatial scales.

TABLE OF CONTENTS

LIST OF TABLES

Table A1. Summary of algae occurrence database.

LIST OF FIGURES

Chapter 1: General Introduction

One of the oldest avenues of ecological research is the study of geographic patterns of biodiversity. This field tackles questions fundamental to the understanding of nature, such as how many species are present at a location, what their identities are and what processes control these patterns. The data, which have been accumulating since the earliest studies of Linneaus, Darwin, and Wallace, now suggest that the distribution of biodiversity across the globe can be described in terms of a relatively small number of "ecological laws" (Gaston 2000). Yet, while these patterns are relatively well documented, understanding their cause is one of the most significant challenges of current-day ecological research. This challenge is becoming increasingly pressing as the rates of global climate change and land degradation accelerate. For such reasons it is vital that we understand how diversity is currently distributed and what processes regulate and maintain these patterns in order to effectively respond to and manage future ecosystem changes. The overall aim of this thesis is to use benthic marine algae as a model group to test, on a global scale, some of the most common theories of biodiversity and to improve our understanding of the processes that drive patterns in the marine realm.

1.1 SPECIES RICHNESS PATTERNS

There are … more species of bird breeding in forests than in fields … more species of trees in eastern North America than in Europe. There is an even more dramatic difference in the number of species in the tropics than in the temperate … Will the explanation of these facts degenerate into a tedious set of case histories, or is there some common pattern running through them all?

[MacArthur 1972]

Latitudinal gradients in species richness, which peak in the tropics and decrease towards temperate and polar regions, are among the most well know biodiversity patterns and have been recognised since the early 1800s (Hawkins et al. 2003a). In terrestrial systems, they are well described in plant and animal communities at a variety of scales and throughout geological time (Willig et al. 2003). In the marine environment, tropical richness peaks are evident in shallow benthic, open-ocean and deep-sea systems (Rosen 1988, Gaston 2000). Few exceptions to the classic latitudinal gradient have been documented, and these are generally associated with studies conducted over short latitudinal spans. However, for a small number of taxa, including parasitoids and aquatic plants, species richness does not appear to be related to latitude in a consistent way across all regions (Willig et al. 2003).

In marine systems, longitudinal gradients are also highly consistent among taxa. Many shallow-water marine taxa reach the highest richness in the central Indo-Pacific Ocean (Roberts et al. 2002). From this central location, species richness decreases monotonically towards the African and South American coastlines. In the Atlantic Ocean, richness is consistently highest along the western coastlines of the Caribbean basin (Macpherson 2002). In conjunction with strong latitudinal gradients, the observed longitudinal patterns of species richness create distinct "hotspots" of marine diversity in the central Indo-Australian Archipelago and in the Caribbean. These diversity hotspots are concordant for many taxa including, corals, reef fishes, mangroves, seagrasses, lobsters, and marine gastropods (Duke et al. 2002, Hughes et al. 2002, Roberts et al. 2002, Green and Short 2003).

Species richness patterns cannot be considered in isolation from the geographic ranges of taxa, which overlap to produce the observed patterns. Species' geographic ranges can be considered the basic unit of biogeography (Brown et al. 1996). At continental to global scales, most species

richness patterns are identified first by mapping individual species' ranges and then compiling them to estimate species richness across broad geographic regions (Hawkins et al. 2003a). Moreover, species of differing range size commonly make different contributions to richness patterns. In many terrestrial systems, highly range-restricted endemics constitute a large proportion of the taxa in biodiversity hotspots, which has significant implications for conservation practises (Myers et al. 2000). For birds and mammals, small and large ranging taxa display quite different richness patterns, with most of the overall spatial patterning in richness caused by relatively few widespread species (Lennon et al. 2004, Vazquez and Gaston 2004).

1.1.1 Drivers of species richness patterns

One of the largest and on-going challenges of ecological research is to identify the processes that create and maintain species richness patterns. At the most fundamental level, global variation in species richness is a direct result of variation in the rates of speciation and extinction in different regions. Thus, most theories that aim to explain the location of biodiversity hotspots rely on explanations of why one area will promote speciation or reduce extinction relative to another. To date, over 120 hypotheses have been put forward to explain the observed gradients in species richness (Palmer 1994). Some of these apply only to specific taxa in isolated regions and others are flawed due to circular arguments (Rohde 1992). However, explanations associated with energy, geographic area and geometric constrains appear to be broadly applicable and to have substantial potential to explain species richness patterns (Rahbek and Graves 2001).

The species-energy hypothesis is a climatically based hypothesis, according to which energy availability generates and maintains richness gradients. There are several forms of the species-energy hypothesis, which involve speciation rates, productivity and species' physiological

tolerances. Rohde (1992) proposed the speciation rate theory based on the idea that rates of metabolism, and therefore gene mutation and speciation, will be faster in the warmer temperatures of the tropics. A particular strength of this hypothesis is that models of how cellular kinetics and generation time change with temperature can be used to produces explicit predictions about the shape of the species-energy (temperature) relationship (e.g. Allen et al. [2002] model species richness as a linear function of temperature and predict the slope of the regression line to be -9) . The productivity version of the species-energy hypothesis suggests that increased energy inputs in the tropics facilitate a greater resource base for species to utilize and this in turn, allows species to build larger populations, which are less susceptible to extinction (Hawkins et al. 2003a). A third alternative is that the species-energy hypothesis may mediate species richness via organisms' physiological responses to temperature. This version of the theory is based on the concept that high latitude environments have climates that are harsh and outside the optimum conditions of many taxa, as well as having characteristically large intra-annual variability in environmental conditions (Willig et al. 2003). Consequently, few organisms can maintain populations in high latitudes.

There is evidence for all of the above processes operating in natural systems. Trees, amphibians, ants and some aquatic taxa display species richness-temperature relationships that conform to the predictions from the speciation rate version of the species-energy hypothesis, suggesting that elevated temperatures increase the number of species in a region by accelerating the biochemical reactions that determine speciation rates (Allen et al. 2002, Kaspari et al. 2004). On a global scale, productivity, measured in terms of temperature-water dynamics, appears to be very important in determining species richness patterns of terrestrial groups such as angiosperms and birds (Francis and Currie 2003, Hawkins et al. 2003b). However, for both plants and animals,

the relative importance of productivity compared to ambient energy may be reduced at high latitudes (Hawkins et al. 2003a).

The species-area hypothesis has been proposed as a major alternative to the species-energy hypothesis. This theory posits that larger areas are able to support more individuals and populations of a species, thus reducing the likelihood of extinction. Additionally, larger areas are more likely to contain barriers, which promote allopatric speciation, as well as a greater diversity of habitats, which also promote specialisation, adaptation and speciation. Strong species-area relationships have been documented at local to regional scales, and since the Earth's area is greatest in the tropics, this relationship should also hold at continental to global scales (Rosenzweig 1995). However, there is on-going debate about whether the species-area relationship holds at the very broadest geographic scales, and the extent to which it contributes to latitudinal variation in species numbers (Chown and Gaston 2000, Hawkins and Porter 2001).

Evidence for the species-area hypothesis is inconsistent at continental to global scales. Rosenzweig (1995) presents evidence that as the size of biogeographic provinces or continental islands increases, so too does the richness of rainforest angiosperms, birds and arthropods. He then argues that since the species-area relationship holds at all spatial scales, the cause of latitudinal gradients in species richness must be related to area, because the tropics cover more area than any other climatic zone. However, when the globe is partitioned into terrestrial biomes at a finer scale based on vegetation type rather than latitudinal cut-offs, there is no general latitudinal gradient in biome size. Moreover, the species richness of birds within these biomes shows no relationship with area (Hawkins and Porter 2001). In the marine realm, evidence for the species-area hypothesis is also mixed. In the Indo-Pacific Ocean, the area of coral reef habitat is well correlated with species richness of both corals and reef fishes (Bellwood and

Hughes 2001, Bellwood et al. 2005), providing support for the species-area hypothesis. In contrast, no species-area effects could be detected for mollusc species richness patterns on either coastline of the Americas (Roy et al. 1998).

Species richness gradients have traditionally been viewed as a function of environmental variables such as energy and area, as discussed above. However recently, the mid-domain effect (MDE) has been suggested as a biogeographic null model (Colwell and Hurtt 1994, Colwell and Lees 2000). The significance of MDEs to understanding diversity patterns is that species richness gradients, similar to those seen in nature, may be generated without invoking any environmental variables. Instead, species richness gradients arise due to geometric constraints on the size and location of species' geographic ranges within a bounded domain in the absence of environmental variables. Thus, the location of a species range is constrained only by the size of the range relative to that of the domain; while large-ranged species will occupy much of the domain, small-range species may occur anywhere in the domain, with all locations being equally likely. Consequently, the probability of range overlap is highest in the middle of the domain and lowest at the edges. This produces a quasi-parabolic gradient in species richness, which peaks at the centre of the domain (the mid-domain effect), and is highly reminiscent of many of the species richness gradients observed in nature.

Mid-domain effects are highly controversial. There is on-going debate about formulation of MDE models and the extent to which MDEs influence empirical species richness patterns (see Colwell et al. [2004] for a review). Numerous methods have been employed to formulate MDEs and while each produces a quasi-parabolic gradient in species richness, the height and shape of that gradient differs with the model's specific assumptions (e.g. Koleff and Gaston [2001], Connolly [2005]). The influence of MDEs on species richness gradients has been investigated in

numerous systems and in ~75% of studies the correspondence between MDE predictions and empirical patterns was moderate to substantial (Colwell et al. 2004). For example, MDEs have been shown to account for significant portions of the variation in bird, mammal, coral and reef fish species richness (Willig and Lyons 1998, Jetz and Rahbek 2002, Connolly et al. 2003, Mora and Robertson 2005). As with any null models, an additional advantage of examining species richness patterns relative to a MDE is that deviations of the empirical patterns from null model predictions may provide valuable insight into instance in which other determinants play an important role in regulating species richness patterns (Colwell et al. 2004).

1.1.2 Challenges for biodiversity research

The study of biodiversity patterns and their determinants involves two major challenges. Firstly, ecological data across very large scales are difficult to generate and standardized data collection is only rarely conducted over broad spatial scales (Hurlbert and White 2005). Consequently most continental and global patterns of species richness have been generated using range maps (Hawkins et al. 2003a). Accurately constructing geographic ranges presents the challenge of compiling distributional data from disparate sources, which may have inconsistencies in sampling effort or the identification of taxa. However, it is generally reasonable to assume that the errors do not dominate the data, so long as the distribution data are complied in a consistent and comprehensive manner (Pielou 1977, Brown et al. 1996).

The second major challenge associated with biodiversity research is devising informative test of hypotheses about the processes that create and maintain the observed richness patterns. At broad biogeographic scales manipulative experiments are neither feasible nor ethical. Thus, studies of broad-scale richness patterns must use statistical techniques to tease apart complex patterns of covariation among suites of variables. Methodologically, this involves overcoming problems

associated with spatial autocorrelation within, and collinearity among, predictor variables as well as the formulations of appropriate biogeographic null models. A further challenge is teasing apart causative factors from what are essentially correlative studies.

1.2 COMMUNITY STRUCTURE

An alternative way to understand biodiversity patterns is to explore patterns of community structure. Here the focus is on the relative abundance of species rather than their total number. In studies based on presence-absence data, such as in this thesis, community structure is defined by the incidence of species at sites. Species' incidences may be treated individually or combined into the relative species richness of groups (functional groups or higher taxonomic levels). One of most prominent themes of research into community structure is an attempt to understand how species are sorted into local communities from a larger common species pool (Weiher and Keddy 1999).

Species may enter communities independently of one another. In such cases, communities are comprised of species that disperse to a site and tolerate the local conditions and the probability of colonisation for a newly arrived species is independent of the species already present in the local community (Mason 1947, Connor and Simberloff 1979, Hubbell 2001). Thus, community assembly is an individualistic process, governed largely by the effects of historical and environmental conditions on individual species' colonisation and extinction dynamics. Alternatively, species' colonisation success may be affected by the composition of the existing community (MacArthur 1972, Brown et al. 2000). If competitive interactions are strongest among species with similar traits, as is assumed by niche-tradeoff models (Fargione et al. 2003), then as communities are assembled, established species will most strongly compete with, and inhibit the settlement of, arriving species with similar resource requirements. Such processes

need not limit the overall number of species in communities, but may cause communities to assemble towards specific relative abundances of different functional guilds (Wilson 1989, Fargione et al. 2003). Thus, community structure is governed via species interactions, rather than individualistic processes.

Evidence for the extent to which species interactions govern community structure comes largely from observational data. While many of these observational studies date back to the beginning of the 20th century (Lomolino et al. 2004), Diamond's (1975) study of the avian fauna of islands in the Bismarck Archipelago initiated the formal study of assembly rules. From presence-absence data alone, Diamond proposed a set of rules defining the colonisation process in terms of permissible and forbidden species combinations and invoked interspecific competition as the process underlying the distribution of particular species pairs across islands. Subsequently, a variety of assembly rules have been proposed to describe repeatable patterns of community structure, and species interactions have typically been inferred as the processes underpinning such structure (Weiher and Keddy 1999).

Assembly rules formally describe repeatable patterns in how local communities are assembled from larger species pools. One of the major challenges associated with assembly rules is rigorously testing for the existence of a rule relative to the null expectation of individualistic community assembly. Connor and Simberloff (1979) were highly critical of Diamond's (1975) initial assembly rules, since the patterns observed in avian community structure could not be distinguished from patterns generated by the random assemblage of species from the regional species pool. Since these initial studies, a range of sophisticated statistical tools have been developed to assess patterns in community structure and to identify the extent to which species interactions influence the observed structure (see Weiher and Keddy [1999] for examples). A

feature common to all these tests is the biological null hypothesis that communities are assembled via individualistic processes and therefore constitute random samples of the regional species pool.

Some assembly rules appear to be more pervasive in nature than others. For example, Gotelli and McCabe (2002) used a null model to test the predictions of several of Diamond's rules and found evidence for significantly non-random community structure in the majority of the 96 datasets examined. In contrast, assembly rules based on the relative number of species in functional groups such as Fox's favoured states rule (Fox 1987) or Wilson's guild proportionality rule (Wilson 1989) have proved more difficult to identify. Fox's favoured states rule appears to hold in north American desert rodent communities (Brown et al. 2000), however, tests are highly sensitive to the null models used (Wilson 1995, Stone et al. 2000). Guild proportionality rules are sought primarily in plant communities, and to date, evidence for such rules has been detected only at very small spatial scales (Wilson 1999). It is noteworthy, however, that results from a manipulative study of community assembly and invasion suggest that competition within functional guilds results in community structure consistent with a guild proportionality-like rule (Fargione et al. 2003).

1.3 BENTHIC MARINE ALGAE

In marine systems, patterns of species richness, geographic range dynamics and community structure are poorly described for all but a few key groups and there is still disagreement as to the major drivers which create and maintain the patterns, (Fraser and Currie 1996, Bellwood and Hughes 2001, Hughes et al. 2002, Mora et al. 2003). For marine macroalgae, most studies have been conducted at local to regional scales and in temperate locations. Of notable absence is a

global-scale synthesis of macroalgal diversity patterns, as has been presented for groups such as angiosperms, birds and corals (e.g. Stehli and Wells [1971], Francis and Currie [2003]).

Much of the study into the broad-scale distribution patterns of marine macroalgae has revolved around the demarcation of biogeographic regions and provinces. Gradients in species richness have been documented at regional scales, and the direction of the relationship between richness and latitude depends largely on the region of study (Pielou 1977, 1978, Gaines and Lubchenco 1982, Santelices and Marquet 1998). Bolton (1994) attempted to synthesize the global relationship between algal species richness and latitude by plotting the total number of species within regions against the latitude of the region. No clear picture of consistent gradients in algal richness emerged, leading him to conclude only that "rich and poor floras occur across the globe". Longitudinal patterns have only been investigated along relatively short stretches of coastline e.g. southern Africa and southern Australia (Womersley 1990, Bolton 1996) making ocean-wide analyses impossible.

In the absence of a global view of algal richness patterns, the processes that control algal richness patterns have been discussed predominantly on a region-by-region basis. Many authors focus largely on the size and positioning of the oceans in the geologic past to explain present day distributions of algae (van den Hoek 1984, Santelices and Abbott 1987, Bolton 1996). Present day climatic features, primarily temperature but also currents, the extent of benthic habitat and day length, are also frequently cited as causes of the observed patterns of algal distribution (Luning 1990, Womersley 1990, Breeman and Pakker 1994, Phillips 2001). Few formal tests of the drivers of algal richness patterns have been conducted at a global scale. The one exception is Adey and Steneck (2001), who developed a thermogeographical model which

suggests that sea surface temperature and area, both integrated over evolutionary time, are important determinants of crustose coralline distribution and abundance.

The structure of algal assemblages has only been investigated at very coarse scales. The R:P (Rhodophyte:Phaeophyte) ratio is commonly reported as a measure of the "tropical-ness" of assemblages and has been shown to decline with latitude (Luning 1990). Alternatively, an assemblage is described according to the proportions of the flora that have tropical, temperate and polar affinities. For example, the southern Australian Chlorophyta has been reported to be comprised of 85% temperate algae, 10% tropical algae and 5% polar algae (Phillips 2001). However, the biological significance of such statistics is questionable. For instance, although *Caulerpa* is considered to be a tropical genus, 15 of the 19 species found in southern Australia are endemic to this temperate region (Womersley 1984).

1.3.1 Benthic marine algae as a model group

Benthic marine algae are an ideal group to use as a model system with which to test major ecological hypotheses about biodiversity and community structure. Algae are globally distributed, occurring throughout tropical, temperate and polar regions, in both the Indo-Pacific and Atlantic Oceans. To date, marine macroecological analyses have focused on molluscs along the coastlines of the Americas and on tropical corals and reef fishes in the Indo-Pacific (e.g. Roy et al. [1998], Hughes et al. [2002]). Thus, examining macroalgae allows me to both test the taxonomic generality of existing results and to address them at a truly global scale. Furthermore, macroalgal genera display some unique patterns (namely temperate richness peaks, Chapter 2) which allow for testing of ecological hypotheses outside of conventional systems (i.e., those which display typical low latitude richness peaks).

On a practical level, there is a long history of phycological study at many locations world wide and species inventories are readily available. This facilitates the development of a global database of algal distributions, from which geographic ranges can be constructed and data on richness and community structure can be extracted. At the genus level, there are >1000 genera of benthic marine algae, providing large enough numbers with which to conduct meaningful analyses without compromising accuracy by using species-level taxonomy (which may be inconsistent for many groups). Some groups, however, are conspicuous and well enough studied to enable testing of the concordance between genus and species level patterns. Furthermore, the wide geographical and habitat distributions of higher level taxa allows robust multi-scale comparisons of community structure. Finally, macroalgae genera can be characterized into functional groups based on rates of biomass production, canopy formation and disturbance response (Steneck and Dethier 1994) and this allows analyses of diversity at an additional level.

1.4 THESIS AIMS AND OUTLINE

The overall aim of this thesis is to use benthic marine algae as a model group to test, on a global scale, some of the most common theories of biodiversity and to improve understanding of the processes that drive the observed patterns in the marine realm. More specifically; the first major aim of this thesis (Chapter 2) was to quantify global and oceanic gradients in macroalgal diversity, using a custom-built database of algal presence-absence from 391 sites world-wide. Analyses were performed on two levels; for all genera of benthic marine macroalgae and for species in the Order Bryopsidales; a group of predominantly reef-associated algae, which is both well surveyed and taxonomically stable (Littler and Littler 2003). In order to gain a first approximation of the processes that drive the observed patterns, in this chapter I have also examined algal richness patterns in the context of sizes and locations of algal geographic ranges.

The second aim of the thesis (Chapter 3) was to understand patterns of marine biodiversity more generally. The order Bryopsidales has species richness patterns in the tropical Indo-Pacific Ocean that are strikingly concordant with those of corals and reef fishes. I apply a recent model developed to explain the drivers of coral and fish richness patterns to bryopsidalean algae. Specifically, I test for the generality of mid-domain effects and area variables as drivers of marine diversity patterns.

The third major objective was to examine the nature of assemblage structure in macroalgal communities. More specifically, I aimed to identify the assembly rules according to which macroalgal communities are structured, the geographical scales over which assembly rules are shared, and if environmental and historical differences among regions can give rise to differences in assembly rules among biogeographic regions. This required the development of new tests of guild proportionality rules (Chapter 4) and is followed by their application to the algal data (Chapter 5).

Chapter 2. Global biodiversity patterns of benthic marine algae[1]

2.1 INTRODUCTION

Global-scale patterns in species richness are one of the most studied phenomena in ecology and consistent trends are repeatedly documented in both terrestrial and marine environments (Gaston 2000). The latitudinal gradient of increasing species richness from polar to tropical regions is arguably the most well known of these patterns and few taxa display contrasting patterns (Willig et al. 2003). Longitudinal gradients, although less well studied, can also be distinct across individual continents and oceans (e.g. Jetz and Rahbek [2001] for birds, Roberts et al. [2002] for marine taxa). However, there is much controversy about the mechanisms that underlie the observed patterns of biodiversity, with upwards of 30 explanations for the latitudinal richness gradient alone (Rosen 1988, Willig et al. 2003).

Most theories that aim to explain the location of biodiversity hotspots typically rely on logical explanations for why one area will promote speciation or reduce extinction relative to another. For example, the species-area hypothesis suggests that larger areas can support more individuals and populations thereby reducing extinction risk, whilst also containing more barriers that promote allopatric speciation (Rosenzweig 1995, Chown and Gaston 2000). The species-energy hypothesis asserts that higher numbers of species in the tropics result from faster metabolic and speciation rates associated with warmer temperatures (Kaspari et al. 2004). Theories that revolve around climatic stability suggest that the tropics are a stable and relatively benign environment where species can specialize on predictable resources and persist when rare, compared to harsh temperate and polar regions where extinction rates are high (Hawkins et al.

[1] This chapter is published as: Kerswell AP (2006) Global biodiversity patterns of benthic marine algae. Ecology **87**:2479-2488.

2003a). The species-productivity hypothesis suggests that greater energy inputs will support more individuals and promote specialization, although whether the relationship between species numbers and productivity is linear or hump-shaped appears to be highly scale dependent (Rosenzweig 1995, Chase and Leibold 2002).

Recently mid-domain effect models have been proposed as biogeographic null models (Colwell and Lees 2000, Colwell et al. 2004). Under these models, species richness gradients arise due to geometric constraints on the size and location of species' geographic ranges within a bounded domain in the absence of environmental factors. For species with wide ranges, the geometry of the domain has significant impact on species richness patterns, in that their ranges are most likely to overlap in the centre of the domain resulting in high numbers of species in the mid-domain region. This has been demonstrated both for African birds and Indo-Pacific corals and reef fishes (Jetz and Rahbek 2001, Connolly et al. 2003). In contrast, small ranging taxa can occur anywhere inside the domain boundaries. Under a mid-domain effect null model, they should be uniformly distributed across the domain, however, in nature, small ranging taxa frequently cluster in common locations. For example, small ranging birds cluster in pockets along the margins of the African continent, peripheral to the middle of the domain (Jetz and Rahbek 2001). Small ranging African Proteaceae also cluster away from the domain centre in southern temperate regions (Laurie and Silander 2002). Such deviations from null model predictions suggest the need to evaluate alternative causes of species richness patterns (Colwell et al. 2004).

To date, biogeographic studies of benthic marine macroalgae have been restricted to regional scales, with little synthesis of worldwide trends. Japan, southern Australia and western Europe are consistently highlighted as regions of high algal diversity, with the polar oceans, west Africa

and south-east Pacific identified as depauperate areas (Silva 1992, Bolton 1994, Santelices and Marquet 1998). Searches for latitudinal gradients in algal richness have also been performed on regional scales, with the most extensive of these studies indicating possible temperate richness peaks along the coastlines of the Americas (Pielou 1978). There is some indication that the most widespread red algal genera have extremely large ranges which extend across entire oceanic domains (Joosten and van den Hoek 1986). Small ranging algal genera appear to be clustered in temperature latitudes along both the east and west coasts of the Americas (Pielou 1978). In contrast, Santelices and Marquet (1998) found some evidence of increasing range size towards higher latitudes along the European coastline.

Algal distribution patterns and regional assemblage composition have been explained largely in terms of historical processes. Early studies of algal biogeography discuss present day patterns as a direct result of tectonic changes over geological time and shifts in species ranges as sea levels and temperature regimes fluctuated (Joosten and van den Hoek 1986). Recently, Adey & Steneck (2001) developed a model that defines thermogeographic regions based on temperature and habitat area since the Pleistocene. They also define biogeographic regions based on the presence, abundance and level of endemism of crustose coralline algae and find that the two definitions produce matching regions. This suggests that energy and habitat area play an important role in determining the present day macroecological patterns observed for benthic marine algae.

The major aim of this chapter was to quantify global and oceanic gradients in macroalgal diversity. On a global-scale, macroalgae display diversity gradients with unique features, such as temperate richness peaks. To understand how such unique patterns are generated, I quantify the manner in which geographic ranges of individual taxa combine to produce the observed

patterns. These analyses were performed on two levels; for all genera of benthic marine macroalgae and for species in the Order Bryopsidales; a group of predominantly reef-associated algae, which is both well surveyed and taxonomically stable (Littler and Littler 2003). While hypotheses about the causes of diversity gradients predict a positive correlation between environmental variables and species numbers, they also contain underlying assumptions about how species' ranges are distributed within biogeographic realms (e.g. Chown and Gaston [2000]). Yet patterns in the distribution of species' ranges are often neglected when searching for processes that create and maintain diversity patterns. Therefore, in this chapter, I have investigated algal richness patterns in the context of sizes and locations of algal geographic ranges in order to gain a more comprehensive understanding of the processes that determine marine diversity.

2.2 METHODS

A global database of benthic marine algae occurrence records was compiled from 191 species lists sourced from the primary literature (141 peer-reviewed papers, 23 books and 14 university-published scientific reports). In total, the database contained 387 sites throughout the Atlantic, Indo-Pacific and Southern Oceans, which spanned 140^{o} of latitude (Figure 2.1, see Appendix 1 for full list of species list references and corresponding sites). At each site, genus-level data were compiled for all fully marine macroalgae in the Classes Rhodophyceae, Phaeophyceae and Chlorophyceae and species-level data were compiled for all algae in the Order Bryopsidales. Algae were entered under the taxonomic classification listed in Guiry et al. (2005). Drift specimens and records noted as questionable by species list authors were excluded. The resulting database contained 1069 genera of marine algae and 388 species of Bryopsidales.

The occurrence database was interfaced with a Geographic Information System (GIS, ArcView 3.2a, ESRI). Each site and algal taxon were unique records, linked via >45,000 occurrence records. Records were verified against the primary literature if a single point occurred in the tropics when all other points were in temperate areas (and vice-versa) or if an isolated point occurred in any ocean basin where there were no other occurrence records (Atlantic, Pacific and Indian Ocean basins). To err on the side of caution, such records were deleted unless special reference was made by the authors as to the veracity of the record (e.g. *Sporochnus moorei* in Hawai'i (Abbott 2004)).

Genus and species ranges were constructed in the GIS. A range was defined by outlining the boundary of all sites at which an alga was reported to occur and was plotted on an equal-area projection basemap. Ranges were not extended over vast oceanic expanses e.g. the east Pacific barrier, nor over areas for which reliable species lists could not be obtained, e.g. Antarctic Ocean (outer contours in Figure 2.2 A delineate the maximum range boundary). The area of the geographic range (in km^2) was determined for each taxon in the GIS and ranges were partitioned into Indo-Pacific and Atlantic elements for algae that occur in both oceans. Range size frequency distributions are presented on logarithmic plots. Because a direct comparison of Indo-Pacific versus Atlantic range sizes would be confounded by the vastly different areas of the two oceans, algal range areas were also expressed as a proportion of the total area of each ocean. The distributions of standardized range sizes were then compared between oceans using two-sided Kolmogorov-Smirnov testing ($R_{2.0.1}$, R Development Core Team [2004]). The size and location of species' geographic ranges were explored for all algal taxa by plotting the latitudinal and longitudinal range extent against the location of range midpoints. In order to clearly identify clustering of algae, mid-point analyses are presented as 2-dimensional density plots ($R_{2.0.1}$, kde2d function).

Diversity patterns of benthic marine algae were explored by generating contours of algal richness. An estimate of diversity at each of the 387 sites in the database was generated by summing the number of ranges which overlapped that site. Range-derived diversity estimates were then used to interpolate contours of genus and species richness using the Inverse Distance Weighted interpolator (ArcView 3.2a GIS, 0.5^{o} cells, nearest neighbour technique with 12 neighbours, 3^{rd} order power, no barriers). Thus contours represent the maximum diversity of a region, within which local sites may have lower diversity.

Endemics were defined as taxa reported at only one location or with a geographic range size $<1x10^{6}$ km^{2} for genera and $<0.5x10^{6}$ km^{2} for species (c.f. Hughes et al. [2002]). These areal cutoffs are smaller than 0.5% of the largest geographic range recorded for each group. The location of endemics was explored by producing contour maps of the number of endemic taxa occurring at sites across the Indo-Pacific and Atlantic oceans. Contours were generated using the Inverse Distance Weighted interpolator (as above). Range-derived estimates of richness and the number of endemics used to generate contour maps are available in Appendix 1.

2.2.1 Data quality control

Two major challenges in biogeographic studies are changes and inconsistencies in taxonomic identification and the delineation of geographic ranges from a set of sampling locations at biogeographic scales. When analyzing data on biogeographic scales, small differences arising from human error are not likely to alter conclusions significantly (Brown et al. 1996). Nevertheless, I have implemented several strategies to minimize bias due to taxonomy and sampling effects in the delineation of geographic range boundaries.

The taxonomic status of each genus/species was verified in AlgaeBase (www.algaebase.org, Guiry et al. [2005]) and older taxonomic classifications were updated to reflect the name assigned by Guiry et al. (2005). Genus, rather than species level classifications were used to determine richness patterns of all benthic marine macroalgae. Genus level classifications were deemed more robust to mis-identification and changes in systematics, and patterns identified at the genus level are often matched by species level data (e.g. see Veron [1995] for corals). In order to directly compare richness patterns between a number of marine groups, species level patterns were analyzed for the Order Bryopsidales. The bryopsidales were chosen due to their relative taxonomic stability (e.g. when plotting the number of new bryopsidalean species against time, the curve reaches an asymptote around 1970, with very few new species being described after this time. Only 3% of species list pre-date 1970). Moreover, when patterns for the bryopsidales are analyzed at the broader genus level, the patterns are highly consistent with those generated by species level data.

Sampling bias is an important concern when creating biogeographic ranges from species lists at particular locations, particularly when areas of high diversity coincide with areas that are traditionally well studied (Bolton 1994). However, very few taxa have been sufficiently intensively surveyed to generate global richness patterns based on occurrence records alone (Hurlbert and White 2005). An alternative method is to use geographic ranges to estimate richness across sites. This method is recognized as generating more realistic estimates of diversity than raw occurrence data, which tend to be more inconsistent and/or incomplete (McAllister et al. 1994) and has been used extensively for analyses conducted at continental to global scales (e.g. 80% of broad-scale studies on terrestrial plants, vertebrates and invertebrates used range-derived richness estimates (Hawkins et al. [2003a]) as have several studies of corals and reef fishes (McAllister et al. [1994], Veron [1995], Bellwood et al. [2005])). Since the

global effort of phycological study is patchy (Silva 1992, Bolton 1994), I used geographic ranges to estimate diversity at each site to minimize the bias associated with sampling effort. Furthermore, I tested for a dependence of algal genus richness on either the number of sites or phycological studies in 14 regions world-wide. Linear regression showed no relationship between richness and either the number of sites (p=0.47, $R_{2.0.1}$) or the number of phycological studies (p=0.365, $R_{2.0.1}$) (see Figure 2.1A,B for regions and regressions), indicating that genus richness was not merely a reflection of variation in sampling intensity.

2.3 RESULTS

2.3.1 Centres of diversity and endemism

Centres of genus diversity for benthic marine algae occur in temperate oceans. In the Indo-Pacific Ocean, centres of diversity occur in southern Australia and Japan, each containing 350-450 genera of algae (Figure 2.2A). The Indo-Australian Archipelago (IAA) and southern Indian Ocean have moderate richness of approximately 250-300 genera. Thus, there is a band of high algal diversity running longitudinally between 110°-160°E. Richness attenuates to the east and west of this band, reaching ~150 genera in the Red Sea and along the Chilean coastline. The areas of lowest diversity occur in the polar regions where fewer than 100 genera have been recorded. Algal richness gradients in the Atlantic Ocean are both latitudinally and longitudially asymmetrical. The eastern coastline has higher diversity than the west with the major Atlantic biodiversity hotspot located along the European coastline, extending south to Morocco (250-300 genera). Additionally, twice as many genera occur in the northern versus the southern Atlantic. Endemic algal genera cluster in areas of high diversity within the respective oceans (Figure 2.2B). At most of the 387 sites across both the Indo-Pacific and Atlantic Oceans, there are fewer than two endemic genera, however this number increases to as many as 21 inside the Japan

biodiversity hotspot. Endemic genera comprise <6% of the flora at any site, including those within biodiversity hotspots.

In contrast to the patterns of all algal genera, centres of diversity for the Order Bryopsidales are located in the tropics and species richness diminishes both latitudinally and longitudinally away from these hotspots (Figure 2.3A). In the Indo-Pacific, species richness is highest in the IAA, while the Atlantic Ocean centre of diversity is located in the central Caribbean. However, diversity in the Atlantic is low, with the majority of areas containing <30 species compared to >90 in the Indo-Pacific. While Indo-Pacific species endemics occur mainly outside areas of high diversity (in India, Japan and Hawaii) most of the endemics in the Atlantic occur in the Caribbean, where species diversity is highest (Figure 2.3B).

2.3.2 Geographic range size distributions

Range size frequency distributions (RSFD) are left skewed on a logarithmic scale (Figure 2.4), highlighting that many algae have very large geographic ranges. Endemics comprise a small proportion of the total number of genera in each ocean (12% in the Indo-Pacific and 7% in the Atlantic). In contrast, 40% of Indo-Pacific and 33% of Atlantic genera have ranges > 10 million km^2 (i.e., ranges 10x larger than endemics). A large percentage of bryopsidalean species are also wide-ranging (44% in the Indo-Pacific and 23% in the Atlantic have ranges > 5 million km^2 or 10x the endemic range size). In contrast to genus level patterns, endemic species are significant with bryopsidales comprising 23% and 20% of the Indo-Pacific and Atlantic flora respectively.

There are significant differences between RSFDs of all algal genera and bryopsidalean species in the Indo-Pacific and Atlantic Oceans (genera: D=0.1814, p<0.001; species: D=0.150, p=0.023). RSFDs in the Indo-Pacific are bimodal with peaks in the middle and the largest size

classes (Figure 2.4A,C). In the Atlantic, the proportion of algae in each range size class increases from left to right (i.e., from small to large), but decreases sharply in the largest size categories (Figure 2.4B,D).

2.3.4 Location and overlap of geographic ranges

There are striking patterns in the location and overlap of algal geographic ranges. Under a mid-domain effect null model, 2-dimensional density plots of range extent versus mid-point should be horizontally uniform within the triangular domain boundaries (Lees et al. 1999). Yet in both the Indo-Pacific and Atlantic Oceans, there is obvious clustering of algal ranges. In the Indo-Pacific, algal genera and species of Bryopsidales with large latitudinal extents cluster near the equator (Figure 2.5A,C). In contrast, small ranging genera cluster in both the northern and southern hemispheres away from the middle of the domain in temperate latitudes (Figure 2.5A). Small ranging bryopsidales are spread throughout all latitudes of the Indo-Pacific with a cluster in the southern temperate region (Figure 2.5C). Longitudinally, algal genera with small and large range extents are centred in the middle of the Indo-Pacific between 110-170^{o}E (Figure 2.5B). Longitudinally restricted bryopsidales are also centred within this band, however, larger ranging species are centred to the western side of the domain, leaving the eastern side relatively species poor (Figure 2.5D).

The Atlantic Ocean is latitudinally highly asymmetrical, with the majority of genera and species having range mid-points north of the equator (Figure 2.6A,C). Algae with large latitudinal extents are centred around 10^{o}N, while latitudinally restricted genera occur predominantly around 50^{o}N in the temperate ocean. A large proportion of small ranging bryopsidales have latitudinal mid-points centred in the northern hemisphere tropics (~20^{o}N). There is a striking contrast between the longitudinal clustering of genera and species in the Atlantic Ocean. The

majority of genera are centred in the middle of the Atlantic domain, while most species of Bryopsidales have ranges centred on either the eastern or western ocean margins (Figure 2.6B,D).

2.4 DISCUSSION

2.4.1 Marine biodiversity patterns

Latitudinal and longitudinal gradients

Diversity maps of the global distribution of benthic marine algae reveal distinct gradients in species and genus richness. Prior to this analysis, the documentation of latitudinal gradients in algal diversity was restricted to small regional scales (Pielou 1978, Santelices and Marquet 1998), leading to speculation that macroalgae do not exhibit global latitudinal gradients (Willig et al. 2003). The results of this chapter clearly show that this is not the case. In the largely reef-associated Order Bryopsidales, diversity peaks at tropical latitudes and decreases steadily towards the poles, in a manner well documented for other tropical marine organisms (Rosen 1988). However, when all algal genera are considered, temperate regions consistently have higher algal richness than tropical areas. In both the Indo-Pacific and Atlantic Oceans, algal centres of diversity occur in temperate areas, with richness decreasing towards the tropics and polar regions. Interestingly, this trend was also identified by early studies using more geographically restricted data sets (Pielou 1978, Vermeij 1978, Gaines and Lubchenco 1982). The peak in algal genus richness at mid latitudes on a global scale makes benthic marine algae an exceptional group, in that there are very few taxa that have diversity peaks outside of the tropics (Willig et al. 2003).

Marine algae also display distinct longitudinal richness gradients. In the Indo-Pacific, algal richness peaks at the same longitudes as the richness of many other taxa (Rosen 1988, Roberts et al. 2002) resulting in a band of exceptionally high diversity between 110^{o}-160^{o}E. Similarly in the Atlantic, the Order Bryopsidales, along with many other coastal marine taxa (Macpherson 2002), reaches peak richness on the tropical western coastline. However, when all algal genera are considered, the greatest diversity occurs on the eastern Atlantic coastline. While this is not unique within temperate regions (Macpherson 2002), the richness of the north-eastern coast is usually significantly less than that of the tropical Caribbean. This is clearly not the case for algal genera, further highlighting the exceptional nature of global algal diversity patterns.

Size and location of geographic ranges

Patterns in the size and location of geographic ranges can provide insights into the mechanisms regulating diversity that are unavailable solely from examining variations in species numbers across a biogeographic realm. Endemic algae occur throughout the Indo-Pacific and Atlantic Oceans and yet represent only a minor element of the total flora at any site. This is in stark contrast to many terrestrial systems, where diversity hotspots are generated largely by an accumulation of endemic taxa (Myers et al. 2000). Furthermore, in the Indo-Pacific, bryopsidalean endemics tend to be clustered outside of the IAA hotspot, in more peripheral depauperate locations, which is consistent with patterns documented for coral and reef fish endemics (Hughes et al. 2002, Paulay and Meyer 2002, contra Mora et al. 2003).

It is also important to consider the observed size and location of ranges in comparison to an appropriate null expectation. Where deviations from such an expectation occur may provide further insight into mechanisms that regulate diversity patterns (Colwell et al. 2004). For example, in the Atlantic Ocean, large ranged taxa are clustered in the centre of the ocean and

small ranged taxa on the eastern and western boundaries. The geometry of the Atlantic Ocean does not allow these taxa to be located anywhere else and thus, no other explanation of richness patterns need be invoked. However, algal ranges are located predominantly in the northern Atlantic and in the Indo-Pacific large ranged algae are clustered within tight latitudinal limits in the middle of the domain while small ranging taxa are disproportionately clustered away from the centre in temperate areas. A similar clustering is also evident for Indo-Pacific corals and reef-fishes (Connolly et al. 2003) and such distributions of range locations can be generated when environmental gradients are incorporated along with geometric constraints in process-based models (Connolly 2005). Thus, identifying deviations from patterns generated solely by geometric constraints provides a basis from which to assess environmental drivers of species richness patterns.

4.4.2 Causes of biodiversity patterns

A consideration of patterns of algal richness and range size and location indicates that the climatic stability, species-energy and competition hypotheses are unlikely to be major drivers of algal richness patterns. The key assumption of the climatic stability hypothesis, that specialization in benign tropical habitat leads to a decline in range size towards low latitudes (Stevens 1989), is clearly not reflected in either genus or species level patterns of algal range size. Furthermore, overall algal richness is not highest in the tropics where metabolic processes are thought to enhance speciation (Kaspari et al. 2004). Competition with corals is commonly cited as an explanation of lower algal richness in the tropics relative to temperate regions (e.g. Fraser and Currie [1996], Miller and Hay [1996]). However, bryopsidalean richness peaks in the tropics in a manner very similar to corals (compare Figure 2.3A with coral richness patterns in

Roberts et al. 2002). Bryopsidales are predominantly reef-associated algae and are likely to be in direct competition with corals for space (Littler and Littler 2003).

While the nature of the productivity-diversity relationship is highly scale dependent (Chase and Leibold 2002), there appears to be little evidence for productivity as a driver of macroalgal diversity at a global scale. Phytoplankton net primary productivity (NPP) is greatest at high latitudes in the northern hemisphere, along western continental boundaries and in the sub-tropical convergence zone, due primarily to enhanced nutrient availability associated with major oceanic upwelling (Field et al. 1998, Behrenfeld et al. 2001). While the three centres of algal genus richness are all located within these areas of peak productivity, areas of high NPP also encompass regions of both moderate (e.g. southern Africa and California) and low algal richness (e.g. Chile and New Zealand). Moreover, oceanic productivity is very low in the Caribbean basin and the IAA where bryopsidalean richness is greatest.

The species-area hypothesis can explain some, but not all, of the global patterns of algal richness. Bryopsidalean species richness patterns closely mirror those of corals and reef fishes, suggesting a common regulatory mechanism. Recently, the area of coral reef has been shown to account for a large proportion of the variation seen in coral and reef fish richness patterns (Bellwood et al. 2005). Within temperate areas, regions of highest algal richness also coincide with large areas of suitable habitat (Silva 1992). However, while the species-area hypothesis can account for the location of peaks in both algal genus and species richness, it cannot explain why temperate hotspots support more genera than do equivalent tropical regions.

Major ocean currents may play an important role in determining the location of algal richness hotspots through propagule dispersal and alteration of oceanic conditions. Ocean gyres flow clockwise in the northern hemisphere and anticlockwise in the southern hemisphere and as a

result, surface circulation in tropical regions is dominated by westward flowing equatorial currents. These currents leave the tropics, travelling poleward along western ocean boundaries and back towards the tropics along the eastern edges. If currents are influential in determining richness patterns through dispersal then the greatest richness of tropical algae should occur in western ocean regions along with depauperate tropical floras in the east. Furthermore, if equatorial currents extend the geographic ranges of tropical algae into temperate regions, then high overall algal richness will occur where tropical and temperate floras overlap. This will be most prominent on the western ocean margins where both tropical and temperate floras are species rich compared to the eastern margins where tropical floras are depauperate. Current driven richness patterns should also result in distinct patterns of range size. Ranges of tropical algae centred at low latitudes are expected to be large and the ranges of algae that originate in temperate regions small, as temperate to tropical dispersal will be restricted.

The richness and range size patterns observed for algae in the Indo-Pacific Ocean are consistent with those predicted in an ocean currents-driven system. An exception is the high diversity of tropical bryopsidales in the eastern Indian Ocean, which is unexpected under an ocean-gyres model. However, the coastal east-Indian Ocean is not dominated by the northward flowing Indian Ocean gyre. Rather, the tropical Leeuwin Current flows from the equator towards the south pole along the east-Indian ocean margin (west Australian coastline), on the inside of the northward flowing Indian Ocean gyre. Furthermore, the Leeuwin Current flows directly into the south Australian temperate hotspot from equatorial regions. Similar mechanisms have been suggested to influence richness patterns of corals and reef fishes (Connolly et al. 2003). Thus currents in the Indo-Pacific appear to play a role in creating and maintaining algal hotspots in temperate regions, where tropical and temperate floras overlap.

Algal richness is low throughout the southern Atlantic, however, unlike the southern Indo-Pacific, the influence of equatorial currents is only modest compared to sub-polar currents and upwelling systems (Pickard and Emery 1990). In the tropical Atlantic, bryopsidales species richness is highest in the western ocean (Caribbean) as expected, however, overall genus richness peaks in the north-eastern temperate regions. The north Atlantic is dominated by the Gulf Stream, which originates in western tropical regions then flows north and east into temperate areas. Water originating in the tropics extends across the north Atlantic and flows north to Scotland and south to Africa (Pickard and Emery 1990), where algal richness is highest. In contrast, the low diversity north-west Atlantic coasts are dominated by the Labrador Current, which originates in the Arctic and extends south to Cape Hatteras (~35^{o}N) (Pickard and Emery 1990). Hence ocean currents also appear to influence algal richness patterns in the Atlantic Ocean.

4.4.3 Conclusions

This chapter provides the first world-wide assessment of patterns in algal richness which incorporates a quantification of the size and location of algal geographic ranges. Macroalgal genera show atypical latitudinal richness gradients and this trend is consistent across oceans and biogeographic realms, suggesting that it is truly an exceptional global phenomenon rather than a regional anomaly. Hypotheses about the causes of diversity gradients generally assume a straightforward positive correlation between environmental variables and the number of taxa found at sites along that gradient. However, underlying these theories are fundamental assumptions about the relationship between the environmental variables and the size and location of species' geographic ranges (Stevens 1989, Chown and Gaston 2000). Yet patterns of geographic range size and location are rarely considered as evidence for or against the various

hypotheses. When they are, we gain additional insights into the drivers of richness patterns that would have been unavailable from analyses of species numbers alone (Lees et al. 1999, Connolly et al. 2003). This chapter explicitly quantifies both richness gradients and the distribution of range sizes and locations that give rise to those gradients. By conducting such analyses, especially for a group of organisms with exceptional richness gradients, important inroads have been made into gaining a more comprehensive understanding of how richness gradients are created and maintained.

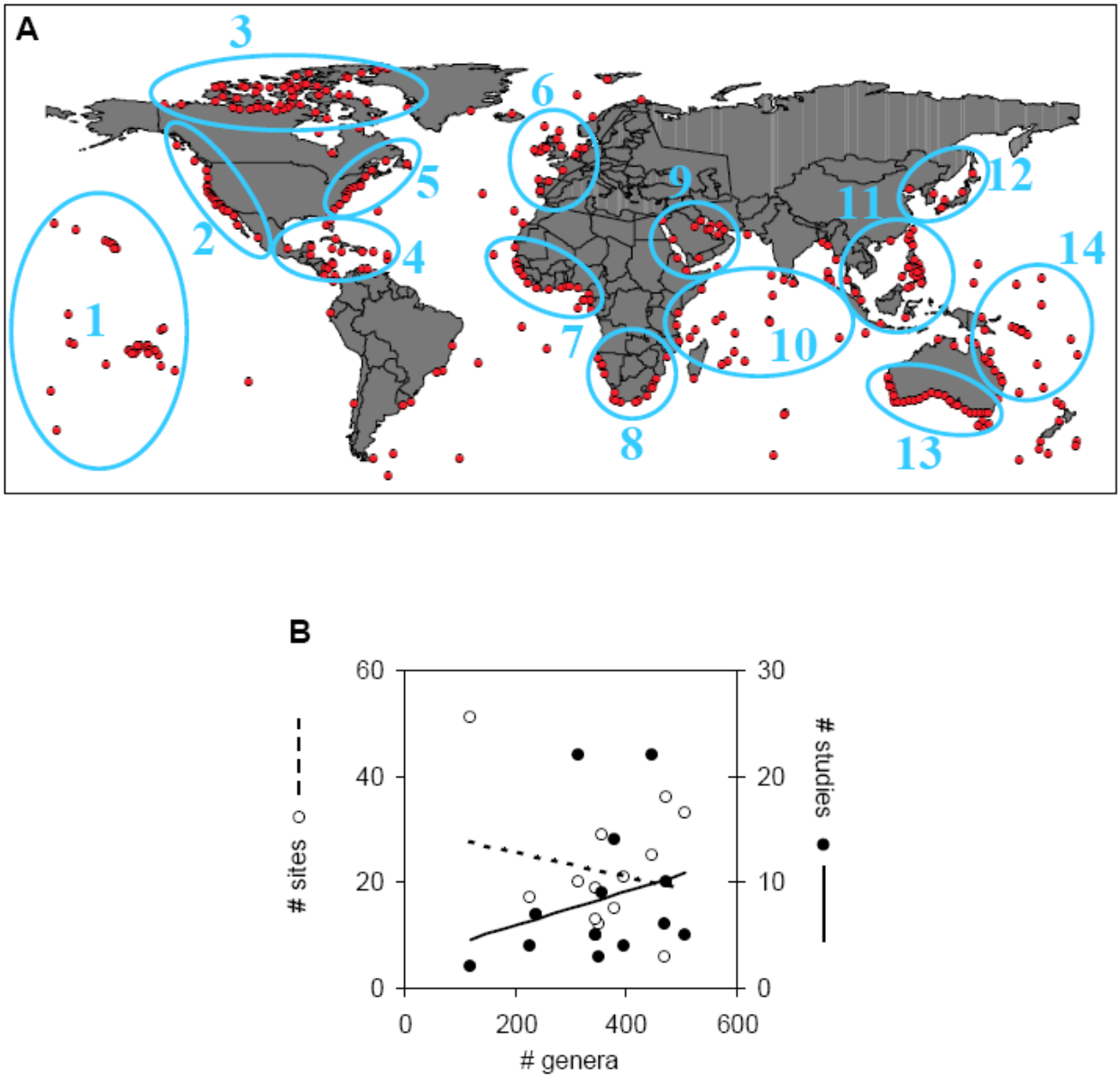

Figure 2.1. A) Map of the Indo-Pacific and Atlantic Oceans showing sites marked as red points and regions delineated by blue lines. B) Linear regressions of the genus richness of regions and the number of sites and phycological studies in the region.

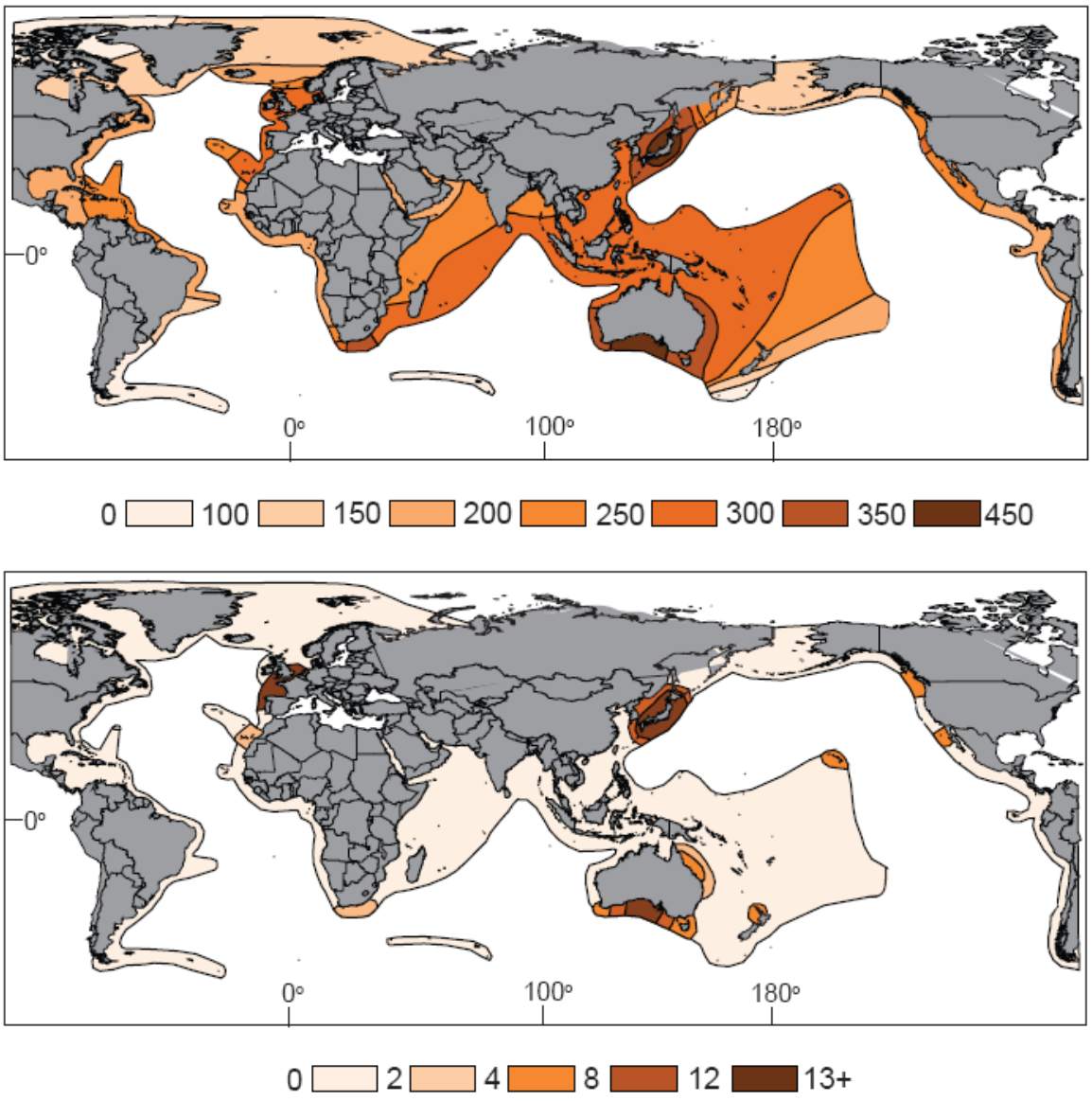

Figure 2.2. Map of the Indo-Pacific and Atlantic Oceans showing contours of A) algal genus richness and B) clusters of endemic genera. Unshaded areas represent oceanic expanses or regions for which reliable data were not available.

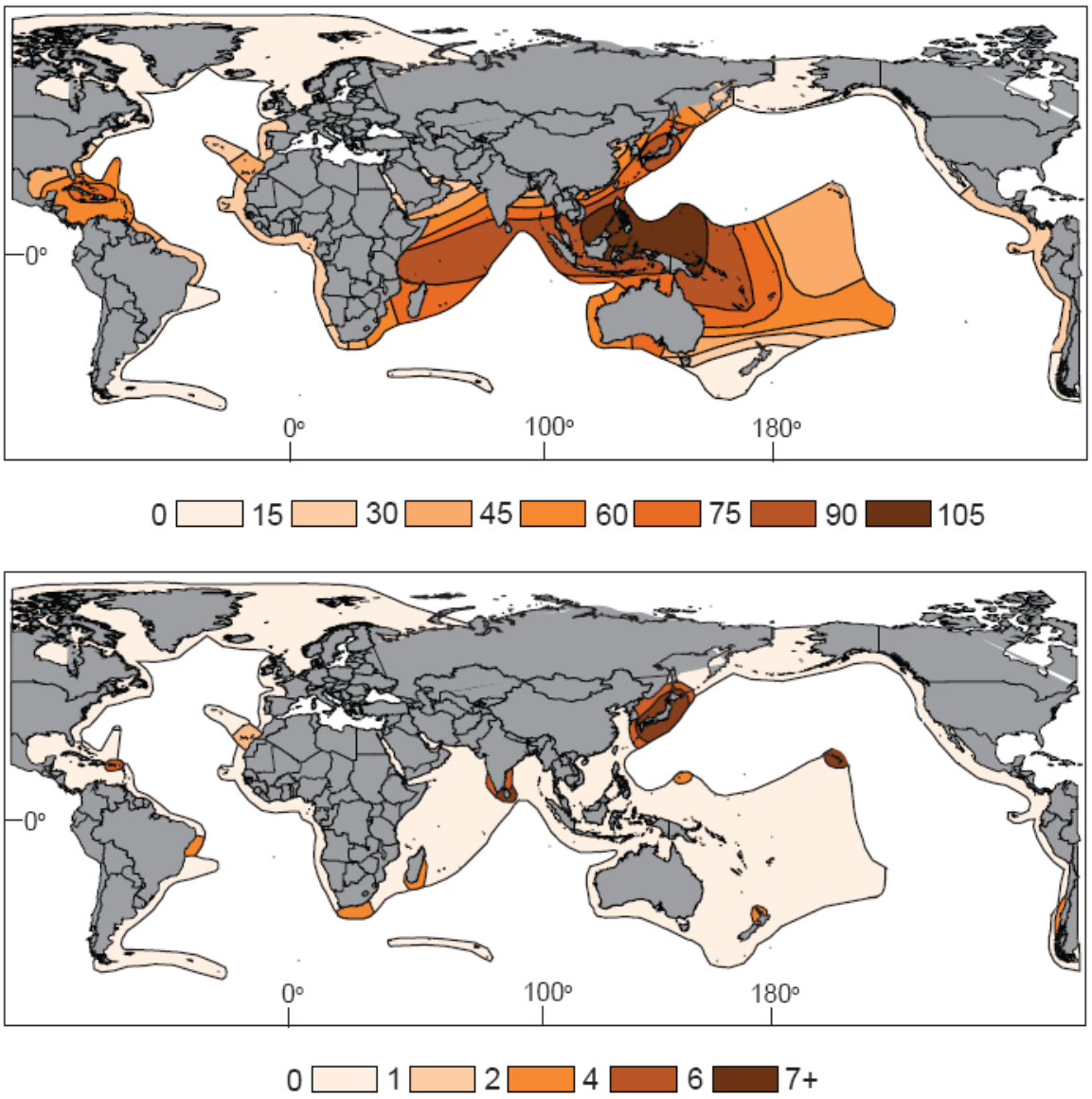

Figure 2.3. Map of the Indo-Pacific and Atlantic Oceans showing contours of A) species richness and B) clusters of endemics within the Order Bryopsidales. Unshaded areas represent oceanic expanses or regions for which reliable data were not available.

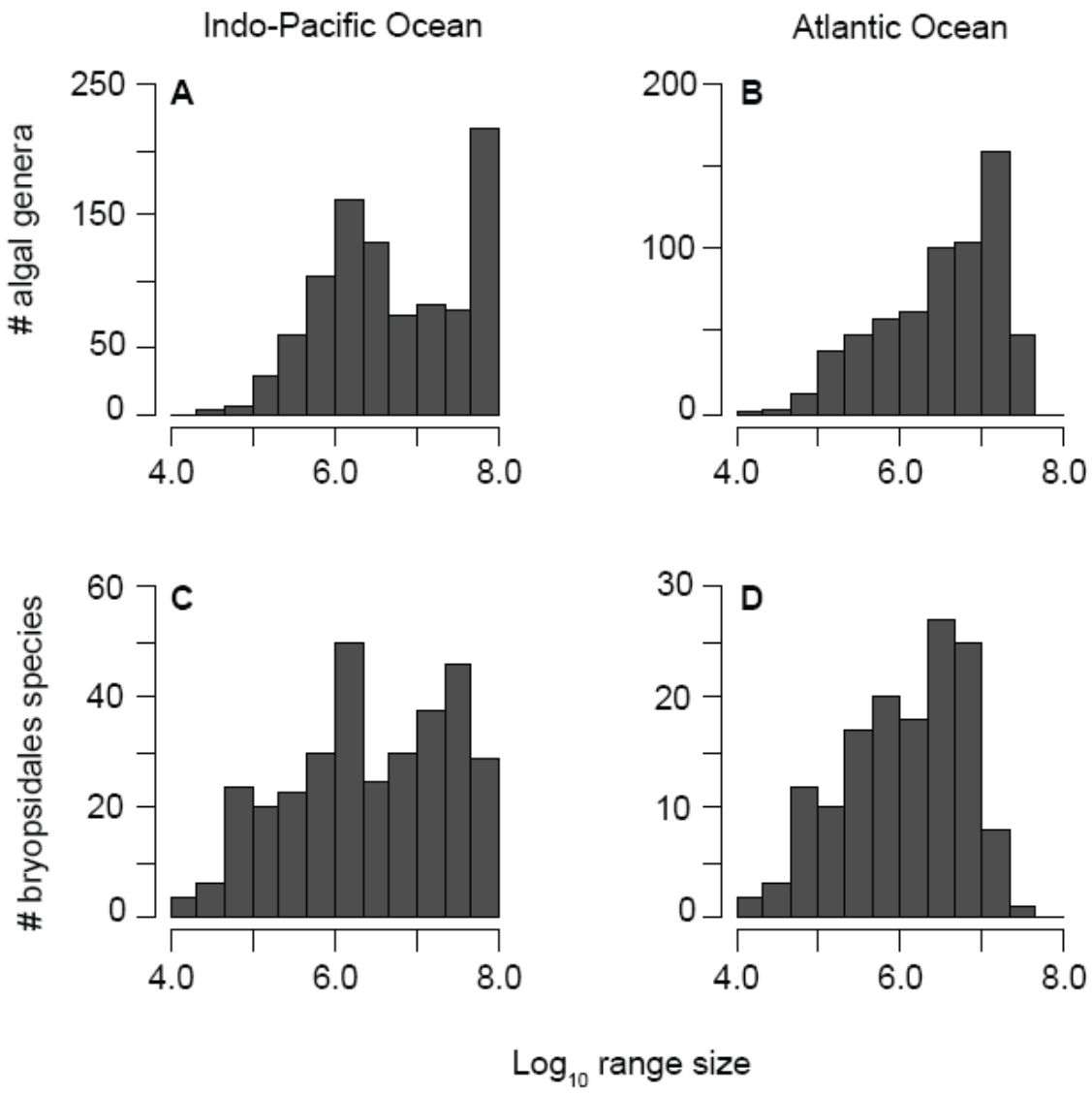

Figure 2.4. Frequency distribution of geographic range size (RSFD) of all algal genera and Bryopsidalean species in the Indo-Pacific (A, C) and the Atlantic Ocean (B, D) on a logarithmic scale.

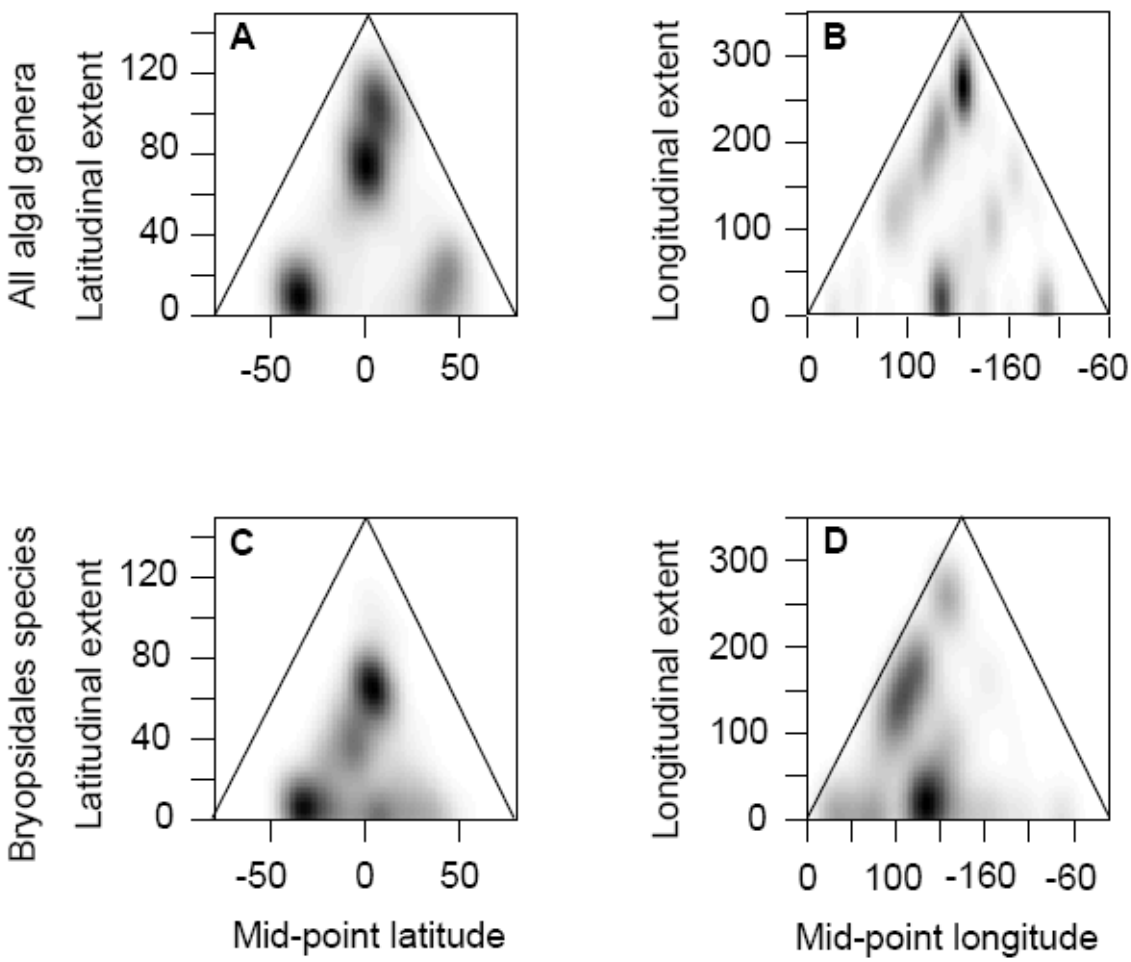

Figure 2.5. Density plots of range extent and range midpoint location for the algal genera (A, B) and bryopsidalean species (C, D) in the Indo-Pacific Ocean. Dark areas represent high concentrations of midpoints, whilst in white areas there are very few midpoints. Triangles indicate domain boundaries (as per Mid-domain Effect (Colwell and Lees [2000])).

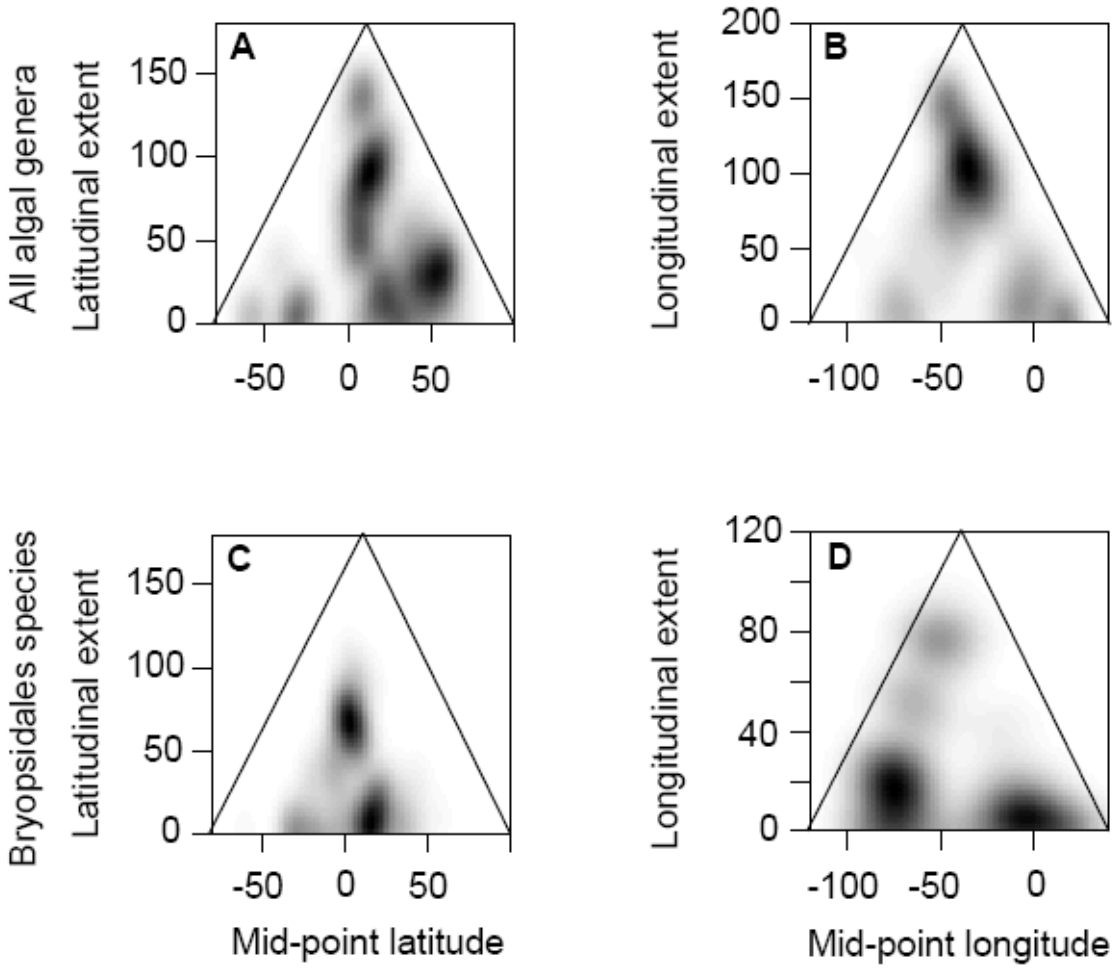

Figure 2.6. Density plots of range extent and range midpoint location for the algal genera (A, B) and bryopsidalean species (C, D) in the Atlantic Ocean. Dark areas represent high concentrations of midpoints, whilst in white areas there are very few midpoints. Triangles indicate domain boundaries (as per Mid-domain Effect [Colwell and Lees (2000)]).

Chapter 3. Drivers of marine biodiversity patterns: testing predictions with macroalgae[2]

3.1 INTRODUCTION

On land, the relationship between global diversity patterns and gradients in climate is well established (Hawkins et al. 2003b, Field et al. 2005). Temperature and water availability are key mechanisms controlling species richness of the both primary producers (angiosperms and woody plants) and animal taxa (Currie 1991, Hawkins et al. 2003b). In contrast, in the marine realm, the relationship between climate and species richness is more ambiguous. Sea surface temperature and productivity have been suggested to underlie patterns of mollusc and foraminifera richness (Roy et al. 1998, Hunt et al. 2005, Rex et al. 2005), while habitat area has been demonstrated as a key predictor of richness in other groups (e.g. reef fishes and corals Bellwood et al. [2005], Mora & Robertson [2005]). Nevertheless, in both systems, understanding the mechanisms that underlie these relationships, and building appropriate statistical models to describe the relationships are ongoing ecological challenges (Evans et al. 2005, Field et al. 2005).

The positive correlation between species richness and global climate was first noticed as early as the 19th century (Hawkins et al. 2003b) and since then, hundreds of authors have proposed that spatial variation in energy availability controls species richness. There are two major lines of reasoning about how this control occurs. First is a productivity-based argument, whereby increased energy inputs (i.e., resources) facilitate larger populations, which are less vulnerable to extinction (i.e., a "more individuals" hypothesis). Second is a physiological tolerances mechanism, related to ambient energy rather than resource availability, whereby benign tropical

[2] This chapter is in prep as: Kerswell AP and Connolly SR (in prep) Drivers of marine biodiversity gradients: testing predictions with macroalgae. Global Ecology and Biogeography.

climates fulfil the physiological requirements of more species than do extreme temperate or polar climates (Francis and Currie 2003, Hawkins et al. 2003b). The finer details of how the mechanisms operate may differ depending on the taxa involved and the scale of study (Evans et al. 2005) and the different mechanisms are not mutually exclusive. On land the best climatic predictor explains, on average, over 60% of the variance in species richness and the relative importance of productivity versus temperature varies with latitude (Hawkins et al. 2003b). In the marine realm, energy related factors have been shown to explain as little as <10 % of the variation in species richness in some groups (Bellwood et al. 2005) or as much as 90% in others (Roy et al. 1998).

The species-area hypothesis is proposed as a major alternative mechanism to species-energy relationships. Species-area relationships, as proposed for terrestrial systems, are based on the argument that because the tropics form a belt north and south of the equator, they are the largest climatically similar area on the planet (Rosenzweig 1995). Larger areas support more individuals and populations, thereby reducing extinction risk. Larger areas also contain more barriers, which promote allopatric speciation (Rosenzweig 1995, Chown and Gaston 2000). Area is undoubtedly important in determining the number of species present at local to regional scales (Gaston 2000). However, there are few examples of area as major determinant of species richness patterns at a broad geographic scale (Chown and Gaston 2000).

More recently the importance of evaluating species richness patterns in the context of biogeographic null models has been recognized with the advent of mid-domain effect theory (Colwell and Hurtt 1994, Colwell and Lees 2000). Under the mid-domain effect (MDE), species richness gradients arise due to geometric constraints on the size and location of species' geographic ranges within a bounded domain in the absence of environmental factors. The extent

to which MDEs predict species richness patterns varies greatly (see Colwell et al. [2004] for a review). Species with large geographic ranges are geometrically constrained to exhibit strong correlations with MDE predictors, while small ranged taxa have the potential to deviate greatly from the predictions (e.g. Jetz & Rahbek [2002], Mora & Robertson [2005]).

The influence of MDEs relative to environmental predictors of richness patterns may be determined by incorporating MDEs into spatial linear regression models, which include other predictor variables (Colwell et al. 2004). However, few studies have actually done so (Jetz and Rahbek 2002, Bellwood et al. 2005, Mora and Robertson 2005) and there is currently considerable controversy about how MDEs should be incorporated into models of species richness gradients (Zapata et al. 2003, Colwell et al. 2004). Most frequently, species richness expected under a MDE is estimated via randomisation and compared to the observed richness values. However, MDE randomisation models may be formulated in multiple ways and according to different assumptions (see Colwell et al. [2004] for a review). Moreover, each method of modelling MDEs produces a different species richness gradient. For example, 1-dimensional MDEs generated using a "spreading dye" model (Jetz and Rahbek 2001), are typically shallower than those generated under fully stochastic models, due to an increased probability of ranges abutting domain boundaries under the spreading dye model (Connolly 2005). Furthermore, Koleff and Gaston (2001) demonstrated that the height of MDEs may vary up to 4-fold depending on the randomisation procedure and, as a consequence, the fit of the empirical data to the MDE null model varied drastically. Thus, the best way to incorporate MDEs into models of species richness is a significant and unresolved problem.

Recently, Bellwood et al. (2005) developed a spatial regression model that examines the relative influence of MDEs along with energy and area variables on the species richness patterns of

corals and reef fishes in the Indo-Pacific Ocean. The observed species richness patterns of corals and reef fish are strikingly concordant, as are the models which best explain these patterns. Reef area and geometric constraints (i.e., MDE) are the factors which best predict coral and fish diversity patterns. Corals and reef fish have markedly different life histories, yet the similarity of both richness patterns and the concordance of factors that are hypothesised to drive them, prompted Bellwood et al. (2005) to suggest that the same processes may apply more widely across a range of marine taxa.

Here I test the generality of an area-MDE model, by applying the Bellwood et al. (2005) model to richness patterns of tropical marine algae. The order Bryopsidales contains species which have a high affinity with coral reefs and largely tropical distributions (Littler and Littler 2003). Moreover, the observed species richness patterns of Bryopsidales are highly concordant with those of corals and reef fishes throughout the Indo-Pacific Ocean (Chapter 2, Kerswell 2006). Thus they are an ideal group with which to test for the generality of MDEs and area variables as drivers of marine diversity patterns.

3.2 METHODS

The determinants of Bryopsidales species richness patterns were investigated at 140 sites in the tropical Indo-Pacific (30^{o}S-30^{o}N) (Figure 3.1). At each site, the number of Bryopsidales species was determined from an existing global database of algal distribution records (Kerswell 2006). Bellwood et al. (2005) tested a variety of variables in order to determine the best predictors of energy, area and MDE. I also tested several energy-related variables to determine the two best energy predictor variables to include in the combined models. The values of mean annual sea surface temperature, mean annual irradiance, mean annual nitrate concentration and mean annual chlorophyll *a* concentration for the 1^{o}x1^{o} degree grid cell which overlapped each site

were taken from the World Ocean Atlas (Conkright et al. 2002). Area was measured as the total area covered by coral reefs within a 600km radius of each site and log-transformed for use in analyses. Bellwood et al. (2005) digitized reef area from published maps and stored the data in a geographical information system. I extracted area estimates for the Bryopsidales analyses from the same GIS.

Colwell et al. (2004) argue that the influence of MDEs relative to other environmental predictors of species richness gradients may be determined by incorporating MDEs into spatial linear regression models. However, given the differences between individual MDE models outlined above, it is unclear which particular randomisation procedure is the most appropriate to use. Bellwood et al. (2005) offer one solution by recognising that despite the differences in shape, all MDE models display a common feature: a monotonic, bilaterally symmetric, decrease in species richness from the mid-domain towards domain boundaries. Thus, instead of choosing a particular randomisation procedure, these authors include MDEs as a function of the distance of a site from the middle of the domain and constrain the models to also show a monotonic, bilaterally symmetric, decrease in species richness from the mid-domain towards domain boundaries. Consequently, in the Bryopsidales analyses, I adopted a similar approach and modelled MDE as the normalised distance of each site from the mid-domain (0^{o}N, 162^{o}E, Figure 3.1). That is, the latitudinal and longitudinal distances of each site from the mid-domain were normalized to the respective sizes of the domain, and the diagonal distance calculated using Pythagoras' Theorem (see Bellwood et al. [2005] for details).

Formulating an MDE as the distance to the middle of the domain may overestimate the magnitude of its effects on species richness gradients, since no constraints are placed on the height of the MDE gradient. Therefore, to assess whether my formulation of the MDE is

analogous to other methods, I compare the fit of the normalised distance to the mid-domain predictor with species richness gradients generated under several other methods. Species richness expected under the Bokma et al. (2001) MDE model was calculated at each of the 140 sites and plotted separated for latitude and longitude. A 1-dimensional analytical formulation of the spreading dye model was calculated from equation 2 in Connolly (2005) using the observed Bryopsidales range extents. It should be noted that species richness estimated using this 1-D spreading dye model represents richness in latitudinal bands (rather than at discrete sites). Thus the magnitude of spreading dye model richness (i.e., the height of the MDE) will always be greater than the estimates of this study and Bokma et al. (2001) and therefore only the shape of the MDEs are compared. Finally, the 1-dimnesional species richness gradients resulting from the Poisson null model of Connolly (2005, equation 9) were calculated both latitudinally and longitudinally.

Spatial regression models were used to explore the relationship between predictor variables and Bryopsidales richness. There has been increasing concern about the effects of spatial autocorrelation on the power of regression models to account for patterns of species richness (Legendre 1993). In order to overcome these problems I used regression models that explicitly account for the spatial nature of geographic data (implemented using function likfit, package geoR, in the software program R: see Ribeiro & Diggle [2001] for details). Specifically, richness relationships are modelled as Gaussian random fields of the form:

$$y(x) = \mu(x) + S(x) + \varepsilon \qquad (3.1)$$

Where $\mu(x)$ (the mean component) is a linear function of the predictor variables as in a standard linear regression and the residual variation is divided into spatially-autocorrelated $S(x)$ and non-spatial (ε) error components. Thus, $S(x)$ is a function of the spatial distance separating locations

and was modelled according to an exponential decay function. All analyses were performed using both the spatially explicit regression model and a non-spatial equivalent (i.e., a traditional linear regression model).

Collinearity between predictor variables may also be problematic in multiple regression analyses. I quantified the potential effects of collinearity between predictor variables and employed two approaches to mitigate its effects. Model selection procedures are particularly sensitive to collinearity in models when many predictor variables are included (Graham 2003). Consequently, I reduced the number of variables in my models, by choosing the two energy variable that best predicted species richness and using these in the final analyses. Collinearity is also problematic when models are built using stepwise regression procedures, because such procedures are highly sensitive to the order in which predictor variables are added to (or removed from) models (Graham 2003). An alternative is to consider all possible combinations of variables and use model selection techniques, such as Akaike's Information Criterion (AIC), to choose the best model (Neter et al. 1996). Hence, in the final analyses, I analysed all possible combinations of variables, including a no trend or constant model in which species richness patterns were predicted only from spatial autocorrelation. This approach allowed me to choose the model that best predicted species richness patterns using AIC and Akaike weights (wAIC)(Akaike 1985). In all analyses, predictor variables were included in regression models as first order terms only. Second order combinations of variables had consistently worse fit (by AIC) to the empirical data than first order terms and were therefore not included in the final models.

In order to explore the richness patterns predicted by the best linear regression model, predicted richness was determined at each site according to each model. Agreement between predicted

and observed richness was assessed by fit to the unity line. Also, maps of predicted richness throughout the Indo-Pacific Ocean were generated to identify where richness predictions consistent with observed values were estimated from the models. Richness estimates were used to interpolate contours of predicted species richness in a GIS using the Inverse Distance Weighted interpolator (ArcView 3.2a GIS, 0.5^{o} cells, nearest neighbour technique with 12 neighbours, 3^{rd} order power, no barriers). Contour maps of observed species richness were generated using the same procedure.

3.3 RESULTS

The best model for Bryopsidales species richness included three variables: the distance to the mid-domain, sea surface temperature and nitrate. Comparisons between the spatial and non-spatial regression models indicated that spatial regression models, which incorporated spatial autocorrelation, were consistently better at predicting Bryopsidales species richness. As shown in Table 3.1, the AICs for spatial regression models are consistently much lower than those for standard linear regression models using the same predictor variables and no spatial effects. Thus all subsequent analyses were based on the spatial regression models. Of the energy variables considered in the initial regression analyses, sea surface temperature and nitrate had the most support from wAIC (Table 3.2). Hence, sea surface temperature and nitrate were identified as the best energy-related predictors of Bryopsidales species richness and used as the energy predictor variables in the combined analyses.

The combined analyses considered temperature, nitrate, reef area and MDE as predictors of Bryopsidales species richness. Collinearity between these variables was moderate, with Spearman's rank correlation coefficient greatest between temperature and MDE at -0.66 (Figure 3.2). This reinforces the importance of using model selection techniques that are more robust to

the influence of collinearity than stepwise regression procedures. Model selection by AIC indicated that the temperature-nitrate-MDE model was the best model (Table 3.3). The trend component of the this model explained 71% of variation in Bryopsidales species richness. Adding reef area to the model (i.e., the full model) accounted only for an extra 1% of variation. Thus, as annual sea surface temperature and nitrate concentrations increase, and as sites get closer to the mid-domain, species richness also increases, while there is a negligible influence of reef area on Bryopsidales richness.

Comparisons of the various MDE species richness gradients confirm that all MDEs display a monotonic, bilaterally symmetric, decrease in species richness from the mid-domain towards domain boundaries (Figure 3.3). Furthermore, the magnitude (i.e., height) of the MDE from this study is within the range of species richness gradients from other MDE formulations, suggesting that modelling species richness as a function of the distance from the mid-domain is a approximation for mid-domain effects.

Incorporating MDEs as the distance to the mid-domain may have overestimated the influence of MDEs, particularly since MDEs are correlated with other predictor variables (Bellwood et al. 2005). Thus, in order to explore the relative influence of MDEs and energy variables in predictions of Bryopsidales species richness, I compared the observed species richness to that predicted by the temperature-nitrate-MDE model, the temperature-nitrate, temperature-only and nitrate-only models (Figure 3.4). When MDE is removed, the power of the environmental variable-only models to predict Bryopsidales species richness is severely reduced.

Species richness patterns predicted from the best temperature-nitrate-MDE model are presented in Figure 3.1B and capture the major features of observed Bryopsidales richness. For both observed and predicted patterns, species richness is highest in the Indo-Australian Archipelago

(IAA) and Micronesia and decreases both latitudinally and longitudinally towards the domain boundaries. However, there are some noteworthy deviations of the predicted from the observed patterns. Even though richness peaks in the central Indo-Pacific in both the observed and predicted data, the temperature-nitrate-MDE model underestimates the magnitude of this richness peak by approximately 10-15%. Additionally, observed algal richness is high throughout the tropical Indian Ocean, with an equatorial band of high richness extending west from the IAA to the east African coastline where up to 90 species of Bryopsidales occur. This feature is not present in the predicted model. Rather, predicted richness decreases steadily from the IAA, to reach between 30 and 45 species on the east African coastline. In both the tropical east Pacific and Hawaii species richness is over-predicted by the temperature-nitrate-MDE model.

3.4 DISCUSSION

I find that sea surface temperature, nitrate and geometric constraints (i.e., a MDE) have a significant influence in shaping species richness patterns of Bryopsidales algae. The best model of geographic variation in algal species richness has striking similarities to that of Bellwood et al. (2005). Mid-domain effects are important factors in determining richness patterns for algae as well as corals and reef fishes. Additionally, all "best models" predict richness patterns that deviate from the observed richness patterns in consistent ways. However, there is also a fundamental difference between the model developed here for Bryopsidales and the Bellwood et al. (2005) model for coral and reef fishes. I find that energy (i.e., temperature and nitrate) is the primary environmental variable driving algal species richness, whereas for corals and reef fishes, habitat area is of primary importance. The comparison of these models suggest that at the very broadest scale, the geometry of the Indo-Pacific domain is a critical factor in predicting

species richness gradients across a variety of taxa, but that the actual environmental factors which influence species richness patterns may be more taxon specific.

The key determinants of species richness patterns for terrestrial primary producers are energy variables related to temperature and water availability (Francis and Currie 2003, Field et al. 2005). For macroalgae (a major group of marine primary producers), temperature is also influential in determining species richness patterns and indeed, temperature has been suggested to be a major driver of species richness in a variety of other marine taxa (e.g. molluscs: Roy et al. [1998], Rex et al. [2005], foraminifera: Hunt et al. [2005]). Terrestrial productivity and species richness are tightly linked with water availability (Francis and Currie 2003, Hawkins et al. 2003a). Of course in the marine realm water availability is unlikely to be limiting; however, nutrients, and in particular nitrates, are often a key factor limiting marine productivity (Howarth 1988). Thus, like terrestrial systems, both ambient energy (i.e., temperature) and productivity (i.e., nitrate) are important environmental variables for predicting species richness of tropical marine algae.

Temperature has previously been implicated in influencing the distribution of many marine algae (Breeman 1988, Luning 1990, Adey and Steneck 2001) although direct tests at large geographic scales have been limited. Tropical algal species are restricted to temperatures between 18-33^{o}C, with optimal growth at ~27^{o}C and sensitivity at the lower thermal range has been shown to affect the geographic distribution of tropical algae in the Atlantic (Pakker et al. 1995). This suggests a physiological tolerances mechanism for the species-energy hypothesis. Additionally, for the Bryopsidales, there is some evidence that population size varies directly with both temperature fluctuations and increased nitrate concentrations. In the eastern Pacific, the summertime abundance of *Caulpera* increased with abnormally warm temperature in El

Nino years and decreased with abnormally cold temperatures in La Nina years (Scrosati 2001). Also, increased inputs of nitrogen have been shown to increase both the abundance and species richness of algal communities in manipulative field-based experiments (Worm et al. 2002, Bracken and Nielsen 2004). This highlights that a "more individuals" mechanism of the species-energy hypothesis may also apply.

Mid-domain effects are highly controversial, and incorporating them into models of species richness gradients is not as straightforward as is sometimes assumed. Here I show that incorporating MDEs as the distance to the mid-domain approximates both the shape and the height of MDEs derived under both randomisation and processed-based modelling procedures. The relative importance of MDEs in predicting species richness gradients has been questioned, particularly in systems where energy variables appear to highly important (Hawkins and Diniz 2002). Here I demonstrate that when MDEs are removed from the temperature-nitrate-MDE models, the fit of the observed to the predicted species richness is substantially worse. This suggests that MDEs are not acting as a surrogate for environmental variables as has been previously suggested (Zapata et al. 2003), but rather are a key predictor of species richness patterns in their own right.

The "best" temperature-nitrate-MDE model for algae and the area-MDE model for corals and fishes deviate from the observed patterns in similar ways. This suggests that important factors omitted from both these analyses may affect corals, fishes and bryopsidalean algae in similar ways. In particular, the geographic pattern of the deviations suggests a potential role for ocean currents as important determinants of marine richness patterns. Jokiel and Marintelli (1992) developed an ocean currents based dispersal model for reef organisms. Under this model, richness is predicted to accumulate on western boundary margins and be lowest at high latitudes

and in isolated locations. These patterns coincide well with the observed patterns of both algae and corals and reef fish, and could explain the lack of fit of the predictive models in the western Indian Ocean, Hawaii and the east Pacific. Furthermore, corals, fishes and Bryopsidales have very similar patterns in the concentration of geographic range endpoints. For all three groups, range endpoints are disproportionately common at high latitudes and along the east African coastline (Connolly et al. 2003, Kerswell 2006). While the predominance of range endpoints in high latitude locations may be attributed to lack of coral reef habitat (corals and fishes) or unsuitable temperature regimes (algae), the concentration of range endpoints along the African coastline is more likely due to ocean gyres than environmental effects (Connolly et al. 2003, Kerswell 2006).

Despite the similarities, comparison of the algae and corals and fish models also raises the question of why the environmental factors that best correspond to richness patterns differ among taxa. As discussed above, temperature and nitrate appear to have influential effects on both the distribution and population sizes of tropical algae. In comparison with corals and reef fishes, tropical algae are less habitat specific. In addition to colonizing reef environs, algae are also found on tropical rocky shores, in mangrove forests and in seagrass beds (Littler and Littler 2003). Thus while the area of coral reef available for habitation is important for corals and reef fish (as revealed in Bellwood et al. [2005]), energy variables may be more important for habitat generalists such as marine algae.

Climate change and anthropogenic effects are currently impacting ecosystems at an unprecedented rate. Thus it is critical to understand what mechanisms are driving contemporary species richness patterns in order to best manage and conserve diversity into the future. Here I confirm the role of geographic domain boundaries as a major predictor of marine species

richness patterns across a variety of taxa. My analyses also highlight that environmental variables influence marine species richness patterns, but that the relative importance of such environmental variables may differ with the taxa in question. Moreover, it appears that not all of the variation in species richness for tropical taxa in the Indo-Pacific Ocean is explained by mid-domain effects and environmental variables and I suggest that broad-scale factors such as ocean currents may also play a role. This presents the challenge of incorporating new types of variables into analyses of species richness patterns in order to understand the processes that create and maintain diversity across the globe.

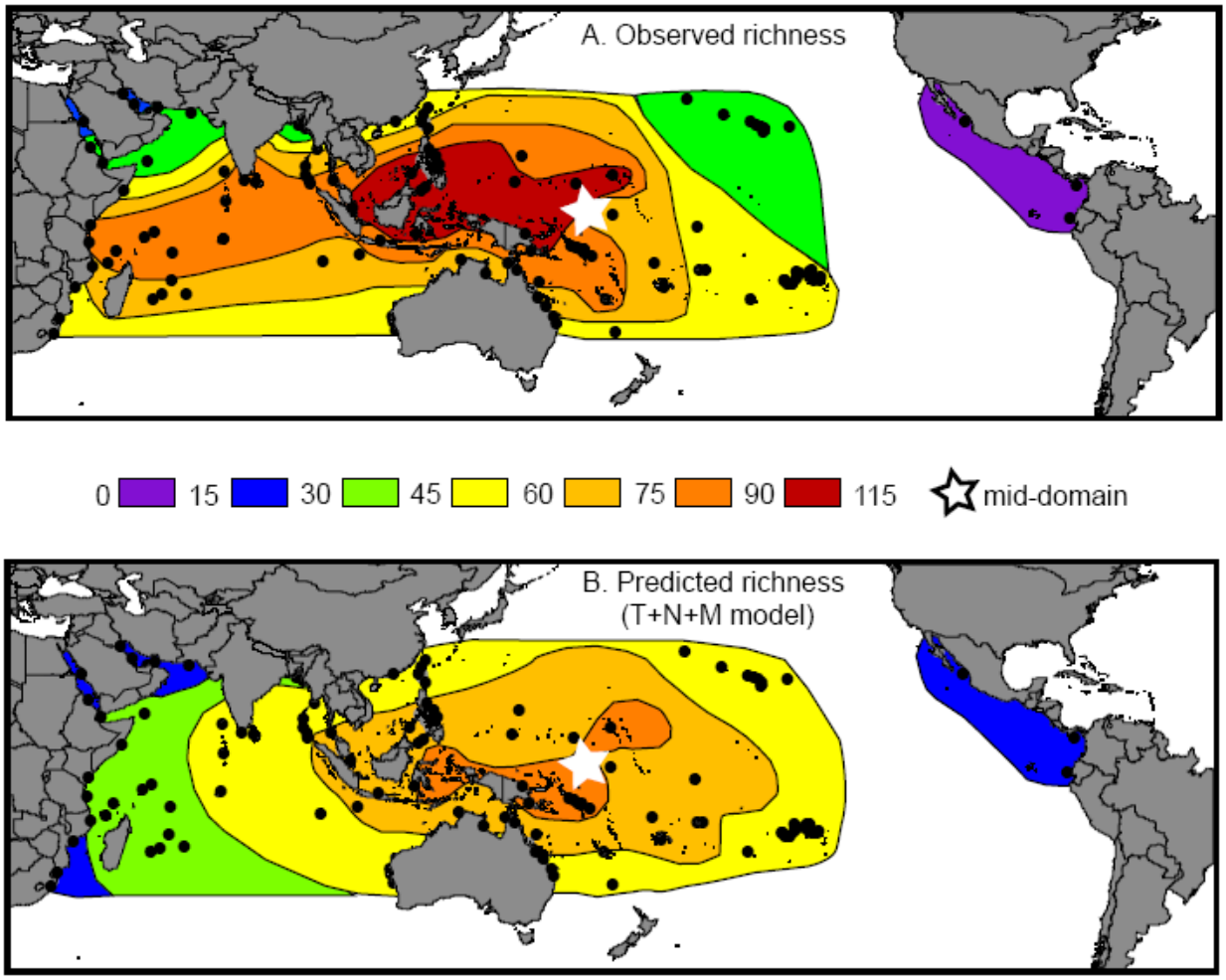

Figure 3.1: Species richness patterns for the order Bryopsidales in the tropical Indo-Pacific Ocean (A. observed richness, B. richness predicted by the best model [temperature-nitrate-MDE]). Black dots represent sampling locations and the white star is the location of the mid-domain.

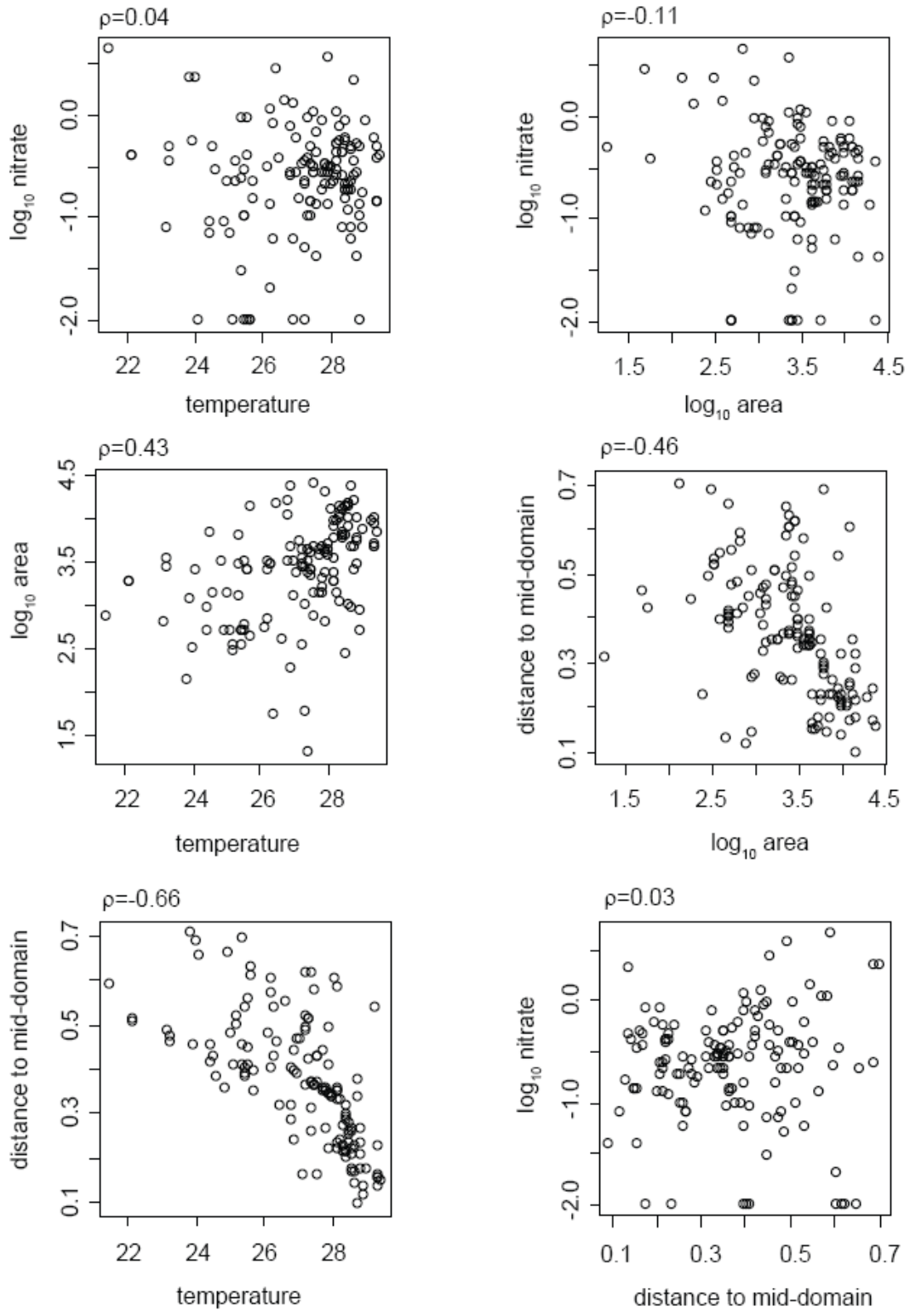

Figure 3.2: Collinearity between temperature, nitrate, reef area and distance to mid-domain (MDE). Spearman's rank correlation (ρ) is indicated for each pair-wise comparison.

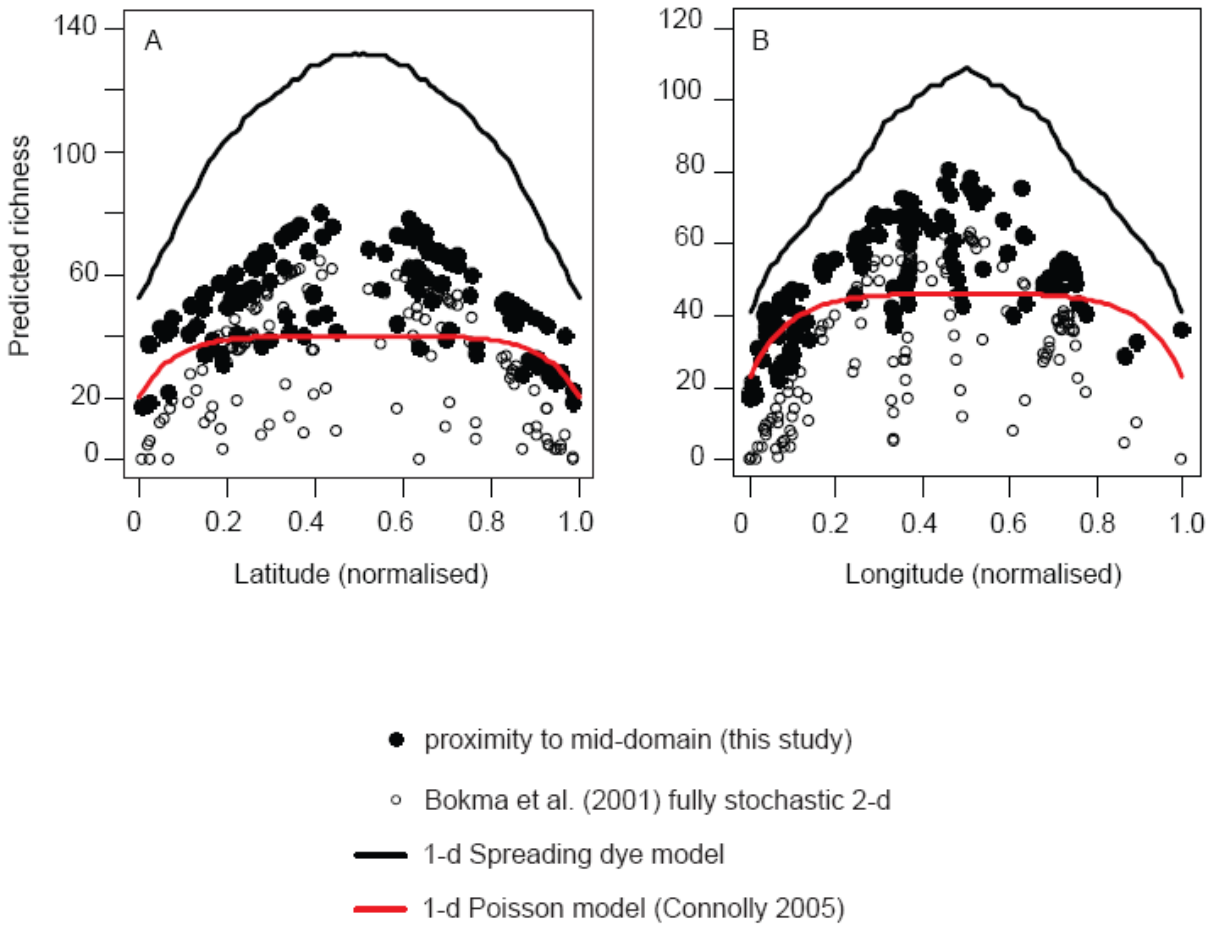

Figure 3.3: Latitudinal (A) and longitudinal (B) gradients in species richness predicted under the various MDE formulations (see legend). Note that in both cases the shape of the MDE derived from the proximity to mid-domain (this study) is consistent with the general shape of MDEs generated via randomisation procedures. Also note that the height of the MDE derived in this study falls within the bounds of other MDE models.

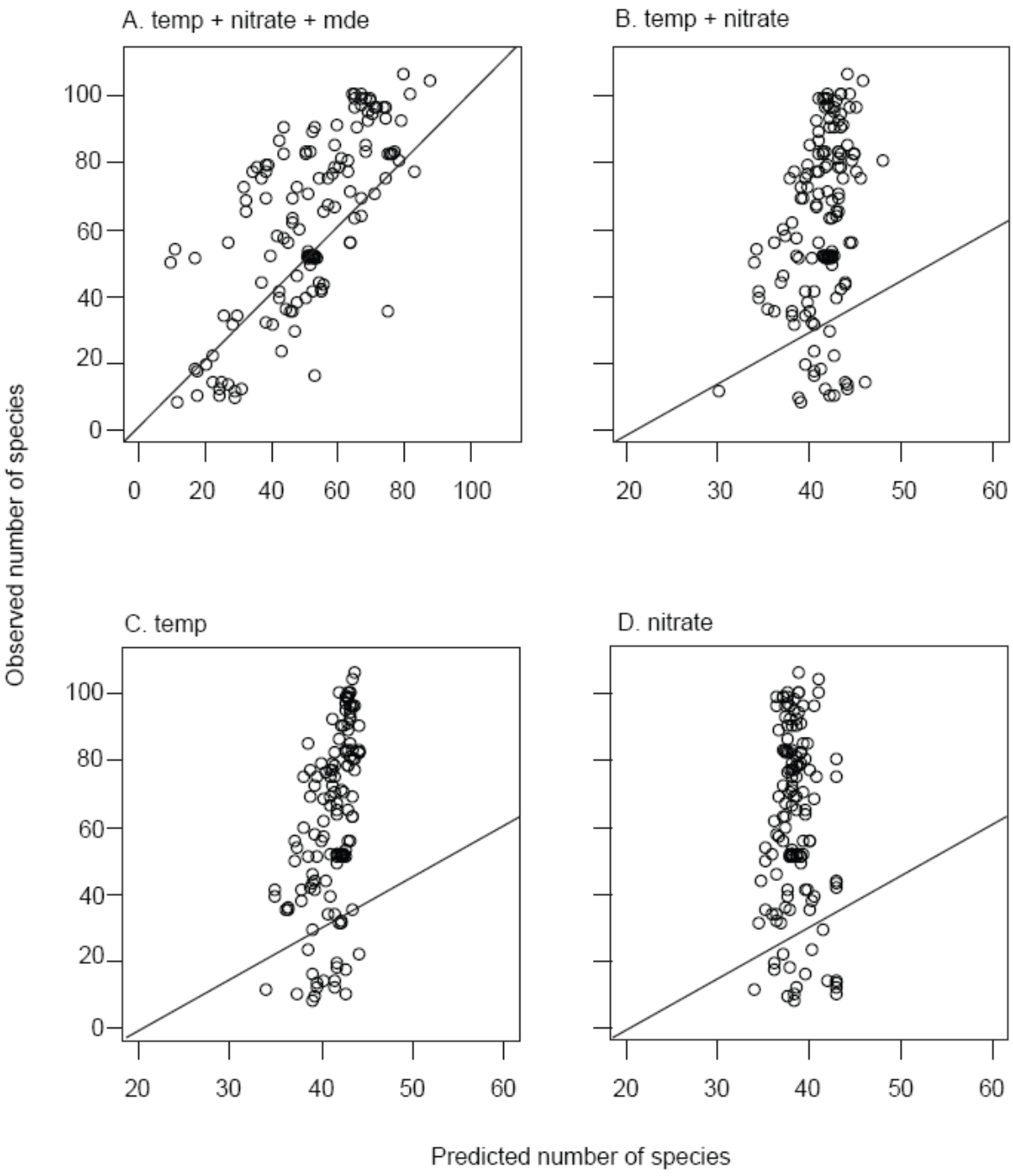

Figure 3.4: Comparison of the observed species richness and that predicted by the temperature-nitrate-MDE model (A) and models including only environmental variables (B-D). Note that removing the MDE predictors from the species richness model substantially reduces the agreement between observed and predicted values.

Table 3.1: AIC from spatial and non-spatial regression models. Note that in all cases spatial models have substantially lower AIC, indicating better support for the spatial regression approach (only results from single variable models are presented but AIC comparisons are analogous for all combinations of predictor variables).

model	AIC	
	non-spatial	spatial
constant	1314	986
temp	1273	986
nitrate	1315	984
area	1296	986
mde	1206	966

Table 3.2: ΔAIC and wAIC (i.e., relative support) for each energy model (np is the number of parameters in the model).

model	ΔAIC	np	wAIC
constant	2.9	4	0.10
temp	0	5	0.42
light	2.0	5	0.09
nitrate	0.6	5	0.32
chl a	5.0	5	0.08

Table 3.3: ΔAIC and wAIC (i.e., relative support) for each model (np is the number of parameters in the model).

models	Δ*AIC*	*np*	*wAIC*
constant	24.8	4	0
temp	25.0	5	0
nitrate	22.4	5	0
area	24.5	5	0
mde	4.9	5	0.03
T+A	25.5	6	0
T+M	3.5	6	0.06
T+N	22.8	6	0
A+M	6.0	6	0.02
A+N	22.7	6	0
M+N	1.8	6	0.14
T+A+M	3.7	7	0.06
T+A+N	23.7	7	0
T+N+M	**0**	**7**	**0.36**
A+M+N	3.3	7	0.07
all	0.7	8	0.26

4.1 INTRODUCTION

A long running controversy in community ecology concerns the extent to which species interactions influence the structure of assemblages. Part of this controversy is centred around the search for "assembly rules". Assembly rules refer to the description of consistent patterns in the number and identity of species seen among replicated sets of communities (Wilson 1999) and mechanisms are subsequently inferred from the presence of these patterns. Most assembly rules aim to identify instances in which competitive species interactions constrain communities to a subset of possible species combinations. Examples of such assembly rules include forbidden species combinations (Diamond 1975), species per genus ratios (genera are represented by only one species at a site even though more may be present in the total species pool (Gotelli and Colwell 2001)), and species nestedness (the composition of species poor sites is an exact subset of species rich sites (Patterson and Atmar 1986)). Assembly rules may also be evident at higher taxonomic or functional guild levels, where the number of species occurring in a particular guild may be limited by strong competition for common resources between species within the same guild (Fox 1987, Wilson 1989).

Two major themes have dominated the study of assembly rules. The first is a description of the rules and the communities in which they are and are not found. This theme will be addressed for macroalgal communities in the next chapter (5). The second theme is that of testing for the statistical significance of the pattern against an appropriate null hypothesis. Here, the focus has been on statistical methods, since it is rarely possible to test assembly rules with small scale manipulative experiments (Gotelli 2004). The validity of assembly rules, therefore, hinges upon being able to distinguish the rules from patterns that might arise by stochastic colonisation and

extinction of species independent of any constraints imposed by the operation of assembly rules. This is the focus of the present chapter.

The first assembly rules were proposed by Diamond (1975) and consisted of seven observations or "rules" relating to bird community structure that could be applied across the Bismarck Archipelago (Papua New Guinea). These rules were subsequently challenged by Connor and Simberloff (1979), who recognised the importance of comparing the observed patterns to an appropriate null expectation. For assembly rules the null expectation is that species are assembled into communities independently of one another and thus the structure of null model communities reflects a random sample of the larger species pool. Consequently, null model communities are often termed "random" communities and observed communities are compared to a null expectation of "randomness". Such terminology is used in the remainder of this chapter. Connor and Simberloff (1979) highlighted that many of the patterns, which Diamond was proposing as rules, could arise simply through independent colonisation of islands by species. This work was the first to consider community structure relative to a null expectation of randomness and sparked an on-going debate about the precise features of observed data that should be incorporated into null models in order to identify assembly rules in a robust manner (see Gotelli and Graves [1996] for a summary and chapters in Weiher and Keddy [1999] as examples of the debate).

Assembly rules are based on the presence or absence of species at sites, recorded as zeros or ones in a species-by-site incidence matrix. In order to create null model communities, species' presences and absences are shuffled between incidence matrix cells, according to algorithms proposed to mimic species entering communities independently of one another (the resulting null model communities are presented in "randomised incidence matrices"). In order to

formulate robust null models of community structure, certain features of the observed data should be maintained in randomised incidence matrices.

4.1.1 Guild proportionality assembly rules

Guild proportionality rules were initially proposed as tests of the influence of interspecific competition on community structure (Wilson 1989). The link between guild proportionality and competition is typically thought to be niche limitation. If the number of resources is limited, then some species will competitively exclude others within the same guild and as a result the proportion of species in a community which belong to any one guild will be limited (Wilson 1989). Guilds may refer to functional groups or higher taxonomic levels e.g. families and orders (Wilson 1999). While guild proportionality is seldom identified in natural communities (see Wilson [1999] for examples), the concept has considerable appeal for providing a straightforward assembly rule (Weiher et al. 1998).

To test for guild proportionality, species incidences are summed across guilds and tests are performed on guild-incidence matrices (guilds-by-site matrices with cell values equal to the number of species in a guild at a site). Traditionally, to test guild proportionality, the proportion of species within each guild is calculated at all sites for both observed and null model (random) communities. Then, taking each guild separately, the variance in the proportions over the sites is calculated. Guild proportionality is inferred if the variance across observed communities is significantly less than that of the random communities. The significance level of this test is determined as the proportion of randomisations whose variance was greater than the observed variance. This is effectively a one-tailed test of whether the occurrence of guilds is more similar across sites than expected under the null model.

There are two major disadvantages of Wilson's test for guild proportionality. Firstly, it does not test for all possible outcomes of comparison between observed and randomised matrices. Using Wilson's 1-tailed test considers whether guild proportions are more similar than in random matrices (i.e., guild proportions are highly homogeneous across sites) or that observed guild proportions are not different from those of random matrices. However, guild proportions may also be more variable than those of the random matrices (i.e., highly heterogeneous across sites) and determining this requires a 2-tailed test. Secondly, Wilson's method tests individual guilds separately, which can lead to ambiguity in interpreting community-wide patterns if some guilds show significant proportionality and others do not. Furthermore, there may be statistical limitations when testing for guild proportionality on a guild-by-guild basis. For example, in wetland communities, Weiher et al. (1998) could not reject the null (random assemblage) hypothesis due to issues associated with multiple hypothesis testing, even though there was a distinct tendency towards guild proportionality across several functional groups. An aggregate test for guild proportionality for assemblages as a whole (as opposed to multiply guild-by-guild tests) would add substantial statistical power to, and reduce ambiguity associated with the interpretation of, guild proportionality tests.

Schluter (1990) developed a method called species-for-species matching which is essentially a test of guild proportionality across a whole community. The species matching method applies a 2-tailed chi-squared test to assess homogeneity in guilds' frequencies of occurrence across sites (i.e., guild proportionality). Under this method, the degree of randomness in community structure is assessed through the position of the observed test statistic relative to its expected (χ^2) distribution. Statistics that fall in the middle of the distribution (i.e., $0.025<p<0.975$) indicate that variation in assembly composition is not significantly different from a random assortment of taxa available in the total species pool (i.e., random structure). Statistics that fall in the extreme

right tail of the distribution (>97.5th percentile of the χ^2 cumulative distribution function) indicate assemblage composition that is more heterogeneous across sites than expected under a random assortment. Finally, statistics that fall in the extreme left tail of the distribution (<2.5th percentile of cumulative distribution function) indicate highly homogeneous assemblage structure across sites (i.e., significant guild proportionality).

The aim of this chapter is to develop a test for guild proportionality that can be applied across entire communities. Firstly, I outline key features of the observed data that, I argue, should be maintained in randomised matrices. Then, several possible approaches to testing for guild proportionality rules in entire communities are assessed in the context of these four criteria. For a test to be a good test for guild proportionality, it should conform to all four of these criteria, as well as accurately predicting community structure. Since none of the existing approaches to testing guild proportionality meet all of these requirements, I propose a new approach in which randomisations are based on Bernoulli trials as a robust test for guild proportionality in entire communities. Throughout I use a hypothetical dataset to illustrate some of the problems associated with existing approaches to testing guild proportionality in entire communities.

4.2 CRITERIA FOR GOOD NULL MODELS OF GUILD PROPORTIONALITY

In order to formulate robust null models of community structure, certain features of the observed data should be maintained in randomised matrices. The four criteria presented below are key features of the observed data that should be maintained in null models of guild proportionality.

Criterion 1: Sites must contain, on average, the same number of species in random matrices as in the observed data.

Criterion 2: Species must be sampled from the species pool with a probability equal to their frequency of occurrence in the pool.

Not all sites in nature have the same species richness, nor do all species occur with the same frequency of occurrence within a region. If such features are not included in null models, it becomes unclear whether deviations of the observed data from the null model are a result of species interactions, or whether the deviations are simply a result of species occurring at sites with a different frequency in the observed data versus the null models (Wilson 1995, Stone et al. 1996). Moreover, when species are allowed to occur in the null model communities with equal frequency of occurrence (rather than at their observed frequencies), Type I errors are common i.e., assembly rules are found in randomly structured communities (Wilson 1995, Stone et al. 1996). One way of tackling these problems is to exactly maintain the observed site richness and number of species occurrences (i.e., the row and column totals of the observed incidence matrix) in the null models (Connor and Simberloff 1979, Stone et al. 1996). However, this approach has been criticised as being too severe since it is prone to Type II errors i.e., failing to identify assembly rules when they are present (Gotelli and Graves 1996). Instead, either the row or column totals, or both, may be allowed to fluctuate stochastically from one simulation to another, so long as the mean values across all simulations are equal to the observed values. Such an approach has been criticised by some authors as introducing too much flexibility into the null model (e.g. Stone et al. 1996). However, if each randomisation is thought of as a replicate community, then we should not expect row and column totals to remain fixed in each replicate community, since stochastic variation would allow these totals to vary around a mean value. Species richness and the number of sites as which a species is found are not fixed quantities in nature. They will vary according to resource availability and environmental conditions, and so there is no reason to fix site richness in the randomisation trials (Gotelli and Graves 1996).

Criterion 3: Guilds must have, on average, the same frequency of occurrence in random matrices as in the observed data.

Criterion 4: The maximum number of species in a guild at a site cannot exceed the total number of species belonging to that guild in the total species pool.

These constraints are specific to guild proportionality tests, for which species occurrences are aggregated into guilds. Just as it is important to maintain, on average, the observed frequency of occurrence of species, it is also important to maintain the observed frequency of occurrence of guilds in random matrices. This allows some guilds to be widespread and others to be range-restricted, as is observed in nature. Furthermore, since the basis of guild proportionality rules is that competition between species in the same guild constrains community structure, random guilds should contain no more species than belong to that guild in the total species pool.

Below I consider existing approaches for testing guild proportionality in the context of the four criteria for a good null model for guild proportionality tests. Most tests violate at least one of the above criteria and thus, I develop a new test for guild proportionality that fulfils all four of the above criteria for a good null model. All analyses were conducted in $R_{2.3.0}$ (R Core Development Team [2004]) and randomisations were performed 1000 times.

4.3 ILLUSTRATIVE DATASET

Throughout the chapter I use a hypothetical dataset to illustrate some of the problems associated with various tests for guild proportionality in entire communities. The hypothetical species-incidence matrix (Table 4.1), presents the presence/absence of 20 species (a1-d7) belonging to 4 guilds (A-D) at 6 sites. Data were generated for each site by conducting a series of 20 Bernoulli trials to determine presences and absences of species (labelled a1 to d7, as per row 2 of Table

4.1) at sites. In order to generate sites that differed in their species richness, data for sites 1-3 were generated with a different probability of a species being present at a site to sites 4-6 (p=0.3 for sites 1-3 and p=0.8 for sites 4-6) (Table 4.1). In order to test for guild proportionality, the occurrences of each guild at each site were totalled and presented as a guild-incidence matrix (Table 4.2). For both the species- and the guild-incidence matrices, the row totals represent the total species richness of a site and the column totals represent the total number of occurrences of each species or guild. It is important to note that these data were randomly assembled and as such should show no evidence for significant homogeneity or heterogeneity in any guild proportionality test.

Table 4.1: Species-incidence matrix

	Guild A			Guild B				Guild C						Guild D							
	a1	a2	a3	b1	b2	b3	b4	c1	c2	c3	c4	c5	c6	d1	d2	d3	d4	d5	d6	d7	richness
Site 1	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	1	5
Site 2	1	0	0	1	0	0	0	1	0	1	1	0	1	0	0	0	0	1	0	0	7
Site 3	0	1	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	1	1	1	7
Site 4	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	17
Site 5	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	17
Site 6	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	18
occurrence	3	3	3	4	3	3	3	5	3	4	4	3	6	3	2	2	4	5	3	5	71

Table 4.2: Guild-incidence matrix

	A	B	C	D	richness
Site 1	1	1	1	2	5
Site 2	1	1	4	1	7
Site 3	1	1	2	3	7
Site 4	2	4	6	5	17
Site 5	3	2	6	6	17
Site 6	1	4	6	7	18
occurrence	9	13	25	24	71

4.4 WILSON (1989) GUILD-BY-GUILD TESTS

The hypothetical data in Tables 4.1 and 4.2 were first tested for guild proportionality on a guild-by-guild basis using the method of Wilson (1989). For this procedure the total species pool is

represented by all species a1 to d7. Once a species has been sampled at a site, it is removed from the species pool i.e., sampling occurs without replacement. Each species' frequency of occurrence is equal to the number of occurrence of that species across all sites divided by the total number of occurrence of all species in the species pool (e.g. for species a1: 3/71 = 0.042). Species are sampled from the species pool with a probability equal to their frequency of occurrence. Thus, randomised guild-incidence matrices were generated by randomly sampling species from the total species pool (without replacement) and with a probability equal to their overall frequency of occurrence in the total species pool. Species were randomly allocated to sites until the site reached the observed richness level. The total count of species in each guild was recorded at each site in random species-incidence matrices, to produce randomised guild-incidence matrices. This randomisation procedure allows the column totals in randomised matrices to vary stochastically, with an expected (i.e., mean) value equal to the observed. Observed row totals (i.e., observed site richness) are preserved in random matrices.

For each site, in both the observed and the random communities, the number of species in each guild was expressed as a proportion of the total richness of the site. For each guild, the variance among sites in this proportional richness was calculated for both the observed data (Var_{obs}) and for each of the random matrices (Var_{rand}). The ratio of Var_{obs} to Var_{rand} was then calculated separately for each random matrix. Guild proportionality is evident in a guild if Var_{obs} is significantly less than Var_{rand}. Significance is determined based on the proportion of random matrices for which $Var_{obs}<Var_{rand}$. If this proportion is less than α (in this case $\alpha=0.05$), then significant evidence of guild proportionality is inferred.

There is no evidence of guild proportionality in the hypothetical dataset above. For each guild A-D, the observed variance in guild proportions falls within the 95% CI of Var_{rand} (Figure 4.1,

α_A=0.12, α_B=0.12, α_C=0.64, α_D=0.50). Thus the variance in all guild proportions cannot be distinguished from random, consistent with the fact that the data for these communities were generated as a random sample from a series of Bernoulli trials.

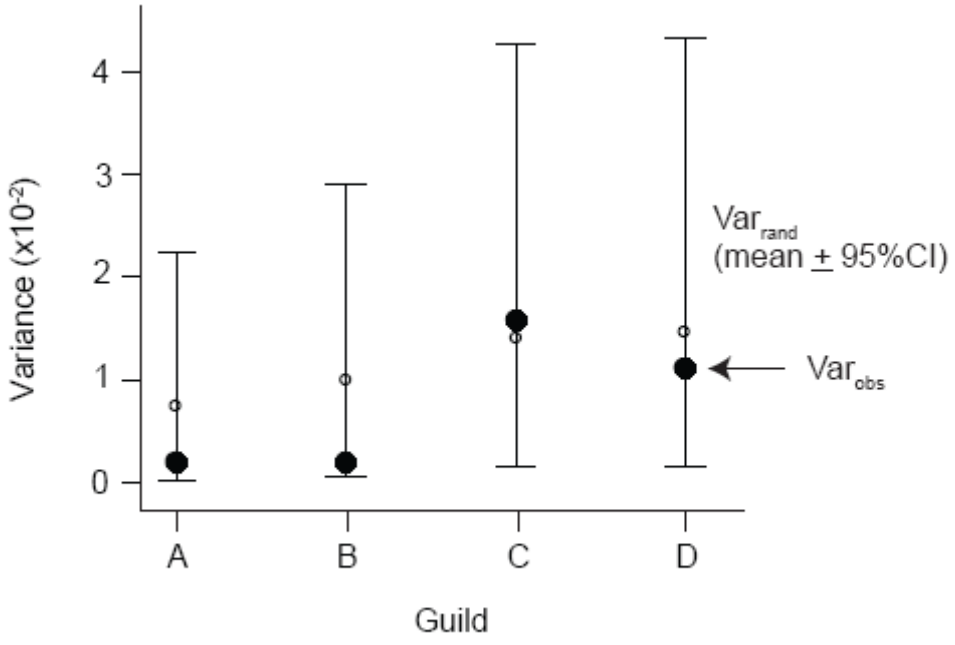

Figure 4.1: Wilson (1989) guild-by-guild test for guild proportionality in hypothetical dataset from Table 4.1. Small open circles are the mean variance in guild proportion across 1000 random matrices ± 95%CI. Solid circles are the observed variance in guild proportions.

4.5 SPECIES MATCHING TEST (SCHLUTER 1990)

The species matching test was proposed by Schluter (1990) as a test for guild proportionality in entire communities. For this test, a chi-squared test for homogeneity is performed on the observed guild-incidence matrix (e.g. Table 4.2) and the location of the observed test statistic (X^2) compared to the chi-squared density function with the appropriate degrees of freedom (calculated as (number of sites-1)×(number of guilds-1)). In order to assess the usefulness of this approach as a test for guild proportionality in entire communities, I consider the

assumptions underlying a chi-squared test for homogeneity and compare them against the four criteria necessary for a good null model.

The chi-squared test compares observed frequencies (i.e the number of species in a guild at a site) with the frequencies expected if there were no relationship between variables in the larger population (i.e. if species occur at sites independently of each other). Under the null hypothesis, the expected frequency in each cell of the guild-incidence matrix is calculated from the observed data by multiplying the row total (i.e. the site richness) by the column total (i.e. the guild's total number of occurrences) and dividing this value by the total number of observations. Under such a formulation, the row and column totals of the expected (null model) guild-incidence matrix are exactly equal to those of the observed guild-incidence matrix. Consequently, criteria 1 and 3, that the observed site richness and the observed frequency of occurrence of guilds, are maintained by the species matching test. This is illustrated with the hypothetical dataset. Table 4.3 presents the expected values from a chi-squared test on the guild-incidence matrix in Table 4.2. Note that the row and column totals of the expected (null model) guild-incidence matrix are equal to the observed values presented in Table 4.2.

Table 4.3: Expected values from a chi-squared test on the observed guild-incidence matrix

	A	B	C	D	richness
Site 1	0.6	0.9	1.8	1.7	5
Site 2	0.9	1.3	2.5	2.4	7
Site 3	0.9	1.3	2.5	2.4	7
Site 4	2.2	3.1	6.0	5.7	17
Site 5	2.2	3.1	6.0	5.7	17
Site 6	2.3	3.3	6.3	6.1	18
occurrence	9	13	25	24	71

To assess criteria 2 and 4, that species are sampled from the species pool with probability equal to the observed frequency of occurrence and that guilds contain no more species in random

matrices than in the observed species pool, we must consider how performing a chi-squared test translates into sampling species from a species pool. Performing a chi-squared test is equivalent to sampling from a multinomial distribution (Biehl and Matthews 1984). Under such a sampling regime, species are drawn from the species pool with replacement. Thus the probability that a species is drawn from the species pool can be set at the species' observed frequency of occurrence, and criterion 2 for a good null model is met. However, sampling species with replacement means that a species may be allocated to a single site more than once. Consequently, when the number of species in each guild is aggregated into a guild-incidence matrix after randomisation, the number of species in a randomised guild at any one site may exceed the total number of species belonging to the guild in the whole region. This clearly violates criterion 4, that guilds contain no more species in random matrices than in the observed species pool. The consequence of this for guild proportionality tests is that there is an over-representation of the variability in the number of species in guilds in random community matrices and thus, random communities are skewed towards being overly heterogeneous.

Problems associated with sampling species with replacement from the total species pool are illustrated below using the hypothetical dataset. Random species-incidence matrices were generated on a site-by-site basis by sampling species from the total species pool with replacement. Species were sampled with a probability equal to their frequency of occurrence in the total species pool and sites were allocated their observed number of species. Species incidence matrices were summarised into guild-incidence matrices. For each guild, I summed the number of times the random guild contained more species than belong to that guild in the entire species pool (i.e., how often criterion 4 was violated across the 1000 randomisation trials). Under the species matching (sampling with replacement) protocol the maximum number of species in a guild was exceeded at at least one site in 42% of trials for guild A, 48% for guild B,

79% for guild C and 50% for guild D. This error is also evident in the expected frequencies from the chi-squared test, which may also be greater than the maximum number of species in a guild (see Table 4.3 guild C at site 6 has expected frequency of 6.3 species, yet guild C only contains 6 species in the observed data).

In order to assess the effects of allowing guilds to contain more species than they posses in the species pool, chi-squared (X^2) statistics were calculated for each random guild-incidence matrix and combined across all randomisations into a frequency distribution of test statistics. The frequency distributions of X^2 statistics were compared to the χ^2 distribution (df=15). Figure 4.2 illustrates that the frequency distribution of test statistics is well characterised by the χ^2 distribution. However, both erroneously suggest highly homogeneous community structure. This highlights how, under a species matching approach, random communities are skewed towards being overly heterogeneous and significant guild proportionality may be falsely inferred.

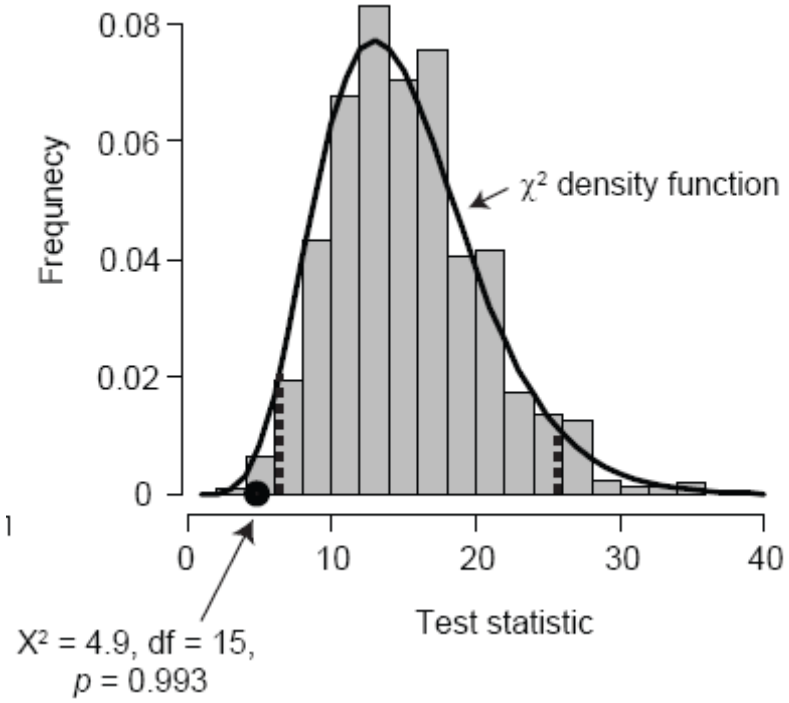

Figure 4.2: Frequency distribution of test statistics from guild-incidence matrices generated by sampling with replacement (grey bars). Note that this distribution corresponds with the χ^2 distribution. The solid circle is the observed (X^2) test statistic and the dashed lines represent the 2.5th and 97.5th percentiles of the cumulative density function. Note that the observed statistic is more extreme than the 2.5th percentile of the cumulative distribution function (p=0.993), erroneously indicating significant guild proportionality.

4.6 MODIFIED SPECIES-MATCHING APPROACH

A potentially robust test for guild proportionality in entire communities would result if the species matching method can be modified in such a way so that random guilds are not allowed to contain more species than are present in that guild in the total species pool, whilst still adhering to the other three criteria for a good null model. Since the problem with the species matching method was that species are sampled with replacement, a modified species matching approach can be developed in which species are sampled *without* replacement from the total species pool. This method effectively samples species from the species pool as recommend by

Wilson (1989, 1995) in his guild-by-guild test, but then combines the randomised data into an entire communities test.

Random matrices can be constructed by sampling species from the total pool without replacement and with a probability equal to their frequency of occurrence in the total species pool as suggested by Wilson (1989). Then to obtain a null distribution of test statistics for the structure of entire random communities, X^2 statistics for each random matrix are calculated and presented as a frequency distribution. This null distribution, rather than the standard χ^2 distribution, can be compared to the observed X^2 test statistic. If the value of the observed statistic is less than the value of the 2.5^{th} percentile of the null distribution, then there is significant homogeneity in the community structure (i.e., guild proportionality). If the value of the observed statistic is greater than the value of the 97.5^{th} percentile of the null distribution, then there is significant heterogeneity in the community structure. Finally, if the observed statistic is between the 2.5^{th} and the 97.5^{th} percentile of the null distribution, the community structure cannot be distinguished from random.

However, this test requires an additional modification. Performing a chi-squared test on the observed and randomised guild-incidence matrices assigns expected frequencies for each matrix that are calculated by multiplying the row total by the column total and dividing this value by the total number of observations (as in a standard chi-squared test). However, as we have seen above, calculating the expected frequencies in such a manner allows guilds to have more species in random matrices than in the total species pool. Therefore, we need to adjust the expected values used to calculate the X^2 statistics so that they accurately reflect the number of species in guilds. This can be done by taking the means for each cell (across all randomisations) as more appropriate expected values. That is, assign the expected value for each cell in the guild-

incidence matrix as the mean cell value over all randomisation trials. Then, in order to calculate the test statistics for both the observed and the random matrices (X^2), a chi-squared statistic can be calculated manually using the mean frequencies from randomisation as the expected values i.e.

$$X^2 = \sum_{i=1}^{G} \sum_{j=1}^{L} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \qquad (4.1)$$

where *G* is a guild and *L* is a site (i.e., location).

This procedure was applied to the hypothetical dataset. In no instances did the number of species in a random guild exceed the number of species contained in that guild in the total species pool. Furthermore, when the observed X^2 statistic was compared to the appropriate null distribution, it fell between the 2.5th and the 97.5th percentile of the null distribution (at the 18.2th percentile), accurately representing the random structure of the hypothetical dataset (Figure 4.3).

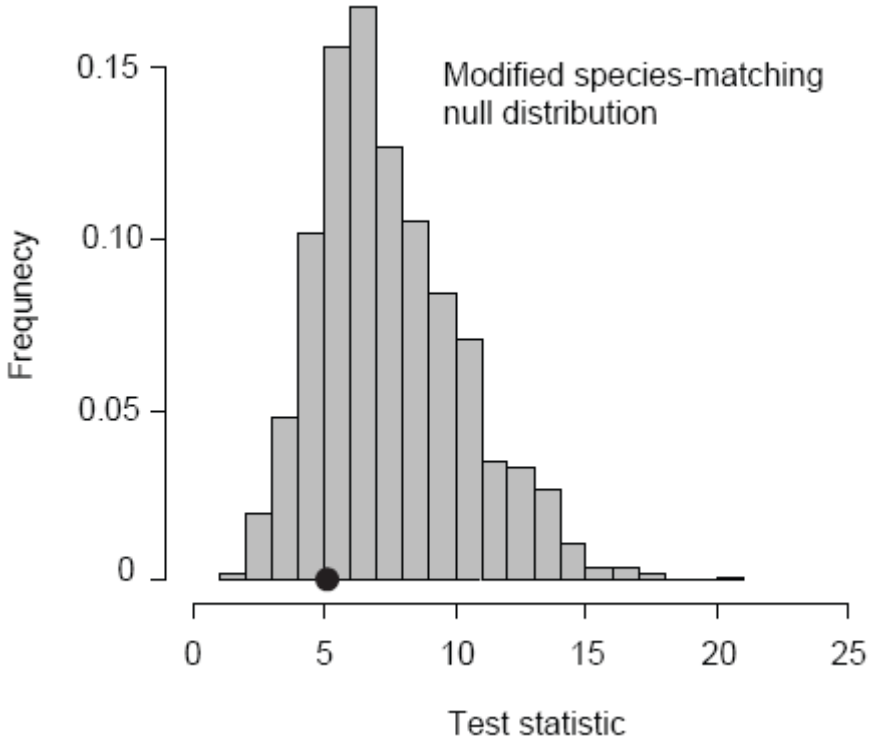

Figure 4.3: Frequency distribution of test statistics from guild-incidence matrices generated by sampling without replacement as per modified species-matching procedure (grey bars). The solid circle is the observed (X^2) test statistic, which accurately falls within the random part of the null distribution.

The modified species-matching method has solved the problem of random guilds containing too many species and thus, as a first approximation it fulfils all the criteria for a good null model. However, in modifying the species matching approach, we have also changed how species are sampled from the larger species pool (i.e., species are now sampled without replacement). It must, therefore, be confirmed that the species are still being sampled with a frequency of occurrence equal to that of the observed data (i.e. fulfilling criterion 2 for a good null model).

Wilson (1989, 1995) emphasises the importance of sampling species according to their frequency of occurrence in the total species pool. This allows some species to be widespread and others to be rare in random matrices and is a key constraint of randomised matrices.

However, sampling species without replacement does not exactly preserve this constraint. Instead, for each randomisation trial, the probability with which a species is drawn from the species pool depends on the order in which species are drawn. For example, take the following very simple incidence matrix:

	spp a	spp b	spp c	richness
site 1	1	0	1	2
site 2	1	1	0	2
site 3	1	0	1	2
occurrence	3	1	2	6

The total species pool from which to sample is {a,a,a,b,c,c} and a species can only be allocated once to a site. So for example, if spp a is drawn first, then the remaining species pool from which to sample is {b,c,c}. To see how the probability of drawing a particular species depends on the order in which species are drawn, let us consider drawing species a and b from the species pool:

$$\Pr(a) \times \Pr(b \mid a) = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6} \tag{4.2}$$

$$\Pr(b) \times \Pr(a \mid b) = \frac{1}{6} \times \frac{3}{5} = \frac{1}{10} \tag{4.3}$$

The frequency of occurrence of species a in the total pool is 1/2, while that of species b is 1/6. However, by drawing species a first (equation 4.2), the probability of species b being drawn next becomes 1/3 (not 1/6) and by drawing species b first (equation 4.3), the probability of species a being drawn next becomes 3/5 (not 1/2). Hence, the overall probability of drawing one species a and one species b is different depending on which order the species are sampled. However, in guild proportionality tests we need to know the probability with which each species occurs in the total species pool and verify that it is the same as the species' observed frequency

of occurrence. This probability can be calculated by averaging the individual probabilities of a species over all the possible sampling orders.

Total Pr(a) = Pr(a first) + Pr(b first)×Pr(a second) + Pr(c first)×Pr(a second)

$$\text{Total } \Pr(a) = \frac{1}{2} + (\frac{1}{6} \times \frac{3}{5}) + (\frac{2}{6} \times \frac{3}{4})$$

$$\text{Total } \Pr(a) = \frac{17}{20}$$

However, the expected frequency of species a in the total species pool is 3/4 (i.e., Pr(drawing a first = 1/2) + Pr(not drawing a first and drawing a second = 1/2×1/2)), which is not equal to 17/20. This highlights that when sampling occurs without replacement from the total species pool, species may be sampled with a probability different to their overall frequency of occurrence. One complication of this is that once species occurrences are totalled across the randomised species incidence matrix, the values may not match the observed values (i.e., the column totals are do not match between observed and randomised matrices). Continuing with the simple example;

The expected frequency with which species a will occur in the randomised matrix (i.e., the expected column total of the randomised matrix for species a) is equal to sum of the probabilities of it being sampled at no sites, one site, two sites and all sites:

$$E(\Theta(a)) = \sum_{i=0}^{3} ip_i$$

$$= \left(\frac{17}{20} \times \left(\frac{3}{20}\right)^2 \times 3\right) + \left(2 \times \left(\frac{17}{20}\right)^2 \times \frac{3}{20} \times 3\right) + \left(3 \times \left(\frac{17}{20}\right)^3\right)$$

$= 2.55$

Since the observed value is 3, we can see that the column totals of the observed incidence matrix may not be equal to the expected column total of the random incidence matrix. This inequality may also affect the column totals of guild-incidence matrices since a guild's frequency of occurrence is calculated as the sum of all the occurrences for the species that belong to that guild. This problem is illustrated using the illustrative dataset (Tables 4.1 and 4.2). The species incidence matrix (Table 4.1) was randomised by sampling species with replacement and then combined into guild-incidence matrices. The expected frequency of each guild is then equal to the mean cell values in the guild-incidence matrix (over all randomisations). As is illustrated in Table 4.4, the expected values do not match the observed. Thus both the second and third criterion for a good null model (preserving observed species and guild occurrences) have been violated and the modified species matching approach does not appear to be a robust test for guild proportionality in entire communities. As a potential alternative I explore a method for randomising species-incidence matrices developed by Stone et al. (1996, 2000).

Table 4.4: Mean (expected) cell frequencies from 1000 random guild incidence matrices using the modified species matching method. Note that expected number of occurrences from the random matrices are not equal to the observed values, i.e., compare column totals.

	A	**B**	**C**	**D**
Site 1	0.7	0.9	1.7	1.7
Site 2	1.0	1.3	2.4	2.3
Site 3	0.9	1.3	2.4	2.4
Site 4	2.5	3.4	5.3	5.8
Site 5	2.5	3.4	5.3	5.8
Site 6	2.7	3.6	5.6	6.2
total occurrence (expected)	10	14	23	24
observed	**9**	**13**	**25**	**24**

4.7 GUILD RANDOMISATION METHOD (STONE ET AL. [1996, 2000])

Stone et al. (1996, 2000) present a method for randomising species presence-absence data without changing the observed species-incidence matrix. Under this approach, species are randomly allocated to guilds with equal probability, whilst remaining on exactly the same sites as in the observed data. Thus, the guilds in random matrices have the same total number of species as in the observed data, but the identity of those species changes with each randomisation trial. The authors presented the guild randomisation method as a means of maintaining species' geographic ranges in random matrices. (Under traditional species-incidence randomisation methods, species are constrained to occur at the same number of sites as in the observed data, but those sites may occur well outside the geographic range of the species).

The guild randomisation approach may also be developed into a guild proportionality test in entire communities. Random matrices are created by swapping the groups to which species belong, while maintaining the observed species-incidence matrix. The random species-incidence matrices are then totalled into guild-incidence matrices and an X^2 statistic calculated for each. These X^2 statistics are then combined into a (null) frequency distribution for the guild randomisation method. Like the modified species matching test, using the expected values from a chi-squared test produces inaccurate numbers of species in guilds. Therefore corrected test statistics must be calculated from equation 4.1 with *Eij* equal to the expected frequencies from the guild randomisation method.

The advantage of using the guild randomisation method to create null models is that the observed species incidence matrix is not altered in the randomisation trials. Consequently, both the observed site richness and the observed frequency of occurrence of species are maintained in

the random data. Furthermore, while species are randomly allocated to guilds, the observed number of species in each guild is maintained at the observed value. Thus three of the four criteria for a good null model have been fulfilled.

The fourth criterion for a good null model, that guilds' frequencies of occurrence must, on average, be equal to the observed value, is not fulfilled by the guild randomisation test. Different species have different numbers of occurrence and each time a species is allocated to a new guild, those occurrences contribute to that guild's total occurrences. Since the allocation of species to guilds changes with each randomisation trial, the total number of occurrences for that guild also changes. Consequently, when the total number of occurrences of each guild is averaged over all randomisation trials, the average number of occurrences is not equal to the observed values (compare rows 7 and 9 of Table 4.5). Therefore, while the guild randomisation method takes great care to maintain the observed frequency of occurrence of species, it does not maintain the observed frequency of occurrence of guilds. The severity of this problem will depend on differences in the average number of species occurrences in each guild. For example, swapping guilds for species that have similar numbers of occurrence will have a small effect on the average number, because all guilds are similar to begin with. However, if different guilds have very different numbers of occurrence, swapping the guilds to which species belong will have a large effect of the average.

To illustrate the discrepancies in the frequency of occurrence of guilds in observed and randomised matrices, I performed the guild randomisation procedure on the hypothetical dataset. Table 4.5 presents the mean frequency of occurrence of each guild at each site (i.e., guild-incidence matrix cells averaged over all randomisation trials). A comparison of the observed

and expected frequency of occurrence of each guild reveals that the third criterion for a good null model is violated by the guild randomisation method.

Table 4.5: Expected frequencies from the guild randomisation method. Note that total guild occurrences are not equal to the observed values.

	A	B	C	D	site richness (mean)
Site 1	0.8	1	1.4	1.8	5
Site 2	1.0	1.3	2.2	2.4	7
Site 3	1.1	1.4	1.9	2.6	7
Site 4	2.6	3.4	5.1	5.9	17
Site 5	2.6	3.4	5.1	5.9	17
Site 6	2.7	3.7	5.3	6.3	18
total occurrence (mean)	11	14	21	25	
95% CI	7-14	10-18	17-25	20-30	
observed	**9**	**13**	**25**	**24**	

4.8 BERNOULLI TRIAL APPROACH

The above analyses highlight the importance of maintaining four key features of the observed data when formulating the null distribution for guild proportionality test across whole assemblages. Since none of the above tests fulfil all four of the key criterion for a good null model, I outline a new approach to randomising species-incidence matrices that preserves the four key features detailed above.

The probability of occurrence of a species S within the species pool ($SPP_{S,POOL}$) is equal to its total number of occurrences in the region divided by the total number of occurrence of all species in the region i.e.

$$\Pr(SPP_{S,POOL}) = \sum_{i=1}^{S}\sum_{j=1}^{L}\frac{n_j}{n_{ij}} \qquad (4.4)$$

where n is the number of occurrences, L is a site and S is a species.

The expected frequency of occurrence of a species S occurring only at site L ($E(SPP_{SL})$)is equal to the probability of species S in the total pool multiplied by the total richness of site L:

$$E(SPP_{SL}) = \Pr(SPP_{S,POOL}) \times R_L \qquad (4.5)$$

where R_L is the richness of site L.

Then we can think of whether species S occurs at site L as a Bernoulli trial with probability p, where p is equal to the expected frequency of occurrence of a species occurring at a site, $E(SPP_{SL}))$.

Now that we know p, we can conduct Bernoulli trials for each species at each site (i.e., for each cell of the species-incidence matrix) and determine whether or not a species is present according to probability p (see Appendix 2 for R code to perform these analyses). Once the species-incidence matrix is randomised, the number of species in each guild can be summed to generate a random guild-incidence matrix. The X^2 statistic is then calculated for each random matrix and combine these over all randomisations to produce the null distribution of test statistics against which the observed X^2 statistic is compared.

This approach allows the richness of sites and the total number of species' occurrences to fluctuate stochastically across randomisation trials, however, the expected value of each is equal to the observed value. This is because the expected frequency of each species occurring at each site (i.e., $E(SPP_{SL})$) sums to the observed richness of site L when calculated across all species occurring at a site (i.e., summed across a row) and $E(SPP_{SL})$ also sums to the observed number of species occurrences when calculated for species S across all sites (i.e., summed down a column). Thus the expected row and column totals match the observed, but may fluctuate across

each random matrix. Furthermore, because each cell of the species-incidence matrix is sampled independently, the observed probability of species in the total pool is known exactly. In some instances, the probability of a species occurring at a site ($E(SPP_{SL})$) may be greater than 1 and in this case, $E(SPP_{SL})$ is set at exactly 1. Allocating species to sites under a Bernoulli trial means that species can only be allocated to a site once, thus preventing guilds from having more species in randomised communities than they have in nature. Finally, the mean number occurrence of guilds is equal to the observed value. Thus the Bernoulli trial approach preserves all four of the criteria proposed for a good null model for guild proportionality tests.

To verify the above points, I applied the Bernoulli trial approach to the hypothetical dataset. Table 4.6 highlights that the site richness, guilds' occurrence and the maximum number of species in a guild are consistent between the observed and the randomised data. By remembering that the species probability of occurrence in the Berboulli trials, p, is equal to the species frequency of occurrence ($E(SPP_{SL})$), we see that all four of the criteria proposed for a good null model for guild proportionality tests are maintained. Furthermore, when the obsseved X^2 statistic is compared to the null distribution of test statistics generated under the Bernoulli trial method, it accurately falls within the random part of the distribution (17.7^{th} percentile, Figure 4.4).

Table 4.6: Expected frequencies from Bernoulli trials. Note that the mean site richness and total guild occurrences are equal to the observed values. Also note that guilds do not contain more species than are observed in the observed data.

	A	B	C	D	site richness (mean)	observed
Site 1	0.6	0.9	1.8	1.7	5	5
Site 2	2.2	3.1	6.0	5.7	17	17
Site 3	0.9	1.3	2.5	2.4	7	7
Site 4	0.9	1.3	2.5	2.3	7	7
Site 5	2.2	3.1	6.0	5.7	17	17
Site 6	2.3	3.3	6.0	6.1	18	18
total occurrence (mean)	9	13	25	24		
observed	9	13	25	24		

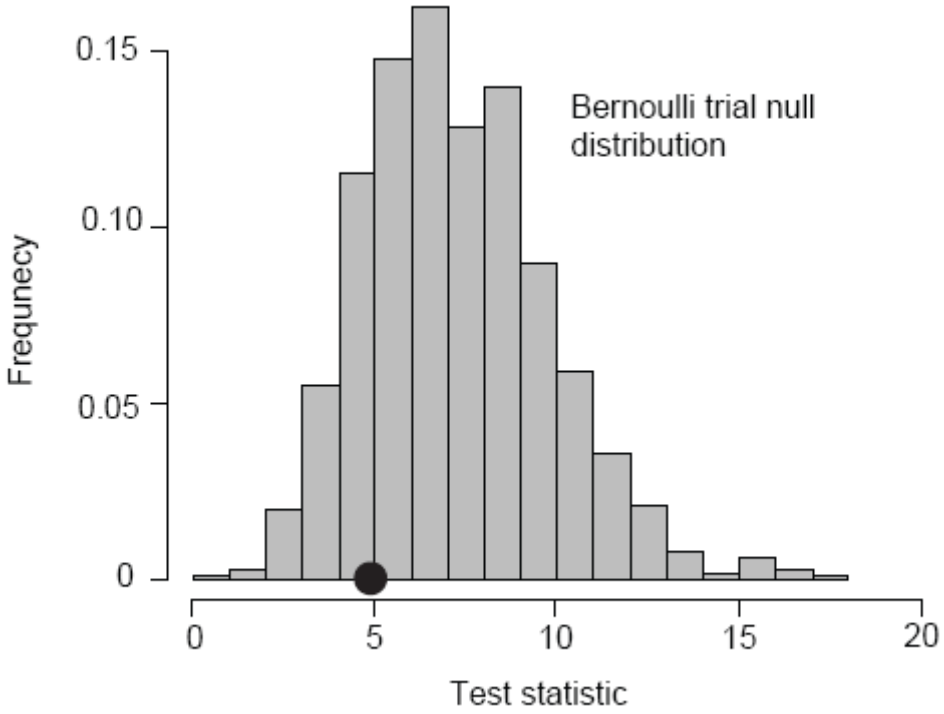

Figure 4.4: Frequency distribution of test statistics from guild-incidence matrices generated by Bernoulli trials (grey bars). The solid circle is the observed (X^2) test statistic, which accurately falls within the random part of the null distribution.

4.9 DISCUSSION

The construction of robust null models of community structure is a rigorously debated topic in the assembly rules literature (e.g. Gotelli and Graves [1996], Wieher and Keddy [1999]). As I have demonstrated above, the evidence of how communities are structured differs with the null model that is used. Thus it is critical that null models of community structure be carefully developed and tested on simulated data. The validity of each of the approaches explored in this chapter (i.e., species matching, modified species matching, guild randomisation, Bernoulli trial) can be determined by considering two factors; firstly, whether the test accurately predicts community structure in simulated data, for which the structure of the assemblage is known and secondly, whether the test preserved important features of the observed data in the null models. These factors are summarised for the four approaches in the tables below.

Table 4.7: Accuracy of tests (assessed using the illustrative dataset).

	Species matching	Modified species matching	Guild randomisation	Bernoulli
Random structure	no	yes	yes	yes

Table 4.8: Preservation of key observed characteristics in null models.

	Species matching	Modified species matching	Guild randomisation	Bernoulli
Site richness (row totals)	yes	yes	yes	yes
Species' frequency of occurrence (column totals)	yes	no	yes	yes
Guilds' frequency of occurrence (column totals)	yes	no	no	yes
Max # species in guilds	no	yes	yes	yes

The species matching approach (Schluter 1990) is clearly not an appropriate test of guild proportionality in entire assemblages. The underlying assumption of this method is that sampling occurs with replacement in randomisation trials. The advantage of sampling with replacement is that species are sampled with probability exactly equal to their observed frequency of occurrence. However, a major disadvantage of sampling with replacement is that species may occur more than once at a single site in random communities. This introduces excess variability into null models and also allows sites to have more species in a guild than occur in that guild in the total species pool. As a consequence of these flaws, the species matching approach overestimates the degree of homogeneity in random or heterogeneously structured communities, which may result in the false assignment of assembly rules.

Randomising incidence matrices through random draws of species from the total species pool (without replacement) was thought to be a robust method of creating null models of community structure, so long as the observed species' frequencies of occurrence and site richness were maintained (Wilson 1995, Brown et al. 2000). However, the analyses performed here have

revealed some inconsistencies with this randomisation method that make the proposed modified species matching approach problematic. In particular, species may not be sampled from the species pool with a probability equal to their frequencies of occurrence. This occurs because when sampling occurs with replacement, the probability of a species being chosen differs with the order in which a species is selected. When these probabilities are averaged over all possible sampling orders the resulting total probability may not be equal to the species' frequency of occurrence. Furthermore, for large species-incidence matrices, averaging each probability over all possible sampling orders will be extremely computationally intensive. As a consequence of this inconsistency, the observed column totals of both species- and guild-incidence matrices may not be equal to the expected column totals in the randomised matrices, in violation of criteria 2 and 3 for a good null model.

The guild randomisation method was developed to construct null models for Fox's favoured states rule (Stone et al. 1996, 2000). Under this rule, the proportion of species within guilds must be homogeneous among separate communities within a region, in addition to all functional groups having equal numbers within each individual community (Fox 1987). However, the guild proportionality rule is less restrictive and does not require all guilds to have the same number of species at any one site, only the same proportion of species between sites. Consequently, when species-incidence matrices are randomised via the guild randomisation method, the observed frequency of occurrence of guilds is not maintained in random matrices.

The Bernoulli trial approach developed above results in accurate predictions of community structure when tested against simulated data. Furthermore, it incorporates the key features of observed incidence matrices that are important to maintain in null models, and is the only one of the four approaches proposed here to do so. In addition to having good statistical properties, the

Bernoulli method is also the most biologically realistic null model. When testing for assembly rules, the biological null model is that species occur at sites independently of one another. This is reflected in the Bernoulli approach which treats each cell of the species-incidence matrix as an independent trial, in contrast to the other methods which randomise on a site-by-site basis. Thus, the Bernoulli trial approach is proposed as a robust test of guild proportionality in aggregated communities and will be employed in the following chapter to test for assembly rules in macroalgal communities.

Chapter 5: Scale transitions and assembly rules in marine macroalgal communities: a global synthesis[3]

5.1 INTRODUCTION

One of the long-standing debates in community ecology concerns the extent to which species interactions regulate community composition. One view is that interactions among species are highly influential in determining community structure and that such interactions constrain the structure of assemblages to a limited suite of possible states (Clements 1936, MacArthur 1972, Brown et al. 2000). An alternative view suggests that interactions are not strong enough to affect community structure. Rather, communities are comprised of individuals that disperse to a site and tolerate the local physical conditions (Mason 1947, Connor and Simberloff 1979, Hubbell 2001). Consequently, community assembly is an individualistic process, determined largely by the effects of historical and environmental conditions on individual species' colonization and extinction dynamics. Communities organized via individualistic processes are often termed "random" assemblages of species, because individualistic processes are thought to give rise to assemblages which resemble random assortments of species from a larger species pool. In contrast, communities in which species interactions are influential are considered to be non-random or deterministic assemblages, since the interactions restrict community composition to a limited suite of possible states.

In 1975 Diamond proposed "assembly rules" as a means of characterizing how individual species combine from larger species pools to form communities. Assembly rules formally describe empirical non-random patterns in community organization (Weiher and Keddy 1999). The first assembly rules concerned patterns of species co-occurrence in avian communities

[3] This chapter is in prep as: Kerswell AP and Connolly SR (in prep) Scale transitions and assembly rules in marine macroalgal communities: a global synthesis. The American Naturalist.

(Diamond 1975), and since then they have been sought more broadly. For example, a recent meta-analysis applied Diamond's assembly rules across 96 terrestrial datasets and revealed that species co-occur in most natural communities less frequently than expected by chance (Gotelli and McCabe 2002). Species interactions are through to be the primary mechanism that determines non-random patterns of community structure i.e., assembly rules (Diamond 1975, Weiher and Keddy 1999), however, it important to note that other factors, such as habitat specificity, may also be influential (Peres-Neto 2004).

Assembly rules may also be present at levels of organization higher than individual species, such as functional groups or higher taxonomic levels (Wilson 1999). In such cases, the relative proportion of species within each group is predicted to remain stable across communities (the "guild proportionality rule" of Wilson [1989]), with assemblages exhibiting a more consistent composition than would be expected on the basis of a random assortment of taxa from the available species pool. A particularly restrictive form of this rule (the "favoured states" rule of Fox [1987]) posits that each guild is equally represented across different local communities and, within each local community, the difference in the number of species in each guild differs by no more than one. Such communities are said to be in a "favoured state".

Establishing the existence of assembly rules requires that the observed assemblages deviate significantly from assemblages constructed at random from the larger species pool. However, specific tests have been conducted in different ways. For Diamond's co-occurrence rules, presence-absence matrices are randomised so that the observed total number of occurrences each species and the observed richness of sites (i.e., row and column totals) are maintained while presence records are swapped among matrix cells. Patterns of species occurrence are then directly compared between observed and randomised matrices (Connor and Simberloff 1979,

Gotelli and McCabe 2002). Tests for the guild proportionality rule also randomise presence-absence matrices, however, species are subsequently pooled into functional groups. If the number of species in groups is less variable across sites in the observed versus the randomised matrices, there is evidence of guild proportionality (Wilson 1989). In this chapter I utilize the Bernoulli trial approach developed in the previous chapter to test for guild proportionality in aggregated macroalgal assemblages.

Patterns of assemblage structure have been investigated at a variety of scales. Temporally, there is evidence for non-random patterns in community structure, which persist over periods ranging from several years (Fukami 2004) to millennia (McGill et al. 2005). Spatially, non-random patterns of community structure have been documented in $<1m^2$ quadrats (e.g. in grass communities [Wilson 1999]) as well as across entire deserts and island archipelagos (e.g. small mammals and birds [Diamond 1975; Kelt et al. 1995]). However, there has been virtually no attention to how patterns of community structure (i.e., assembly rules) change with scale. The few exceptions (Gotelli and Ellison 2002, Mouillot et al. 2005) have focused on patterns of species co-occurrence, and have not yet yielded unambiguous conclusions about how assembly rules change with or depend on spatial scale.

Here I examine assemblage structure in macroalgal communities. Specifically, I conduct a multi-scale test of the guild proportionality rule, using a hierarchy of random assembly models at scales ranging from regional to global. The analyses identify highly non-random structure within macroalgal communities consistent with a guild proportionality assembly rule. Additionally, these analyses reveal the geographical scales over which guild proportionality rules apply, and they also show how environmental and historical differences among regions can give rise to differences in assemblage structure among biogeographic regions.

5.2 METHODS

The composition of benthic marine algal assemblages was determined at 120 sites across the Indo-Pacific and Atlantic Oceans. Sites were located between 50^{o}S and 70^{o}N and assigned to 1 of 6 regions – Indian (20), tropical Pacific (30), north Pacific (10), south Pacific (14), tropical Atlantic (21), north Atlantic (25) (numbers in brackets indicate how many sites occur in each region, see Figure 5.1 for location of sites). Tropical regions extended between 30^{o}S and 30^{o}N, as the 30^{o} parallels are generally recognized as the limit of tropical marine ecosystems (i.e., coral reefs). The south Atlantic was not included, because there were too few sites to conduct meaningful analyses.

The genus composition of algal assemblages was derived from a database detailing the global distribution of 1069 genera of fully marine macroalgae (see Chapter 2, Kerswell [2006]). An algal genus was recorded as present in an assemblage if its geographic range overlapped the site at which that assemblage occurred. Thus local assemblages are comprised of all the genera that potentially occur at a site (due to range overlap) i.e., the local genus pool. Such measures of local communities have been used successfully in previous studies of assembly rules in both marine and terrestrial systems (Kelt 1999, Bellwood and Hughes 2001). Additionally, geographic range-derived estimates of occurrence reduce biases associated with patchy sampling effort of phycological studies across the globe (McAllister et al. 1994, Kerswell 2006).

Algal assemblage structure was defined at every site as the genus richness of nine algal functional groups. The functional groups were based on those of Steneck and Dethier (1994) (see Table 5.1 for full details). All of the functional groups are globally distributed.

5.2.1 Hierarchy of random assembly models

I tested for guild proportionality in algal assemblages by comparing observed assemblages to null models, which assumed random assemblage structure at a variety of scales. Assemblage structure was examined at scales ranging from global to regional and separate models were constructed at each geographic scale. Each scale was chosen to reflect different assumptions about the underlying biological groupings of the source genus pools. For each model I generated a source genus pool (in the form of an incidence matrix) by recording the presence and absence of all algal genera at all sites encompassed by the scale of the model. Thus the manipulation of scale occurred as a scaling down of the source genus pools from which random communities were assembled (see below), while the structure of observed "local" communities remained unchanged. See Figure 5.2 for a visual representation of the relationship between the models.

At the largest scale, the global model tested the simplest scenario whereby all communities follow the same guild proportionality rule; i.e., the proportion of each algal group was more similar across sites than if communities were assembled at random from the global genus pool. The global model was then sub-divided in two ways. Firstly, the global model was sub-divided at the biome scale, since ~30% of macroalgal genera have geographic ranges that are restricted to temperate waters. This division resulted in two biomes models, tropical and temperate, which each tested for guild proportionality rules within either tropical or temperate algal communities. Secondly, the global model was sub-divided at an ocean-basin sale, since ~40% of algal genera are endemic to either the Indo-Pacific or Atlantic Ocean. This division resulted in two oceans models, Indo-Pacific and Atlantic, which each tested for guild proportionality rules within the respective oceans. At the smallest scale, six regional models were developed to test for separate

guild proportionality rules in each region. These models allowed for the evolutionary and physiological separation of algal taxa.

5.2.2 Testing for guild proportionality

I used the Bernoulli trial approach developed in Chapter 4 to test for guild proportionality in aggregated algal communities. Observed genus-incidence matrices were constructed for each model (e.g. global model: 120 sites x 1069 genera, tropical biome model: 71 sites x 650 genera). Each genus-incidence matrix was randomised according to the Bernoulli trial algorithm, to create 1000 random genus-incidence matrices for each model. These matrices were summarized into functional group-incidence matrices by summing all the genera in each functional group at each site included in the model. A null distribution of X^2 test statistics was created by calculating the chi-squared statistic for all functional group-incidence matrices generated for each model. Each observed incidence matrix was also summarized into a functional group-incidence matrix and the corresponding X^2 test statistic calculated.

The position of the observed test statistic for each model was compared with its corresponding null distribution in order to assess homogeneity in the frequency of occurrence of groups across sites at each scale and determined if communities were assembled at random from larger genus pools. Observed statistics that fell in the middle of the null distribution (i.e., between the 2.5^{th} and 97.5^{th} percentile) indicated that variation in assembly composition was not significantly different from a random assortment of the algae available in the larger genus pool (i.e., random structure). Observed statistics that fell in the extreme right tail of the null distribution ($>97.5^{th}$ percentile) indicate assemblage composition that was more heterogeneous across sites than expected under a random assortment and observed statistics that fell in the extreme left tail of the null distribution ($<2.5^{th}$ percentile) indicated highly homogeneous assemblage structure

across sites (i.e., assemblage structure was more similar among sites than expected by chance and consistent with a guild proportionality rule).

In addition to the null model analyses, I used linear discriminate analysis to explore the nature of observed algal assemblages. The first and second discriminant scores were plotted for each site and the location of regions in multidimensional space defined as the boundary around all the sites occurring in that region. All analyses were performed using $R_{2.3.0}$ (R Development Core Team [2004]).

5.3 RESULTS

I found strong evidence that algal assemblages at sites were not a random subset of the larger genus pool, at any scale. At the largest, global scale the observed X^2 test statistic was located to the extreme right of the null distribution (Figure 5.3, Table 5.2), highlighting assemblage composition that was significantly more heterogeneous among sites across the globe than predicted by the corresponding random assembly models (i.e., assemblage structure more different than expected by chance). Similarly at the oceans scale, in both the Indo-Pacific and the Atlantic Oceans models, the observed X^2 statistics were also located towards the far right of the null distribution (Figure 5.4A,B Table 5.2). Hence there is no evidence that a consistent guild proportionality rule applies at any of these very broad spatial scales.

At the biome scale, the location of the observed chi-squared test statistics relative to the null distribution differs for tropical versus temperate models. For the temperate biomes model the observed X^2 test statistic is located to the extreme right of the null distribution (Figure 5.4D, Table 5.2), as in the global and oceans models. However, for the tropical biomes model the observed test statistic falls to the extreme left of the null distribution (Figure 5.4C). This indicates that assemblage composition that was more homogeneous among sites in the tropics

than predicted by the corresponding random assembly model (i.e., assemblage structure more similar than expected by chance) and thus there is evidence for significant guild proportionality within tropical biomes. At the regional scale, there is also a trend towards highly heterogeneous structure in temperate regions and highly homogeneous structure in tropical regions (Figure 5.5, Table 5.2), again indicating guild proportionality among tropical but not temperate sites.

Patterns represented in multivariate plots of the observed algal assemblage data are consistent with the patterns of algal assemblage structure revealed by the null model analyses (Figure 5.5). Each region occupies a unique area of multidimensional space (i.e., no two regions are exactly superimposed upon one another) and while there is noticeable overlap of tropical regions, temperate regions tend to be distinct both from the tropics and from one another. Biomes are separated along the first discriminant axis, with the majority of temperate sites lying to the right and the majority of tropical sites lying to the left. This axis alone explains ~50% of the between region variance in algal assemblage structure.

5.4 DISCUSSION

5.4.1 Assembly rules in macroalgal communities

The results of this chapter reveal highly non-random structure of macroalgal communities. Both within individual tropical regions and across the tropics as a whole, algal assemblages show significant consistency in functional group composition across sites, suggesting the existence of guild proportionality at these scales. Previously, evidence for guild proportionality has not been found at scales greater than a few hectares and has only rarely been found in plant communities (Wilson 1999). The finding that the guild proportionality rule applies to algal assemblages at scales of thousands of kilometres highlights how examining community assembly at a macroecological scale can complement the insights afforded by finer-scale studies.

Algal assemblage structure changes substantially with scale. At the very broadest scales, guild proportions vary greatly among sites and there is no evidence of global or ocean-basin guild proportionality rules. In contrast, algal assemblages within tropical regions conform to region-specific guild proportionality rules, despite large differences in overall algal richness at this scale (e.g. genus richness of assemblages varies over 2-fold in all tropical regions: Kerswell [2006]). In contrast to the tropics, regional and biome guild proportionality rules do not hold in temperate areas. There is large variation in guild proportions within individual temperate regions. For example, most groups vary in guild proportions by at least 2-fold across sites within a single region, but this difference may be as great as 8-fold in some groups (e.g. gelatinous functional group in the north Atlantic). At a biome scale, these difference are even greater, e.g. up to 9-fold in the fine branching functional group. Thus there is no evidence for guild proportionality within temperate areas at any of the scales investigated in this study.

5.4.2 Mechanisms of community assembly

Instances of non-random patterns of community structure are generally used to infer that interspecific competition shapes community structure (Weiher and Keddy 1999). Indeed the guild proportionality rule was initially proposed as a specific test for interspecific competition as a driver of community structure (Wilson 1989). The link between guild proportionality and competition is typically thought to be niche limitation. If the number of resources is limited, then some species will competitively exclude others within the same guild and as a result the proportion of species in a community which belong to any one guild will be limited (Wilson 1989). In macroalgal communities, light and available substrate are highly limiting resources and at local scales there is direct evidence that competition for these factors can be influential in structuring macroalgal communities in predictable and consistent ways (Luning 1990, Irving

and Connell 2006). However, patterns alone do not uniquely imply process and it is therefore important to consider what other processes may drive communities towards non-random configurations (Gotelli and McCabe 2002, McCay et al. 2004).

Factors other than competitive interactions that generate non-random patterns of community structure include shared habitat preferences between taxa, the location of geographic ranges and non-competitive species interactions (McCay et al. 2004, Peres-Neto 2004). It has been shown across several taxa that communities that occur in areas of similar habitat type are likely to be highly convergent (Schluter and Ricklefs 1993, Peres-Neto 2004). Thus it is possible that the highly homogeneous nature of algal assemblage structure within the tropics may be an artefact of individual regions sharing highly similar habitats. However, sites within individual regions and across the tropical biome (where guild proportionality rules are present) encompass a range of habitat types including coral reefs (e.g. Micronesia, Jamaica), rocky reefs (e.g. Taiwan, Brazil), sandy beaches (western Australia, west African coast) and muddy shores (e.g. Bay of Bengal, Mauritania). Despite these large differences in habitat, algal assemblages are highly homogeneous in structure throughout the tropics.

Non-competitive interactions, such as mutualistic associations, may also cause guild proportionality. In the case of macroalgae, ~12.5% of algal genera are obligate epiphytes or parasites of other macroalgae. Since the epiphytic/parasitic genera are not distributed equally across functional groups, I reran the above analyses excluding this 12.5% of taxa. Excluding the epiphytic/parasitic taxa did not alter the results at the global, oceans or biomes scale, and at the regional scale, the outcome of comparisons observed guild proportions and their corresponding null models were the same for all but the Indian Ocean, where support for guild proportionality became statistically non-significant. This suggests that mutualisms do not drive the observed

guild proportionality rules. Additionally, the removal of epiphytic algae from analyses assumes that competitive interactions do not occur within the epiphytic group, yet this is not always the case in natural systems (Morcom et al. 2005).

Some authors argue that highly structured communities (i.e., communities in which assembly rules are evident) may result merely from the location and overlap of species' geographic ranges, rather than from species interactions (Stone et al. [1996, 2000] but see also Brown et al. [2000]). In this study, local algal assemblages were defined by those genera whose geographic ranges overlap a particular site. Therefore, if the location of algal geographic ranges were driving guild proportionality in tropical regions, then the rules should only be present at those scales at which most sites were incorporated by the ranges of most genera. That is, the same genera would contribute to the structure of most sites, resulting in highly similar assemblages within an area. However, in addition to strong regional guild proportionality rules, significant guild proportionality is found throughout the tropical biome. At this scale, the geographic ranges of many taxa (~45%) are restricted to either the tropical Atlantic or Indo-Pacific. Hence sites throughout the tropics (especially those in the Indo-Pacific versus the Atlantic Ocean) differ in which genera contribute to the structure of communities, but functional groups still constitute very similar proportions of local richness across all tropical sites.

At the very broadest scale, historical and environmental factors are likely to be responsible for the lack of consistent guild proportionality within oceans or across the globe. At these large scales, algal communities are highly heterogeneous, reflecting differences between tropical and temperate environments. Temperate regions have long been physically separated, both by land barriers and the tropical Tethys Sea, leading to isolation of species within either the Atlantic or Indo-Pacific Oceans and within separate hemispheres. Such regional isolation is likely to have

resulted in the evolution of unique algal floras in different temperate areas and hence the lack of a temperate biome guild proportionality rule. Within temperate regions, eastern and western coastlines often experience very different temperature and light regimes. For example, in the northern Atlantic Ocean, temperatures are consistently 10^{o}C lower on the western as compared to the eastern coastlines as a result of the polar Labrador Current (Pickard and Emery 1990). Consequently, algal communities in the eastern north Atlantic have species rich floras dominated by warm-temperate taxa, while communities on the western coastlines are characterized by species poor Arctic-associated taxa (Luning 1990). Consequently, guild proportionality rules for macroalgae do not hold at a regional scales (as defined in this study) in temperate areas.

Biological interactions appear to be important in maintaining consistent algal assemblage structure only within the tropics. This finding compares well with patterns of community structure in coral and reef fish assemblages, which are highly similar throughout much of the tropics (Bellwood and Hughes 2001, Bellwood et al. 2002). The tropics are climatically quite stable and all three tropical regions were connected until only 3.2 million years ago (with the closing of the Isthmus of Panama). These factors may have lessened the influence of historical and environmental differences and facilitated homogenization of assemblage structure (i.e., the evolution of a consistent guild proportionality rule) across the tropics to a much greater degree than in temperate regions.

5.4.3 Methodology

The approach employed in this study has a number of advantages. Firstly, the multi-scale analysis allowed several random assembly models to be formulated. This suite of models allowed me to unambiguously identify the scales at which guild proportionality assembly rules

operate. The use of the Bernoulli trial approach for testing guild proportionality rules also has a number of advantages. I was able to test the guild proportionality hypothesis in the aggregate, rather than testing each guild separately. This overcomes the loss of power and potential ambiguity of testing a single ecological hypothesis by means of multiple statistical tests. Furthermore, employing a 2-tailed test allowed me to identify deviations from random structure as either more heterogeneous or more homogeneous than expected by chance. Finally, the multivariate analyses are highly consistent with the model-based analyses, indicating that the results are likely to be quite robust.

In conclusion, much of the debate over assembly rules has focused on a dichotomy: whether individualistic processes or species interactions shape the structure of ecological communities. This study highlights how processes that generate highly homogeneous non-random assemblage structure, such as species interactions, and those that do not, such as environmental or historical differences, can be important at different scales or in different parts of the world, even within a single taxonomic group. In particular, I find that within tropical regions and across the tropical biome as a whole macroalgal assemblages show strikingly concordant composition, consistent with Wilson's (1989) guild proportionality rule, even when genus richness and habitat type vary substantially at these scales. The unprecedentedly large scale at which I find these rules to be operating suggests that community structure may be homogenised over much larger geographical scales than has previously been realized (cf. Wilson [1999]). However, I also identify a role for environmental and historical differences, especially outside the tropics. There is substantial potential for such multiple-scale approaches to provide a framework for understanding how different processes interact across scales to shape global patterns in community structure. For this reason, I believe that such approaches warrant further exploration.

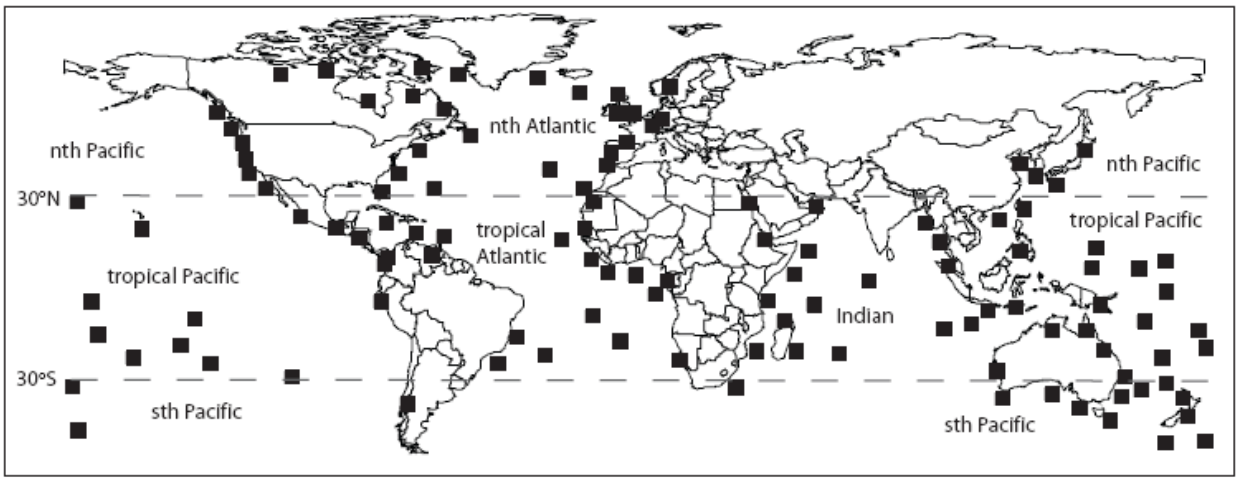

Figure 5.1: Location of the 120 sites at which algal assemblage structure was determined. Sites were pooled into six regions: Indian Ocean, tropical Pacific, north Pacific, south Pacific, tropical Atlantic and north Atlantic. The tropical biome extends from 30^{o}N to 30^{o}S and the temperate biome encompasses areas to the north and south of these latitudes.

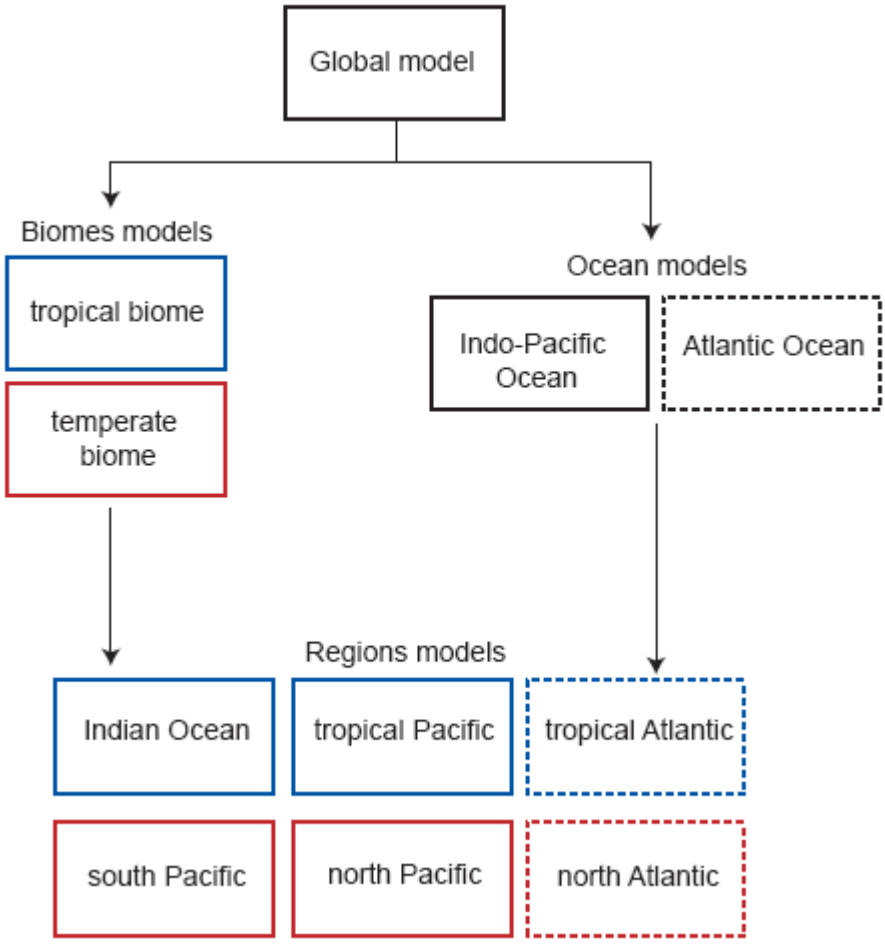

Figure 5.2: Diagram showing the relationships between global, ocean, biome and regional models.

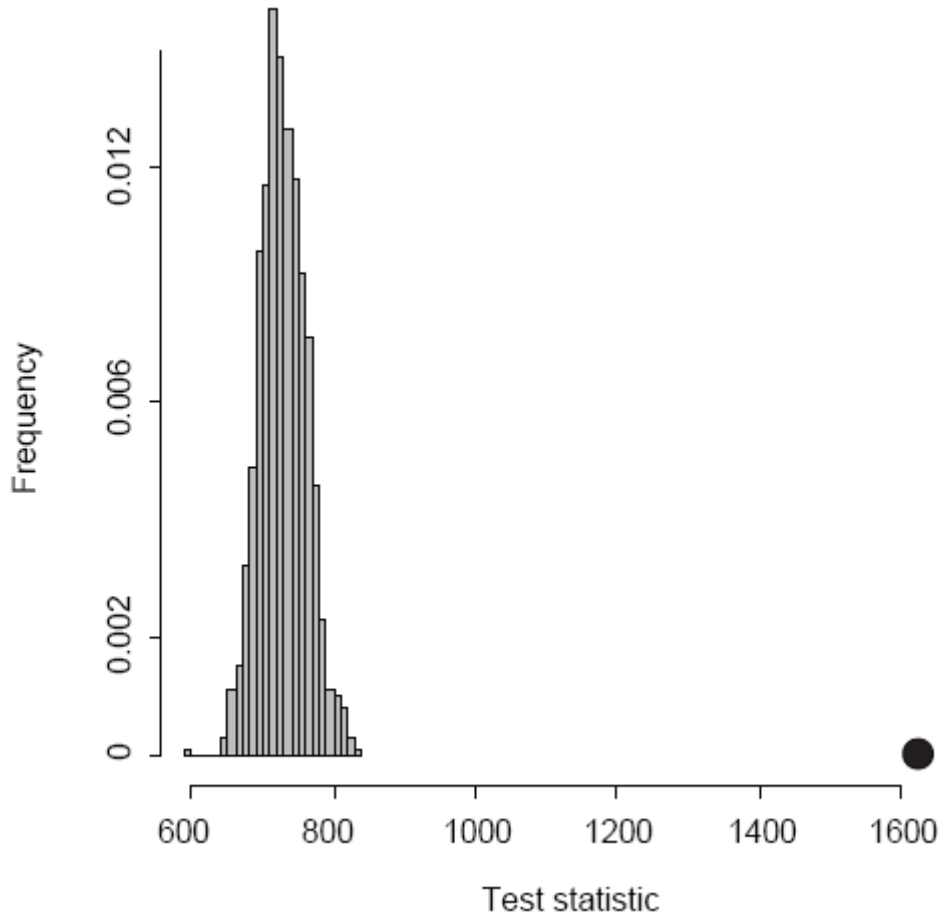

Figure 5.3: Frequency distribution of X^2 statistics from null models at the global scale. Genus-incidence matrices were randomised according to the Bernoulli trial approach (Chapter 4) and aggregated into functional-group incidence matrices. X^2 statistics were calculated for each of these matrices and are presented as the above frequency distributions. The observed X^2 statistic is marked as solid points.

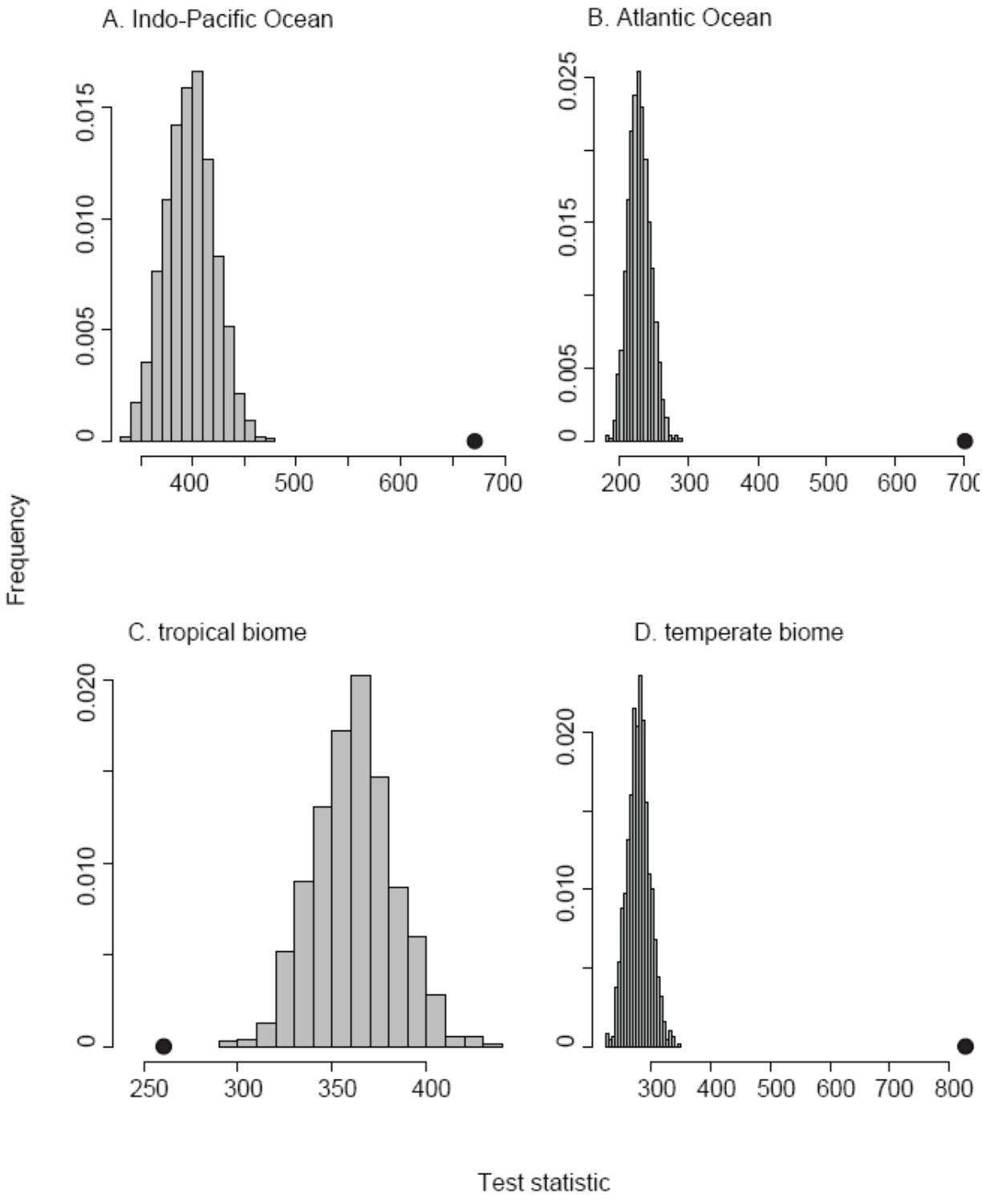

Figure 5.4: See caption on facing page.

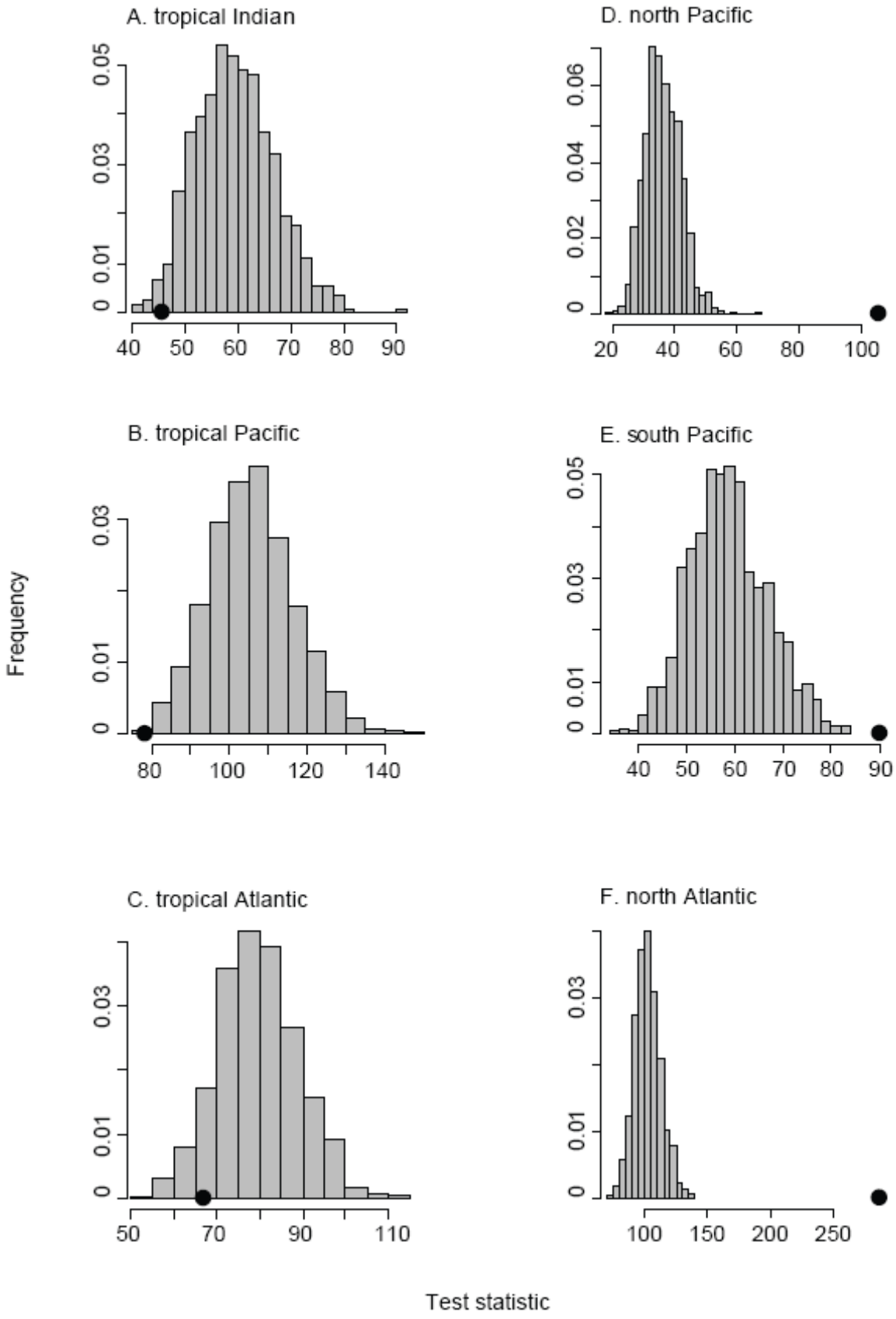

Figure 5.5: See caption on facing page.

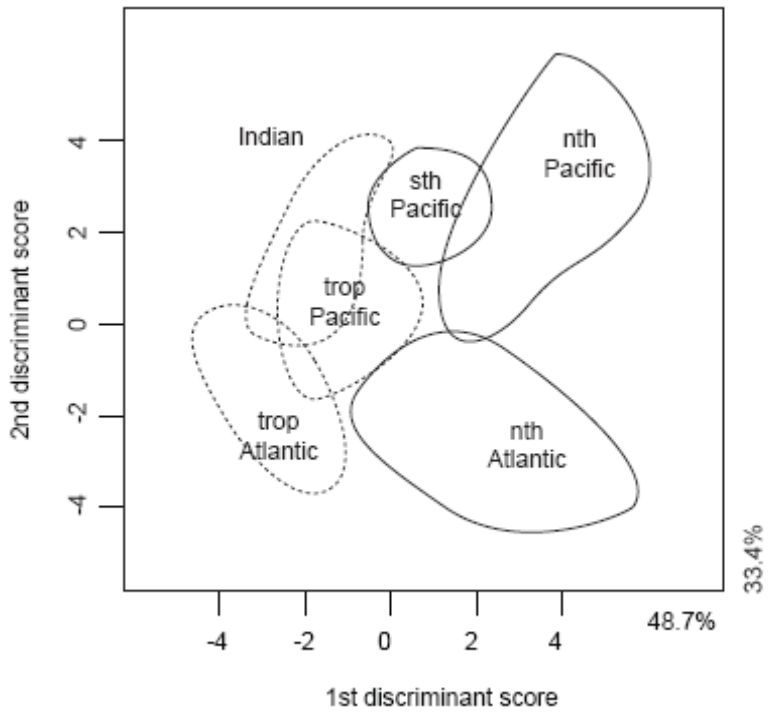

Figure 5.6: Linear discriminant analysis of genus richness at all sites for algal functional groups. The location of individual sites along discriminant axes is not marked, however, shapes outline the outer boundary of sites within each region. Tropical regions are delineated with dashed lines and temperate regions with full lines.

Table 5.1: Functional group classifications used in this study and corresponding classification of Steneck & Dethier (1994).

Functional group (this study)	Steneck & Dethier (1994) classification	Comment	Example
Coarsely branched	Corticated macrophytes	-	*Chondrus, Gigartina*
Crustose	Crustose algae	May be free living or epiphytic	*Peyssonnelia, Ralfsia*
Filamentous	Filamentous algae	-	*Cladophora*
Finely branched	Corticated foliose algae	-	*Dictyota, Padina*
Gelatinous	-	Soft and gelatinous; may be lightly calcified	*Liagora*
Jointed calcareous	Articulated calcareous algae	-	*Jania, Halimeda*
Leathery	Leathery macrophytes	-	*Laminaria, Sargassum*
Sheets	Foliose algae	Single cell layer in sheet or tube	*Ulva, Hypoglossum*
Tufts/feathery	Filamentous algae	Filaments gathered in tufts, often with common stalk or holdfast	*Asparagopsis, Chlorodesmis*

Table 5.2: Observed chi-squared statistics and the corresponding percentile of the null distribution for each model (null models were based on 1000 randomisations).

Scale	Model	Observed X^2 statistic	Percentile of null distribution
Global	Global	1624	>0.999
Biome	Tropical	261	<0.001
	Temperate	828	>0.999
Ocean	Indo-Pacific	672	>0.999
	Atlantic	702	>0.999
Regional	Indian	46	0.018
	tropical Pacific	78	0.010
	tropical Atlantic	66	0.072
	north Pacific	106	>0.999
	south Pacific	90	>0.999
	north Atlantic	287	>0.999

6.1 ALGAL RICHNESS PATTERNS

The first maps of global diversity patterns for benthic marine algae arose from this thesis (Kerswell 2006). These maps show that contrary to previous suggestions, benthic marine algae have distinct richness gradients at a global scale. Algal richness peaks at mid-latitudes and decreases both towards the tropics and polar regions. This highlights marine algae as an exceptional group, since very few taxa have richness peaks outside of the tropics (Willig et al. 2003). Longitudinally, algae also display unusual patterns in the Atlantic Ocean, with the highest richness occurring on the eastern Atlantic coastlines. In the Indo-Pacific Ocean, however, richness "hotspots" occur in southern Australia and Japan, resulting in a band of very high genus richness throughout the central Indo-Pacific, which is consistent with richness patterns of other taxa in this ocean (Hughes et al. 2002, Roberts et al. 2002).

Prior to this study, the general consensus was that macroalgae did not display globally consistent latitudinal richness gradients (Bolton 1994, Willig et al. 2003). However, this is clearly not the case. Previous studies of algal richness were conducted at restricted spatial scales, and when synthesised by Bolton (1994) no clear gradients in algal richness were evident. Yet, while Bolton's (1994) results are consistent with those documented here – areas of high and low richness occur both in the tropics and in temperate regions – the inference that algae do not show latitudinal gradients at a global scale is false. When algal richness is considered in 2-dimensions, i.e., both latitudinally and longitudinally, it becomes clear that the gradients are distinct. The results of this study, therefore, highlight the importance of considering richness patterns both at very broad geographic scales, as well as 2-dimensionally.

Rigorously documenting algal genus richness patterns provids the first opportunity to explore the processes which create and maintain these patterns at a global scale. This study also highlights the importance of considering species' geographic ranges when attempting to understand the processes that drive richness patterns. I analysed algal richness gradients in the context of patterns in range size and overlap to understand the likely drivers of algal richness patterns at a global scale. There is little evidence that energy, productivity, climatic stability or competition are major determinants of algal richness patterns. Rather, the results implicate habitat areas and ocean currents as the most plausible drivers of observed diversity patterns.

In order to understand patterns of marine diversity more generally, I performed a case study of the order Bryopsidales. The order Bryopisdales is reef associated group (Littler and Littler 2003), which has Indo-Pacific richness patterns that are strikingly concordant with those corals and reef fishes. Using a spatial regression-based approach to model the patterns of bryopsidalean richness, and comparing the results to those from an analogous study of corals and reef fishes (Bellwood et al. 2005), I revealed several key features about the processes that create and maintain marine diversity in the tropical Indo-Pacific. Firstly, geometric constraints (i.e., a mid-domain effect) are highly important in predicting richness patterns for all three groups. Secondly, while temperature and nitrate are the environmental variable most well correlated with algal richness, reef area is the environmental variable most well correlated with coral and reef fish richness. This suggests that the environmental variables that are most important in determining species richness patterns are likely to be taxon specific. Finally, by examining the deviations of the richness patterns predicted by both the bryopsidales and the coral and fish model, I identified a potential role of ocean currents in shaping richness patterns in the tropical Indo-Pacific Ocean.

6.1.1 Algal richness patterns – future directions

To date the conventional approach to analysing the drivers of species richness patterns has been to use regression models. However, associated with this approach is the problem of understanding causation from correlative studies. Thus the challenge for future studies which aim to uncover the processes that create and maintain diversity gradients is to conduct the analyses in a more causative framework. A possible way forward has recently been suggested by Connolly (2005), who developed an approach that incorporates both geometric constraints and environmental gradients into process based models of species richness gradients. A complementary approach is that developed by Goldberg et al. (2005), which allows for testing of hypotheses about regional rates of taxon origination, extinction, and dispersal using information on the ages and current distributions of taxa. Now that the most fundamental data on algal distributions are available, there is much scope to apply new and emerging techniques to understand the processes driving the observed patterns.

6.2 ALGAL COMMUNITY STRUCTURE

To date the debate concerning the processes that shape communities has been highly polarised (Weiher and Keddy 1999). However, the new Bernoulli trial-based guild proportionality rule developed in Chapter 4, and its application to algal data across multiple spatial scales (Chapter 5), highlights that processes such as species interactions and historical or environmental differences, may be important at different scales. For marine algal functional groups, guild proportionality rules hold throughout both individual tropical regions and the tropical biome as a whole. At these scales, species interactions appear to be important in shaping communities. In contrast, algal community structure is highly heterogeneous in temperate oceans, at all scales examined in this study. Temperate regions have been long isolated from one another, and within

individual regions, environmental conditions fluctuate greatly, thus preventing the operation of guild proportionality rules.

The new Bernoulli trial-based guild proportionality rule developed in this thesis has numerous advantages over existing tests. Foremost, the test lends itself to application across multiple spatial scales. By applying it to algal data at scales ranging from regional to global, I have identified both the scales at which guild proportionality rules hold and how different processes may interact across spatial scales to shape global patterns in community structure. Methodologically, the new test overcomes several of the problems associated with testing for guild proportionality on a guild-by-guild basis. Furthermore, the test preserves all the key features of observed communities, including the frequency of occurrence of both species and guilds. Most importantly, the Bernoulli trial approach to randomising species-incidence matrices provides a null model that is both biologically and statistically consistent; species enter communities (i.e., are drawn from the larger species pool) independently of one another.

6.2.1 Algal community structure – future directions

Assembly rules are highly controversial and difficult to document (Weiher and Keddy 1999). This is largely due to the reliance on statistical models to draw conclusions about the processes that shape community structure (Stone et al. 1996). Currently there is a dearth of experimental data available to rigorously test and calibrate tests for assembly rules. Consequently, future manipulative studies which directly test the extent to which species interactions influence the structure of communities will be critical both to understanding whether assembly rules exist and also for calibrating the existing statistical tests.

6.3 SUMMARY

The overarching aim of this thesis was to use benthic marine algae as a model group to test on a global scale, some of the most common theories of biodiversity and to improve our understanding of the processes that drive the observed patterns. The first step in this research was to identify the genus richness patterns for marine algae, and in doing so I have revealed that algae have distinct, yet exceptional richness gradients. I have highlighted the value of understanding richness patterns in the context of both the number of taxa at a site and also how species' ranges overlap in geographic space. The results of this thesis highlight that very broad-scale processes such as geometric constraints and ocean currents are likely to be highly important in determining the richness patterns of the tropical Indo-Pacific Ocean, while the environmental factors that are of primary importance appear more taxon specific. Our on-going challenge is to extend the methods used to understand the drivers of richness patterns beyond correlative approaches. Biodiversity may also be considered in terms of community structure and characterised by assembly rules. Here I have developed a new test for the guild proportionality rule, which is robust to many of the statistical problems associated with other assembly rule tests. When applied to macroalgal communities across multiple spatial scales, this test highlights how processes that generate highly homogeneous assemblage structure, such as species interactions, and those that do not, such as environmental or historical differences, can both be important in shaping community structure. Thus this thesis has both furthered our knowledge of how diversity is distributed on earth and made important inroads into understanding the drivers of the observed patterns.

LITERATURE CITED

Abbott, I. A. 2004. Marine Green and Brown Algae of the Hawaiian Islands. Bishop Museum Press, Honolulu.

Adey, W. H., and R. S. Steneck. 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.

Akaike, H. 1985. Prediction and entropy. Pages 1-24 *in* A. C. Atkinson and S. E. Fienberg, editors. A celebration of statistics, The ISI centenary volume. Springer Verlag, Berlin.

Allen, A. P., J. H. Brown, and J. Gillooly. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science **297**:1545-1548.

Behrenfeld, M. J., J. T. Randerson, C. R. McClain, G. C. Feldman, S. O. Los, C. J. Tucker, P. G. Falkowski, C. B. Field, R. Frouin, W. E. Esaias, D. D. Kolber, and N. H. Pollack. 2001. Biospheric primary production during an ENSO transition. Science **291**:2594-2597.

Bellwood, D. R., and T. P. Hughes. 2001. Regional-scale assembly rules and biodiversity of coral reefs. Science **292**:1532-1534.

Bellwood, D. R., T. P. Hughes, S. R. Connolly, and J. E. Tanner. 2005. Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. Ecology Letters **8**:643-651.

Bellwood, D. R., P. C. Wainwright, C. J. Fulton, and A. Hoey. 2002. Assembly rules and functional groups at global biogeographical scales. Functional Ecology **16**:557-562.

Biehl, C. C., and W. J. Matthews. 1984. Small fish community structure in Ozark streams: Improvements in the statistical analysis of presence-absence data. American Midland Naturalist **111**:371-382.

Bokma, F., J. Bokma, and M. Monkkonen. 2001. Random processes and geographic species richness patterns: why so few species in the north? Ecography **24**:43-49.

Bolton, J. J. 1994. Global seaweed diversity: Patterns and anomalies. Botanica Marina **37**:241-245.

Bolton, J. J. 1996. Patterns of species diversity and endemism in comparable temperate brown algal floras. Hydrobiologia **327**:173-178.

Bracken, M. E. S., and K. J. Nielsen. 2004. Diversity of intertidal macroalgae increases with nitrogen loading by invertebrates. Ecology **85**:2828-2836.

Breeman, A. M. 1988. Relative importance of temperature boundaries and other factors in determining geographic boundaries of seaweeds: experimental and phenological evidence. Helgolander Meeresutersuchungen **42**:199-241.

Breeman, A. M., and H. Pakker. 1994. Temperature ecotypes in seaweeds: adaptive significance and biogeographic implications. Botanica Marina **37**:171-180.

Brown, J. H., B. J. Fox, and D. A. Kelt. 2000. Assembly rules: Desert rodent communities are structured at scales from local to continental. American Naturalist **156**:314-321.

Brown, J. H., G. C. Stevens, and D. M. Kaufman. 1996. The geographic range: Size, shape, boundaries, and internal structure. Annual Review of Ecology and Systematics **27**:597-623.

Chase, J. M., and M. A. Leibold. 2002. Spatial scale dictates the productivity-biodiversity relationship. Nature **416**:427-430.

Chown, S. L., and K. J. Gaston. 2000. Areas, cradles and museums: the latitudinal gradient in species richness. Trends in Ecology & Evolution **15**:311-315.

Clements, F. E. 1936. Nature and structure of the climax. Journal of Ecology **24**:252-284.

Colwell, R. K., and G. C. Hurtt. 1994. Non-biological gradients in species richness and a spurious rapoport effect. American Naturalist **144**:570-595.

Colwell, R. K., and D. C. Lees. 2000. The mid-domain effect: geometric constraints on the geography of species richness. Trends in Ecology & Evolution **15**:70-76.

Colwell, R. K., C. Rahbek, and N. J. Gotelli. 2004. The mid-domain effect and species richness patterns: What have we learned so far? American Naturalist **163**:E1-E23.

Conkright, M. E., R. A. Locarnini, H. E. Garcia, T. D. O'Brien, T. P. Boyer, C. Stephens, and J. I. Antonov. 2002. World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures. National Oceanographic Data Center, Silver Spring, MD.

Connolly, S. R. 2005. Process-based models of species distributions and the mid-domain effect. American Naturalist **166**:1-11.

Connolly, S. R., D. R. Bellwood, and T. P. Hughes. 2003. Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. Ecology **84**:2178-2190.

Connor, E. F., and D. Simberloff. 1979. The assembly of species communities: Chance or competition? Ecology **60**:1132-1140.

Currie, D. J. 1991. Energy and large-scale patterns of animal and plant species richness. American Naturalist **137**:27-49.

Diamond, J. M. 1975. Assembly of species communties. Pages 342-444 *in* M. L. Cody and J. M. Diamond, editors. Ecology and Evolution of Communitites. Harvard University Press, Cambridge, Massachusetts.

Duke, N. C., E. Y. Y. Lo, and M. Sun. 2002. Global distribution and genetic discontinuities of mangroves - emerging patterns in the evolution of Rhizophora. Trees-Structure and Function **16**:65-79.

Evans, K. L., P. H. Warren, and K. J. Gaston. 2005. Species-energy relationships at the macroecological scale: a review of the mechanisms. Biological Reviews **80**:1-25.

Fargione, J., C. S. Brown, and D. Tilman. 2003. Community assembly and invasion: An experimental test of neutral versus niche processes. Proceedings of the National Academy of Sciences of the United States of America **100**:8916-8920.

Field, C. B., M. J. Behrenfeld, J. T. Randerson, and P. G. Falkowski. 1998. Primary production of the biosphere: Integrating terrestrial and oceanic components. Science **281**:-237240.

Field, R., E. M. O'Brien, and R. J. Whittaker. 2005. Global models for predicting woody plant richness from climate: Development and evaluation. Ecology **86**:2263-2277.

Fox, B. J. 1987. Species assembly and the evolution of community structure. Evolutionary Ecology **1**:201-213.

Francis, A. P., and D. J. Currie. 2003. A globally consistent richness-climate relationship for angiosperms. American Naturalist **161**:523-536.

Fraser, R. H., and D. J. Currie. 1996. The species richness-energy hypothesis in a system where historical factors are thought to prevail: Coral reefs. American Naturalist **148**:138-159.

Fukami, T. 2004. Community assembly along a species pool gradient: implications for multiple-scale patterns of species diversity. Population Ecology **46**:137-147.

Gaines, S. D., and J. Lubchenco. 1982. A unified approach to marine plant-herbivore interactions. II. Biogeography. Annual Review of Ecological Systems **13**:111-138.

Gaston, K. J. 2000. Global patterns in biodiversity. Nature **405**:220-227.

Goldberg, E., K. Roy, R. Lande, and D. Jablonski. 2005. Diversity, endemism and age distributions in macroevoloutionary sources and sinks. American Naturalist 165:623-633.

Gotelli, N. J. 2004. Assembly rules. Pages 1027-1035 *in* M. V. Lomolino, D. F. Sax, and J. H. Brown, editors. Foundations of Biogeography. University of Chicago Press, Chicago.

Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters **4**:379-391.

Gotelli, N. J., and A. M. Ellison. 2002. Biogeography at a regional scale: Determinants of ant species density in New England bogs and forests. Ecology **83**:1604-1609.

Gotelli, N. J., and G. R. Graves. 1996. Null Models in Ecology. Smithsonian Institution Press, Washington.

Gotelli, N. J., and D. J. McCabe. 2002. Species co-occurrence: A meta-analysis of J. M. Diamond's assembly rules model. Ecology **83**:2091-2096.

Graham, M. H. 2003. Confronting multicollineartiy in ecological multiple regression. Ecology **84**:2809-2815.

Green, E. P., and F. T. Short. 2003. World Atlas of Seagrasses. UNEP World Conservation Monitoring Centre, University of California Press, Berkeley.

Guiry, M.D., Nic Dhonncha, E. & Rindi, F. 2005. AlgaeBase version 3.0. World-wide electronic publication, National University of Ireland, Galway. http://www.algaebase.org

Hawkins, B. A., and J. A. F. Diniz. 2002. The mid-domain effect cannot explain the diversity gradient of Nearctic birds. Global Ecology and Biogeography **11**:419-426.

Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J. F. Guegan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, E. E. Porter, and J. R. G. Turner. 2003a. Energy, water, and broad-scale geographic patterns of species richness. Ecology **84**:3105-3117.

Hawkins, B. A., and E. E. Porter. 2001. Area and the latitudinal diversity gradient for terrestrial birds. Ecology Letters **4**:595-601.

Hawkins, B. A., E. E. Porter, and J. A. F. Diniz. 2003b. Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. Ecology **84**:1608-1623.

Howarth, R. W. 1988. Nutrient limitation of net primary production in marine ecosystems. Annual Review of Ecology and Systematics **19**:89-110.

Hubbell, S. P. 2001. The unified theory of biodiversity and biogeography. Princeton University Press, Princeton, NJ.

Hughes, T. P., D. R. Bellwood, and S. R. Connolly. 2002. Biodiversity hotspots, centres of endemicity, and the conservation of coral reefs. Ecology Letters **5**:775-784.

Hunt, G., T. M. Cronin, and K. Roy. 2005. Species-energy relationship in the deep sea: a test using the Quaternary fossil record. Ecology Letters **8**:739-747.

Hurlbert, A. H., and E. P. White. 2005. Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. Ecology Letters **8**:319-327.

Irving, A. D., and S. D. Connell. 2006. Predicting understorey structure from the presence and composition of canopies: an assembly rule for marine algae. Oecologia **148**:491-502.

Jetz, W., and C. Rahbek. 2001. Geometric constraints explain much of the species richness pattern in African birds. Proceedings of the National Academy of Sciences of the United States of America **98**:5661-5666.

Jetz, W., and C. Rahbek. 2002. Geographic range size and determinants of avian species richness. Science **297**:1548-1551.

Jokiel, P., and F. J. Martinelli. 1992. The vortex model of coral reef biogeography. Journal of Biogeography 19:449-458.

Joosten, A. M. T., and C. van den Hoek. 1986. Worldwide relationships between red seaweed floras: A multivariate approach. Botanica Marina **29**:195-214.

Kaspari, M., P. S. Ward, and M. Yuan. 2004. Energy gradients and the geographic distribution of local ant diversity. Oecologia **140**:407-413.

Kelt, D. A. 1999. On the relative importance of history and ecology in structuring communities of desert small animals. Ecography **22**:123-137.

Kerswell, A. P. 2006. Global biodiversity patterns of benthic marine algae. Ecology **87**:2479-2488.

Koleff, P., and K. J. Gaston. 2001. Latitudinal gradients in diversity: real patterns and random models. Ecography 24:341-351.

Laurie, H., and J. A. Silander. 2002. Geometric constraints and spatial pattern of species richness: critique of range-based null models. Diversity and Distributions **8**:351-364.

Lees, D. C., C. Kremen, and L. Andriamampianina. 1999. A null model for species richness gradients: bounded range overlap of butterflies and other rainforest endemics in Madagascar. Biological Journal of the Linnean Society **67**:529-584.

Legendre, P. 1993. Spatial autocorrelation: Trouble or a new paradigm? Ecology **74**:1659-1673.

Lennon, J. J., P. Koleff, J. J. D. Greenwood, and K. J. Gaston. 2004. Contributions of rarity and commonness to patterns of species richness. Ecology Letters **7**:81-87.

Littler, D. S., and M. M. Littler. 2003. South Pacific Reef Plants. Offshore Graphics, Inc, Washinton, D.C.

Lomolino, M. V., D. F. Sax, and J. H. Brown. 2004. Foundations of Biogeography. Chicago University Press, Chicago.

Luning, K. 1990. Seaweeds: Their Environment, Biogeography and Ecophysiology. John Wiley and Sons, New York.

MacArthur, R. H. 1972. Geographical Ecology: Patterns in the Distribution of Species. Harper and Row, New York.

Macpherson, E. 2002. Large-scale species-richness gradients in the Atlantic Ocean. Proceedings of the Royal Society of London Series B-Biological Sciences **269**:1715-1720.

Mason, H. L. 1947. Evolution of certain floristic associations in western North America. Ecological Monographs **17**:201-210.

McAllister, D. E., F. W. Schueler, C. M. Roberts, and J. P. Hawkins. 1994. Mapping and GIS analysis of the global distribution of coral reef fishes on an equal-area grid. Pages 155-175 *in* R. I. Miller, editor. Mapping the Diversity of Nature. Chapman and Hall, London.

McCay, T. S., M. J. Lovallo, W. M. Ford, and M. A. Menzel. 2004. Assembly rules for functional groups of North American shrews: effects of geographic range and habitat partitioning. Oikos **107**:141-147.

McGill, B. J., E. A. Hadly, and B. A. Maurer. 2005. Community inertia of quaternary small mammal assemblages in North America. Proceedings of the National Academy of Sciences of the United States of America **102**:16701-16706.

Miller, M. W., and M. E. Hay. 1996. Coral-seaweed-grazer-nutrient interactions on temperate reefs. Ecological Monographs 66:323-344.

Mora, C., P. M. Chittaro, P. F. Sale, J. P. Kritzer, and S. A. Ludsin. 2003. Patterns and processes in reef fish diversity. Nature **421**:933-936.

Mora, C., and D. R. Robertson. 2005. Causes of latitudinal gradients in species richness: A test with fishes of the Tropical Eastern Pacific. Ecology **86**:1771-1782.

Morcom, N. F., W. J. Woelkerling, and S. A. Ward. 2005. Epiphytic nongeniculate corallines found on *Laurencia filiformis f. heteroclada* (Rhodomelaceae, Ceramiales). Phycologia **44**:202-211.

Mouillot, D., A. Simkova, S. Morand, and R. Poulin. 2005. Parasite species coexistence and limiting similarity: a multiscale look at phylogenetic, functional and reproductive distances. Oecologia **146**:269-278.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature **403**:853-858.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. W. 1996. Applied linear statistical models. Irwin, Chicago, USA.

Pakker, H., A. M. Breeman, W. F. P. Vanreine, and C. Vandenhoek. 1995. A comparative study of temperature responses of caribbean seaweeds from different biogeographic groups. Journal of Phycology **31**:499-507.

Palmer, M. W. 1994. Variation in species richnes - towards a unification of hypotheses. Folia Geobotanica and Phytotaxonomica **29**:511-530.

Patterson, B. D., and W. Atmar. 1986. Nested subsets and the structure of insular mammalian faunas and archipelagos. Biological Journal of the Linnean Society **28**:65-82.

Paulay, G., and C. Meyer. 2002. Diversification in the tropical pacific: Comparisons between marine and terrestrial systems and the importance of founder speciation. Integrative and Comparative Biology 42:922-934.
Peres-Neto, P. R. 2004. Patterns in the co-occurrence of fish species in streams: the role of site suitability, morphology and phylogeny versus species interactions. Oecologia **140**:352-360.

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.

Pickard, G. L., and W. J. Emery. 1990. Descriptive Physical Oceanogaphy: An Introduction. Pergamon Press, Oxford, UK.

Pielou, E. C. 1977. The latitudinal spans of seaweed species and their patterns of overlap. Journal of Biogeography **4**:299-311.

Pielou, E. C. 1978. Latitudinal overlap of seaweed species: evidence for quasi-sympatric speciation. Journal of Biogeography **5**:227-238.

R Development Core Team 2004. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rahbek, C., and G. R. Graves. 2001. Multiscale assessment of patterns of avian species richness. Proceedings of the National Academy of Sciences of the United States of America **98**:4534-4539.

Rex, M. A., J. A. Crame, C. T. Stuart, and A. Clarke. 2005. Large-scale biogeographic patterns in marine mollusks: A confluence of history and productivity? Ecology **86**:2288-2297.

Ribeiro, P.J. Jr, & Diggle, P.J. 2001. geoR: A package for geostatistical analysis. R-News 1: 15–18. ISSN 1609–3631.

Roberts, C. M., C. J. McClean, J. E. N. Veron, J. P. Hawkins, G. R. Allen, D. E. McAllister, C. G. Mittermeier, F. W. Schueler, M. Spalding, F. Wells, C. Vynne, and T. B. Werner. 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. Science **295**:1280-1284.

Rohde, K. 1992. Latitudinal gradients in species diversity: The search for the primary cause. Oikos **65**:514-527.

Rosen, B. 1988. Progress, problems and patterns in the biogeography of reef corals and other tropical marine organisms. Helgolander Meeresuntersuchungen **42**:269-391.

Rosenzweig, M. L. 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge.

Roy, K., D. Jablonski, J. W. Valentine, and G. Rosenberg. 1998. Marine latitudinal diversity gradients: Tests of causal hypotheses. Proceedings of the National Academy of Sciences of the United States of America **95**:3699-3702.

Santelices, B., and I. A. Abbott. 1987. Geographic and marine isolation: An assessment of the marine algae of Easter Island. Pacific Science **41**.

Santelices, B., and P. A. Marquet. 1998. Seaweeds, latitudinal diversity patterns and Rapoport's Rule. Diversity and Distributions **4**:71-75.

Schluter, D. 1990. Species-for-species matching. American Naturalist **136**:560-568.

Schluter, D., and R. E. Ricklefs. 1993. Convergence and the regional component of species diversity. Pages 230-240 *in* R. E. Ricklefs and D. Schluter, editors. Species Diversity in Ecological Communities: Historical and Geographical Perspectives. University of Chicago Press, Chicago, Illinois, USA.

Scrosati, R. 2001. Population dynamics of *Caulerpa sertularioides* (Chlorophyta : Bryopsidales) from Baja California, Mexico, during El Nino and La Nina years. Journal of the Marine Biological Association of the United Kingdom **81**:721-726.

Silva, P. 1992. Geographic patterns of diversity in benthic marine algae. Pacific Science **46**:429-437.

Stehli, F. G., and J. W. Wells. 1971. Diversity and age patterns in hermatypic corals. Systematic Zoology **20**:115-126.

Steneck, R. S., and M. N. Dethier. 1994. A functional group approach to the structure of algal dominated communities. Oikos **69**:476-498.

Stevens, G. C. 1989. The latitudinal gradient in geographic range: how so many species co-exist in the tropics. American Naturalist **133**:240-256.

Stone, L., T. Dayan, and D. Simberloff. 1996. Community-wide assembly patterns unmasked: The importance of species' differing geographical ranges. American Naturalist **148**:997-1015.

Stone, L., T. Dayan, and D. Simberloff. 2000. On desert rodents, favoured states and unresolved issues: Scaling up and down regional assemblages and local communities. American Naturalist **156**:322-328.

van den Hoek, C. 1984. World-wide latitudingal and longitudinal seaweed distribution patterns and their possible causes, as illustrated by the distribution of Rhodophytan genera. Helgolander Meeresutersuchungen **38**:227-257.

Vazquez, L. B., and K. J. Gaston. 2004. Rarity, commonness, and patterns of species richness: the mammals of Mexico. Global Ecol Biogeography **13**:535-542.

Vermeij, G. J. 1978. Biogeography and Adaptation. Harvard University Press, Cambridge Massachusetts, USA.

Veron, J. E. N. 1995. Corals In Space and Time: The Biogeography and Evolution of the Scleractinia. UNSW Press, Sydney.

Weiher, E., P. Clarke, and P. Keddy. 1998. Community assembly rules, morphological dispersion, and the coexistence of plant species. Oikos **81**:309-322.

Weiher, E., and P. Keddy. 1999. Ecological Assembly Rules: Perspectives, Advances, Retreats. Cambridge University Press, Cambridge.

Willig, M. R., D. M. Kaufman, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. Annual Review of Ecology, Evolution, and Systematics **34**:273-309.

Willig, M. R., and S. K. Lyons. 1998. An analytical model of latitudinal gradients of species richness with an empirical test for marsupials and bats in the New World. Oikos **81**:93-98.

Wilson, J. B. 1989. A null model of guild proportionality, applied to stratification of a New Zealand temperate rainforest. Oecologia **80**:263-267.

Wilson, J. B. 1995. Null models for assembly rules: The Jack-Horner Effect is more insidious than the Narcissus Effect. Oikos **72**:139-144.

Wilson, J. B. 1999. Assembly rules in plant communities. Pages 418 *in* E. Weiher and P. Keddy, editors. Ecological Assembly Rules: Perspectives, Advances, Retreats. Cambridge University Press, Cambridge.

Womersley, H. B. S. 1984. The Benthic Flora of Southern Australia. Part I. Government Printer, Adelaide.

Womersley, H. B. S. 1990. Biogeography of Australasian marine macroalgae. Pages 367-381 *in* M. N. Clayton and R. J. King, editors. Biology of Marine Plants. Longman Australia Pty Ltd, Melbourne.

Worm, B., H. K. Lotze, H. Hillebrand, and U. Sommer. 2002. Consumer versus resource control of species diversity and ecosystem functioning. Nature **417**:848-851.

Zapata, F. A., K. J. Gaston, and S. L. Chown. 2003. Mid-domain models of species richness gradients: assumptions, methods and evidence. Journal of Animal Ecology **72**:677-690.

APPENDIX 1: DATABASE SUMMARY

Table A1: Summary of algae occurrence database. The number of taxa at each site represents the range-derived data used to create diversity maps (Figure 2.2 & 2.3). References are the original species lists from which occurrence data were extracted and geographic ranges complied.

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
1	215	0	44	0	12.47	53.87	140
2	261	0	51	0	-26.01	32.58	32, 36
3	247	2	24	1	-33	17.56	141
20	289	0	83	0	-8.12	156.83	184
21	289	0	82	0	-9.05	159.15	184
22	290	0	82	0	-8.63	158.15	184
23	287	0	82	0	-9.07	160.13	184
24	288	0	82	0	-9.59	160.18	184
25	286	0	80	0	-10.44	162	184
26	285	0	96	1	-5.09	119.28	167
27	299	0	104	1	-5.15	145.82	26, 27, 30, 33
28	170	1	51	0	16.75	-88.08	92, 93
29	269	0	83	0	-6	72	45, 146
30	164	1	46	0	19.74	-87.67	43
31	253	0	69	1	-12.5	177.08	120
33	195	0	31	0	14.71	-16.31	58, 73
34	278	1	19	1	43.37	-8.38	13
35	267	0	58	0	25.63	122.07	190
36	177	0	54	1	12.52	-82.71	128
37	239	2	16	0	20.3	-156.25	62
38	277	0	72	1	21.53	121.35	172, 173
39	275	0	69	1	22.38	121.3	172, 174
40	193	0	28	1	19.87	-16.61	97
41	197	1	23	0	-27.12	-109.37	130, 137
42	339	3	36	1	-33.49	26.17	9
43	449	21	85	8	36.16	133.16	76
44	174	0	12	0	26.68	45.88	15, 16
45	177	0	18	0	7.58	-12.18	8, 73
46	197	0	51	0	-21.15	-159.45	23
47	275	0	41	0	-29.05	167.59	23, 99
48	182	0	62	0	9.33	-79.37	46, 150, 188, 189
49	205	0	51	0	-16.3	-151.45	123
51	171	0	23	0	-23.1	-135	123
52	181	0	39	0	-18.04	-141	123
53	205	0	52	0	-16.45	-151	123
54	199	0	52	0	-15.47	-145.1	123
55	203	0	52	0	-16.1	-148.14	123

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
56	168	0	16	0	-9.22	-140.01	123
57	198	0	51	0	-14.49	-148.34	123
58	204	0	52	0	-17.29	-149.52	123
59	170	0	29	0	-22	-140	123
60	192	0	51	0	-16.4	-142.5	123
61	198	0	52	0	-15	-147.4	123
62	190	0	49	0	-16	-142.25	123
63	198	0	52	0	-15.42	-144.43	123
64	193	0	52	0	-14.35	-145.13	123
65	192	0	51	0	-14.3	-144.58	123
66	203	0	52	0	-17.32	-149.34	123
67	199	0	52	0	-16.45	-144.15	123
68	198	0	51	0	-15	-148.1	123
69	203	0	52	0	-17.23	-150.37	123
70	168	0	15	0	-8.55	-139.32	123
71	199	0	52	0	-17.5	-144.5	123
75	284	0	64	0	-14.66	136.25	126
76	290	1	70	0	-12.62	141.52	126
77	299	0	77	0	-16.93	149.18	100
78	279	0	54	0	-28.31	31.25	41, 86
79	302	7	77	1	-20.29	147.9	127
80	294	0	100	0	-9.3	147.07	28, 29, 102
81	287	2	75	0	-13.96	143.85	60
82	326	7	75	0	-20.85	148.78	60
83	317	6	60	0	-25.36	152.64	60
84	315	7	75	0	-22.93	150.68	60
85	305	6	56	0	-27.16	153.19	60
86	313	7	76	3	-18.95	147	60
87	231	0	56	1	-13.82	-171.75	149
88	229	0	35	0	-3.5	-172.5	152
89	272	0	62	0	-0.52	166.93	154
90	182	0	22	0	16	40	91
92	239	0	65	0	-13.25	40.52	12
93	204	0	55	0	12.3	-70	171
94	209	1	54	0	12.15	-68.27	171
95	230	1	21	3	-35.22	173.95	114
96	160	0	8	0	29.33	34.57	96
97	231	0	56	1	-14	-170.5	148
98	253	0	66	0	-18	179.5	121, 151
99	287	1	85	0	-21.22	165.45	55
100	249	0	20	1	38.3	-28	117, 153
101	275	2	23	1	39.58	-8	153
102	276	0	18	0	42.67	-5.02	153
103	295	9	18	0	46.88	-0.04	153
104	245	0	8	0	54.01	5.93	153

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
106	201	0	8	0	62.31	12.3	153
107	134	0	2	0	67.77	28.77	153
108	289	0	19	0	53.38	-7.76	153
109	286	1	11	0	52.97	-1.53	153
110	254	1	12	0	57.34	-2.13	153
111	227	0	8	0	60.09	-1.09	153
112	186	0	6	0	62.02	-6.47	153
113	108	0	2	0	72.65	5	153
114	78	0	0	0	78.12	15.4	153, 169
115	167	2	4	0	64.63	-17.8	56, 153
116	126	0	2	0	68.8	-32.58	153
117	125	0	2	0	68.22	-56.73	153
118	107	0	1	0	70.3	-68.3	153
119	138	0	2	0	56.28	-61.04	153
120	197	0	6	0	48.84	-54.28	153
121	188	0	6	0	49.65	-61.65	153
122	192	0	6	0	46.57	-66.69	153
123	186	0	6	0	44.07	-68.81	153
124	182	1	7	0	42.54	-70.52	153
125	179	1	7	0	41.56	-70.84	153
126	178	1	7	0	41.14	-71.33	153
127	172	0	7	0	40.83	-72.78	153
128	156	0	7	0	40.02	-74.25	153
129	150	0	7	0	38.84	-75.21	153
130	156	0	7	0	37.92	-75.2	153
131	165	0	9	0	36.98	-75.65	153
132	175	2	4	0	56.28	-132.88	139
133	233	6	7	0	51.19	-126.76	139
134	230	6	7	1	47.09	-123.38	139
135	226	2	6	0	43.79	-123.82	139
136	224	0	7	0	41.39	-124.06	5
137	225	0	11	0	27.7	-113.25	5
138	286	0	36	0	-26.48	113.72	178, 179, 180, 181, 182, 183
139	334	0	39	0	-28.73	114.34	178, 179, 180, 181, 182, 183
140	358	0	41	0	-30.82	115.04	178, 179, 180, 181, 182, 183
141	364	0	45	0	-32.85	115.27	178, 179, 180, 181, 182, 183
142	347	0	44	0	-34.24	115.54	178, 179, 180, 181, 182, 183
143	338	0	41	0	-34.47	116.79	178, 179, 180, 181, 182, 183
144	345	0	42	0	-34.05	118.81	178, 179, 180, 181, 182, 183
145	353	0	44	0	-33.51	120.8	178, 179, 180, 181, 182, 183
146	354	3	44	1	-33.62	122.64	178, 179, 180, 181, 182, 183
147	359	0	44	0	-32.95	124.52	178, 179, 180, 181, 182, 183
148	363	0	44	0	-31.84	126.4	178, 179, 180, 181, 182, 183
149	412	0	46	0	-31.34	128.5	178, 179, 180, 181, 182, 183
150	444	3	56	0	-31.39	130.87	178, 179, 180, 181, 182, 183

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
151	437	2	53	0	-32.05	133.21	178, 179, 180, 181, 182, 183
152	436	16	50	2	-33.47	134.82	178, 179, 180, 181, 182, 183
153	437	0	50	0	-33.47	134.82	178, 179, 180, 181, 182, 183
154	433	2	50	0	-34.8	136.03	178, 179, 180, 181, 182, 183
155	428	1	49	0	-35.37	137.91	178, 179, 180, 181, 182, 183
156	399	0	44	0	-36.91	140.31	178, 179, 180, 181, 182, 183
157	379	0	41	0	-38.32	142.18	178, 179, 180, 181, 182, 183
158	369	0	41	0	-38.29	144	178, 179, 180, 181, 182, 183
159	366	0	39	0	-38.26	145.6	178, 179, 180, 181, 182, 183
160	333	0	35	0	-37.91	147.28	178, 179, 180, 181, 182, 183
161	322	0	33	0	-37.66	147.22	178, 179, 180, 181, 182, 183
162	429	1	49	0	-35.4	137.39	178, 179, 180, 181, 182, 183
163	333	0	32	0	-40.52	146.29	178, 179, 180, 181, 182, 183
164	284	4	24	0	-41.53	147.75	178, 179, 180, 181, 182, 183
165	258	0	21	0	-42.55	146.4	178, 179, 180, 181, 182, 183
166	296	0	23	0	-41.52	144.94	178, 179, 180, 181, 182, 183
167	247	1	23	0	-34.32	20.2	19
168	245	2	23	0	-33.56	18.28	71
169	181	1	7	0	41.2	-70.45	164
170	265	0	78	0	9.1	92.44	72
171	265	0	78	0	8	93.22	72
173	268	0	81	0	7	93.53	72
174	270	0	62	0	22.34	120.36	64
175	387	21	28	8	43.63	142.25	109
176	453	21	88	8	38.21	135.81	191
177	238	0	8	0	54.09	7.52	14
178	232	0	29	4	-35.7	-70.38	131, 138
179	245	2	77	5	18	-66	10
180	176	0	38	4	-14.77	-38.37	40
181	241	0	77	0	-2.99	40.16	31, 146
182	142	0	20	0	19.1	-95.95	85
183	174	0	14	1	27.2	51.63	17
184	187	0	19	1	24.28	54.25	146
185	249	0	79	0	-9.42	46.37	146
186	257	0	86	0	-6	53.17	146
187	107	0	0	0	-37.83	77.52	146
188	258	0	65	0	11.4	92.44	146
189	174	0	13	0	26.05	50.31	146
190	211	0	32	1	21.36	90.63	146
191	250	1	53	0	15.27	95.25	146
192	265	0	79	1	-16.63	59.63	146
193	270	0	71	0	-10.3	105.4	146
194	265	0	67	0	-12.07	96.53	146
195	245	0	75	0	-12.11	44.18	146
196	268	1	82	0	-6.34	72.24	146

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
197	189	0	31	0	11.85	43.1	146
198	269	2	90	8	8.11	77.3	146
199	290	2	83	1	-6.92	110.72	146
200	182	0	18	1	27.6	54.77	146
201	172	0	10	0	29.2	48	146
202	242	0	63	0	9.88	72.75	146
203	268	1	68	3	-18.75	47.24	146
204	272	0	90	0	3.68	101.83	146
205	260	0	83	1	3.03	73.06	146
206	275	1	77	1	-20.14	57.35	146
207	237	1	56	0	-18.4	36.3	146
208	266	0	78	0	8	93.22	146
209	186	0	17	0	25.57	56.28	146
210	208	0	34	2	24.3	64.51	146
211	175	0	14	0	26.03	51.2	146
212	272	0	72	0	-21.08	55.28	146
213	271	0	69	0	-19.75	63.5	146
214	166	0	10	0	22.41	38.38	146
215	268	0	82	0	-10	60	146
216	262	1	90	0	-4.5	55.5	146
217	272	0	94	0	1.32	103.86	146
218	218	0	52	0	5.2	48.3	146
220	269	1	89	3	7.49	80.17	146
221	105	0	0	0	-38.44	77.3	146
222	247	0	78	0	-7.35	39.69	146
223	266	0	80	0	7.99	99.18	146
224	201	2	34	0	15.89	47.46	146
226	246	2	25	1	-31.32	18.14	155
227	142	0	5	0	-47.7	179.08	59
228	153	0	5	0	-48.02	166.57	59
229	133	0	3	0	-50.75	166.17	59
230	138	0	3	0	-49.68	178.8	59
231	115	0	3	0	-52.53	169.13	59
232	269	0	57	0	24	121	87
233	338	3	56	4	-30.41	26.28	145, 146
234	167	0	13	0	-16.39	13.28	73
235	178	0	20	0	10.2	-14	73
236	178	0	22	0	11.6	-15.34	73, 176
237	172	0	17	0	5.42	-9.32	73
238	174	1	15	0	5.57	-0.66	73
239	169	0	14	0	6.18	1.32	73
240	169	0	14	0	6.24	2.31	73
241	166	0	13	0	5.37	5.31	73
242	163	0	13	0	3.36	8.9	73
244	158	0	12	0	-1.5	10.26	73

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
245	163	0	13	0	1.4	9.42	73
246	163	0	13	0	1	7	73
247	163	0	13	0	1.37	7.27	73
248	160	0	12	0	-1.24	5.37	73
249	104	0	1	0	67.5	-100	84
250	28	0	0	0	82.44	-61.3	84
251	31	0	1	0	81.72	-64.64	84
252	52	0	1	0	80.27	-70.2	84
253	57	0	1	0	79.08	-75.3	84
254	56	0	1	0	78.77	-74.77	84
255	71	0	1	0	76.08	-85.55	84
256	65	0	1	0	76.48	-90.42	84
257	65	0	1	0	76.7	-89.52	84
258	66	0	1	0	76.5	-90.5	84
259	81	0	1	0	74.53	-82.45	84
260	92	0	1	0	73.02	-85.15	84
261	96	0	1	0	72.73	-93.8	84
262	96	0	1	0	72.36	-94.13	84
263	103	0	1	0	69.34	-91.3	84
264	96	0	1	0	72.7	-77.85	84
265	97	0	1	0	72.66	-74.44	84
266	114	0	1	0	64.44	-65.27	84
267	117	0	1	0	61.39	-71.11	84
268	108	0	1	0	64.7	-82.2	84
269	108	0	1	0	66.52	-86.23	84
270	110	0	1	0	59.65	-85.16	84
271	110	0	1	0	52.63	-80.49	84
272	104	0	1	0	69.37	-94.86	84
273	106	0	1	0	68.62	-95.88	84
274	106	0	1	0	68.1	-97.57	84
275	106	0	1	0	69.61	-98.46	84
276	101	0	1	0	68.63	-102.91	84
277	94	0	1	0	69.14	-106.17	84
278	96	0	1	0	68.5	-105.72	84
279	97	0	1	0	68.25	-106.9	84
280	97	0	1	0	67.98	-108.82	84
281	82	0	0	0	68.15	-112.78	84
282	71	0	0	0	68.81	-115.08	84
283	59	0	0	0	72.17	-118.48	84
284	58	0	0	0	70.78	-121.19	84
285	4	0	0	0	69.43	-133.05	84
286	2	0	0	0	69.58	-139.03	84
287	23	0	0	0	82.72	-63.34	84
288	26	0	0	0	80.67	-88.6	84
289	45	0	0	0	79.34	-92.88	84

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
290	66	0	1	0	76.86	-89.5	84
291	56	0	0	0	77.18	-95.2	84
292	57	0	0	0	76.74	-96.5	84
293	65	0	0	0	75.94	-96.03	84
294	68	0	1	0	75.75	-92.85	84
295	76	0	1	0	74.56	-94.69	84
296	65	0	0	0	75.34	-97.33	84
297	57	0	0	0	75.39	-102.18	84
298	49	0	0	0	75.89	-106.54	84
299	59	0	0	0	75.34	-105.68	84
300	64	0	0	0	74.03	-106.75	84
301	60	0	0	0	71.85	-113.6	84
302	43	0	0	0	75.82	-111.43	84
303	36	0	0	0	75.51	-118.52	84
304	31	0	0	0	76.15	-119.22	84
305	27	0	0	0	75.73	-121.02	84
306	44	0	0	0	75.58	-110.78	84
307	279	0	70	0	20.25	122	147
309	293	0	96	0	13.54	124.18	147
310	290	0	91	0	15.38	122.53	147
311	294	0	100	0	12.73	120.72	147
312	295	0	99	2	12.21	123.36	147
313	293	0	97	0	11.66	124.89	147
314	286	0	92	0	9.37	118.16	147
315	296	0	100	0	10.82	122.01	147
316	294	0	95	0	10.71	124.4	147
317	293	0	98	0	10.16	123.67	147
318	293	0	96	0	9.48	124.11	147
319	294	0	99	0	9.81	123.11	147
320	292	0	96	0	7.53	124.83	147
321	288	0	96	0	6.4	121.59	147
322	254	0	44	1	32.18	-64.48	142
323	294	4	37	2	28.3	-14.1	57
324	259	1	23	0	32.45	-17	118
325	268	1	12	1	30.72	-114.75	7
326	451	21	86	8	33.5	132.05	108
327	225	0	14	0	39.95	120.8	160
328	294	0	100	0	12.34	120.62	147
331	292	0	99	0	9.13	123.35	147
333	286	0	93	0	5.35	120.38	147
334	296	0	99	0	11	123.5	147
360	287	1	10	0	53.49	-10	133
361	128	0	12	0	-23.02	-42	42
362	268	0	35	0	-25.54	113.35	78
363	361	2	46	1	-32.01	115.28	65, 67

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
364	264	0	35	1	28.12	-177.24	2
365	267	0	38	4	25.69	-169.82	2
366	142	0	11	0	-20.5	-29.32	124
367	316	3	48	0	-31.67	29.43	18
368	295	1	50	0	-29.59	31.03	50
369	321	4	41	1	-32.72	151.29	101, 103, 104, 105
370	301	2	42	0	-31.28	159.09	103, 104, 106
371	307	13	12	3	32.42	-117.7	5, 158
372	216	0	53	0	9.77	-65.48	53
373	253	0	46	0	22.2	114.15	61
374	301	0	85	0	13.5	144.75	95
375	310	0	106	4	7.36	143.21	95
376	205	3	12	1	-40.97	173.11	115
377	181	1	8	0	-44	-176.5	116
378	239	5	19	4	-34.15	172.17	113
379	218	0	50	0	15.32	-61.18	170
380	181	0	58	0	12.4	-81.47	44
381	179	0	56	0	12.17	-81.85	44
382	187	0	60	0	14.27	-80.33	44
383	186	0	60	0	13.57	-80.07	44
384	179	0	40	0	27.5	-82.47	37, 38
385	226	0	51	0	17.08	-61.25	74
386	127	0	11	0	-23.15	-44.23	49
387	280	0	63	0	-12.43	130.85	187
388	218	0	31	1	20.6	92.33	68, 69, 70
389	196	3	20	0	34.23	-77.79	143
390	191	0	19	0	32.73	-79.81	143
391	177	0	19	0	31.12	-81.23	143
392	180	0	29	0	29.21	-80.88	143
393	206	0	64	1	17.9	-77.16	21, 22
394	186	0	60	0	19.32	-80.31	162
395	287	0	77	0	6.95	158.21	63, 98
396	77	0	0	0	-33.91	-53.66	24
397	73	0	0	0	-34.41	-55.8	24
398	280	0	11	0	55.06	-6.8	106
399	312	0	31	0	36.21	127.34	83
400	279	0	11	0	54.76	-5.9	106
401	288	0	11	0	54.23	-5.78	106
402	209	0	20	0	-30.25	-178.5	112
403	332	0	41	0	-28.5	113.82	66
404	156	0	12	0	8.21	-80.1	46
405	172	0	15	0	4.66	-4.87	73
406	270	2	10	0	52.13	4.45	156
407	175	0	51	0	9.94	-82.88	77
409	150	0	9	1	23.2	-106.38	122

Site ID	# Genera	# Endemic Genera	# Bryops species	# Endemic bryops species	Latitude	Longitude	References
410	224	0	7	0	40.67	-124.04	5
412	229	0	7	0	39.24	-123.68	5
413	236	0	7	0	38.34	-122.91	5
414	237	0	7	0	37.83	-122.39	5
415	241	0	7	0	37.31	-122.27	5
416	251	4	8	0	36.82	-121.86	5
417	269	2	9	0	35.92	-121.29	5
418	281	2	9	0	34.98	-120.72	5
419	288	2	10	0	34.41	-119.82	5
420	291	2	12	0	34.17	-119.22	5
421	293	2	12	0	33.78	-118.64	5
422	303	2	12	0	33.01	-118.22	5
423	155	0	16	1	-3.95	-77.3	39
424	224	0	67	0	18.14	-69.91	135
425	187	0	14	0	-21.84	13.29	94
426	173	0	12	0	-18.57	12.24	94
427	185	0	14	0	-21.3	13.65	94
428	185	0	14	0	-24.56	14.7	94
429	294	0	100	0	12.35	120.63	147
430	443	3	56	0	-31.39	130.87	178, 179, 180, 181, 182, 183
431	309	4	12	0	32.23	-116.75	5
432	288	0	82	0	-9.77	160.04	184
433	267	6	42	6	22	-159.5	3
434	268	6	44	8	21.5	-158	3
435	268	6	43	7	21	-157	3
436	260	2	41	5	21	-150	3
437	260	3	41	5	19.5	-156.5	3
438	67	6	1	1	-60	-61	177
439	56	0	0	0	-54	-66	177
440	54	0	0	0	-52	-60	177
441	62	1	0	0	-54	-36	177
442	44	0	0	0	-53	73.5	177
443	51	0	0	0	-54.5	159	177
444	183	1	63	1	22.31	-79.29	159
445	286	0	91	2	9	167	165
446	160	0	13	0	-7.95	-14.37	75
447	157	0	12	0	-15.93	-5.7	75
448	235	1	38	0	-16	-24	75

SPECIES LISTS

1. Abbott, I. A. 1980. Early Collections of Hawaiian Marine-Algae. Pacific Science 34:101-107.

2. Abbott, I. A. 1989. Marine algae of the Northwest Hawaiian Islands. Pacific Science 43:223-233.

3. Abbott, I. A. 1999. Marine Red Algae of the Hawaiian Islands. Bishop Museum Press, Honolulu.

4. Abbott, I. A. 2004. Marine Green and Brown Algae of the Hawaiian Islands. Bishop Museum Press, Honolulu.

5. Abbott, I. A., and G. J. Hollenberg. 1976. Marine Algae of California. Stanford University Press, Stanford.

6. Abbott, I. A., and B. Santelices. 1985. The marine algae of Easter Is (Eastern Polynesia). Pages 71-75 in Fifth International Coral Reef Congress, Tahiti.

7. Aguilar-Rosas, L. E., R. Aguilar-Rosas, A. C. Mendoza-Gonzalez, and L. E. Mateo-Cid. 2000. Marine algae from the northeast coast of Baja California, Mexico. Botanica Marina 43:127-139.

8. Aleem, A. A. 1978. Preliminary List of Marine-Algae from Sierra-Leone. Botanica Marina 21:397-399.

9. Andersen, R. A., and H. Stegenga. 1989. Subtidal algal communities at Bird Is, Eastern Cape, South Africa. Botanica Marina 32:299-311.

10. Ballantine, D. L., and N. E. Aponte. 1997. A revised checklist of the benthic marine algae known to Puerto Rico. Caribbean Journal of Science 33:150-179.

11. Ballantine, D. L., and N. E. Aponte. 2003. An annotated checklist of deep-reef benthic marine algae from Lee Stocking Island, Bahamas (western Atlantic), I. Chlorophyta and Heterokontophyta. Nova Hedwigia 76:113-127.

12. Bandeira, S., C. M. Antonia, and A. T. Critchley. 2001. A taxonomic listing, including new distribution records for benthic intertidal seaweeds from Mecufi, northern Mozambique. South African Journal of Botany 67:492-496.

13. Barbara, I., and J. Cremades. 1996. Seaweeds of the Ria de A Coruna (NW Iberian Peninsula, Spain). Botanica Marina 39:371-388.

14. Bartsch, I., and R. Kuhlenkamp. 2000. The marine macroalgae of Helgoland (North Sea): an annotated list of records between 1845 and 1999. Helgoland Marine Research 54:160-189.

15. Basson, P. W. 1979. Marine-Algae of the Arabian Gulf Coast of Saudi-Arabia (1st-Half). Botanica Marina 22:47-64.

16. Basson, P. W. 1979. Marine-Algae of the Arabian Gulf Coast of Saudi Arabia .2. Botanica Marina 22:65-82.

17. Basson, P. W. 1992. Checklist of Marine-Algae of the Arabian Gulf. Journal of the University of Kuwait-Science 19:217-229.

18. Bolton, J. J., and H. Stegenga. 1987. The marine algae of Hluleka and the warm temperate/sub-tropical transition on the east coast of southern Africa. Helgoland Marine Research 41:165-183.

19. Bolton, J. J., and H. Stegenga. 1990. The seaweeds of DeHoop Nature Reserve and their phytogeographical significance. South African Journal of Botany 56:233-238.

20. Bucher, K. E., and J. N. Norris. 1995. Marine algae new to the Lesser Antilles, including Mazoyerella kraftii sp. nov. (Ceramiales, Rhodophyta). Caribbean Journal of Science 31:1-24.

21. Chapman, V. J. 1961. The Marine Algae of Jamaica Part I. Myxophyceae and Chlorophyceae. The Institute of Jamaica, Kingston, Jamaica.

22. Chapman, V. J. 1963. The Marine Algae of Jamaica Part II. Phaeophyceae and Rhodophyceae. The Institute of Jamaica, Kingston, Jamaica.

23. Chapman, V. J. 1977. Marine-Algae of Norfolk Island and Cook Islands. Botanica Marina 20:161-165.

24. Coll, J. C., and J. C. Oliveira. 1999. The benthic marine algae of Uruguay. Botanica Marina 42:129-135.

25. Compere, P. 1984. Some Algae from the Red-Sea Hills in Northeastern Sudan. Hydrobiologia 110:61-77.

26. Coppejans, E. 1992. Marine algae of Papua New Guinea (Madang Prov.) 2. A revised and completed list of Caulerpa (Chlorophyta - Caulerpales). Blumea 36:382-410.

27. Coppejans, E., O. De Clerck, and F. Leliaert. 2001. Marine brown algae (Phaeophyta) from the north coast of Papua New Guinea, with a description of Dictyita magneana sp. nov. Cryptogamie Algologie 22:15-40.

28. Coppejans, E., O. De Clerck, and C. Van den Heede. 1995. Annotated and illustrated survey of the marine macroalgae from Motupore Island and vicinity (Port Moresby area, Papua New Guinea). I. Chlorophyta. Biologisch Jaarboek Dodonaea 62:70-108.

29. Coppejans, E., O. De Clerck, and C. Van den Heede. 1995. Annotated and illustrated survey of the marine macroalgae from Motupore Island and vicinity (Port Moresby area, Papua New Guinea). II. Phaeophyta. Belgian Journal of Botany 128:176-197.

30. Coppejans, E., F. Leliaert, O. Dargent, and O. De Clerck. 2001. Marine green algae (Chlorophyta) from the north coast of Papua New Guinea. Cryptogamie Algologie 22:375-443.

31. Coppejans, E., F. Leliaert, and O. De Clerck. 2000. Annotated list of new records of marine macroalgae for Kenya and Tanzania since Isaac's and Jaasund's publications. Biologisch Jaarboek Dodonaea 67:31-93.

32. Coppejans, E., F. Leliaert, and T. Schils. 2002. New records of marine benthic algae for the Mozambican coast, collected at Inhaca Is. South African Journal of Botany 68:342-248.

33. Coppejans, E., and A. J. K. Millar. 2000. Marine red algae from the north coast of Papua New Guinea. Botanica Marina 43:315-346.

34. Coppejans, E., and C. Van den Heede. 1996. The Codiales (excl. Codium) (Chlorophyta) from Kenya, Tanzania (Zanzibar) and the Seychelles. Belgian Journal of Botany 129:47-65.

35. Cormaci, M., E. Lanfranco, J. A. Borg, S. Buttigieg, G. Furnari, S. A. Micallef, C. Mifsud, F. Pizzuto, B. Scammacca, and D. Serio. 1997. Contribution to the knowledge of benthic marine algae on rocky substrata of the Maltese Islands (Mediterranean Sea). Botanica Marina 40:203-215.

36. Critchley, A. T., M. E. Aken, S. Bandeira, and M. Kalk. 1997. A revised list of seaweeds from Inhaca Island, Mozambique. South African Journal of Botany 63:426-435.

37. Dawes, C. J. 1974. Marine algae of the west coast of Florida. University of Miami Press, Coral Gables, Florida.

38. Dawes, C. J., and J. F. van Breedveld. 1969. Benthic Marine Algae. Memoirs of the Hourglass Cruises 1:1-30.

39. Dawson, E. Y., C. Acleto, and N. Foldvik. 1964. The seaweeds of Peru. Nova Hedwigia 13:1-111.

40. de Castro, N. J. M. 1998. Checklist of seaweeds from Bahia State, Brazil. Acta Botanica Malacitana 23:5-21.

41. De Clerck, O., H. R. Engledow, J. J. Bolton, R. J. Anderson, and E. Coppejans. 2002. Twenty marine benthic algae new to South Africa, with emphasis on the flora of Kwazulu-Natal. Botanica Marina 45:413-431.

42. de Guimaraens, M. A., and R. Coutinho. 1996. Spatial and temporal variation of benthic marine algae at the Cabo Frio upwelling region, Rio de Janeiro, Brazil. Aquatic Botany 52:283-299.

43. Diaz-Martin, M. A., E. Torres-Mejia, and J. Espinoza-Avalos. 1998. Algae checklist of the Yum Balam protection area, Quintana Roo, Mexico. Revista De Biologia Tropical 46:487-492.

44. Diaz-Pulido, G., and G. Bula-Meyer. 1997. Marine algae from oceanic atolls in the southwestern Caribbean (Albuquerque Cays, Courtown Cays, Serrana Banks and Roncador Bank). Atoll Research Bulletin 448:1-18.

45. Drew, E. 1995. Diversity of the green algal genus Halimeda in the Chagos Archipelago, central Indian Ocean. Aquatic Botany 52:143-150.

46. Earle, S. A. 1972. A review of the marine plants of Panama. Bulletin of the Biological Society Washington 2:69-87.

47. Egerod, L. 1975. Marine Algae of Andaman Sea Coast of Thailand - Chlorophyceae. Botanica Marina 18:41-66.

48. Etcheverry, H. 1986. Algas Marinas Bentonicas de Chile. Instituto de Oceangologia Universidad de Valparaiso, Vina del Mar, Chile.

49. Falcao, C., M. C. Maurat, C. A. G. Nassar, M. T. M. Szechy, and G. J. P. Mitchell. 1992. Benthic marine flora of the northeastern and southeastern coast of Ilha Grande, Rio de Janeiro, Brazil: Phytogeographic considersations. Botanica Marina 35:357-364.

50. Farrell, E., A. T. Critchley, and M. E. Aken. 1993. The intertidal flora of Isipingo Beach, Natal, South Africa and its phycogeographic affinities. Helgoland Marine Research 12:145-160.

51. Frick, H. G., C. F. Boudouresque, and M. Verlaque. 1996. A checklist of marine algae of the Lavezzi Archipelago, with special attention to species rare or new to Corsica (Mediterranean). Nova Hedwigia 62:119-135.

52. Gallardo, T., A. G. Garreta, M. A. Ribera, M. Cormaci, G. Furnari, G. Giaccone, and C. F. Boudouresque. 1993. Checklist of Mediterranean Seaweeds .2. Chlorophyceae Wille S L. Botanica Marina 36:399-421.

53. Ganesan, E. K. 1989. A catalogue of benthic marine algae and seagrasses of Venezuela. Fondo Editorial Conicit, Caracas.

54. Garreta, A. G., T. Gallardo, M. A. Ribera, M. Cormaci, G. Furnari, G. Giaccone, and C. F. Boudouresque. 2001. Checklist of Mediterranean seaweeds. III. Rhodophyceae rabenh. 1. Ceramiales oltm. Botanica Marina 44:425-460.

55. Garrigue, C., and R. T. Tsuda. 1988. Catalogue of marine benthic algae from New Caledonia. Micronesica 21.

56. Gunnarsson, K., and S. Jonsson. 2002. Benthic marine algae of Iceland: revised checklist. Cryptogamie Algologie 23:131-158.

57. Haroun, R. J., M. C. Gil-Rodriguez, J. Diaz de Castro, and W. F. Prud'homme van Reine. 2002. A checklist of the marine plants from the Canary Islands (Central Eastern Atlantic Ocean). Botanica Marina 45:139-169.

58. Harper, J. T., and D. J. Garbary. 1997. Marine algae of northern Senegal: The flora and its biogeography. Botanica Marina 40:129-138.

59. Hay, C. H., N. M. Adams, and C. J. Parsons. 1985. The marine algae of the subantartic islands of New Zealand. A list of species. National Museum of New Zealand Miscellaneous Series 11.

60. Henderson, R. J. F. 2002. Names and Distribution of Queensland Plants, Algae and Lichen. Queensland Herbarium, Queensland Government EPA, Brisbane.

61. Hodgkiss, I. J., and K. Y. Lee. 1983. Hong Kong Seaweeds. Urban Council, Hong Kong.

62. Hodgson, L. M., and I. A. Abbott. 1992. Nearshore Benthic Marine-Algae of Cape Kinau, Maui. Botanica Marina 35:535-540.

63. Hodgson, L. M., and K. J. McDermid. 2000. Marine plants of Pohnpei and Ant Atoll: Chlorophyta, Phaeophyta and Magnoliophyta. Micronesica 32:289-307.

64. Huang, S. F. 1990. The Marine Algal Flora of Hsiao-Liuchiu Island. Botanical Bulletin of Academia Sinica 31:245-255.

65. Huisman, J. M. 1991. Supplement to the catalogue of marine plants recorded from Rottnest Island. Pages 11-18 in F. Wells, D. Walker, H. Kirkman, and R. Lethbridge, editors. Fifth International Marine Biological Workshop: The Marine Flora and Fauna of Rottnest Island, Western Australia. Western Australian Museum, Rottnest Island, Western Australia.

66. Huisman, J. M. 1997. The marine flora and fauna of the Houtman Abrohlos Islands, Western Australia. Pages 177-237 in F. Wells, editor. Seventh International Marine Biological Workshop. Western Australian Museum, Houtman Abrohlos Islands, Western Australia.

67. Huisman, J. M., and D. Walker. 1990. Catalogue of the marine plant of Rottnest Island, Western Australia, with notes on their distribution and biogeography. Kingia 1:349-459.

68. Islam, A. K. M. 1976. Contribution to the Study of the Marine Algae of Bangladesh. J. Cramer, Vaduz, Germany.

69. Islam, A. K. M., and A. Aziz. 1982. Addition to the List of Marine-Algae of St-Martins Island, Bangladesh. 2. Brown, red and blue-green algae. Nova Hedwigia 36:643-656.

70. Islam, A. K. M., and A. Aziz. 1987. Addition to the List of Marine-Algae of St-Martins Island, Bangladesh .3. Red Algae. Nova Hedwigia 45:211-221.

71. Jackelman, J., H. Stegenga, and J. J. Bolton. 1991. The marine benthic flora of the Cape Hangklip area and its phytogeographic affinities. South African Journal of Botany 57:295-304.

72. Jagtap, T. G. 1992. Marine Flora of Nicobar Group of Islands in Andaman Sea. Indian Journal of Marine Sciences 21:56-58.

73. John, D. M., G. W. Lawson, and G. K. Ameke. 2003. The marine macroalgae of the tropical West Africa sub-region. Nova Hedwigia 125:1-213.

74. John, D. M., and J. H. Price. 1979. The marine benthos of Antigua (Lesser Antilles) II. An annotated list of algae species. Botanica Marina 22:327-331.

75. John, D. M., W. F. Prud'homme van Reine, G. W. Lawson, T. B. Kostermans, and J. H. Price. 2004. A taxonomic and geographical catalogue of the seaweeds of the western coast of Africa and adjacent islands. Nova Hedwigia 127:3-139.

76. Kajimura, M. 1987. Deep-Water Flora of Benthic Marine-Algae in the Oki-Islands, Sea of Japan. Botanica Marina 30:373-385.

77. Kemperman, T. C. M., and H. Stegenga. 1986. The marine benthic algae of the Atlantic side of Costa Rica. Brenesia 25-26:99-122.

78. Kendrick, G. A., J. M. Huisman, and D. Walker. 1990. Benthic macroalgae of Shark Bay, Western Australia. Botanica Marina 33:47-54.

79. Khoja, T. M. 1987. New Records of Marine-Algae for the Red-Sea Coast of Saudi-Arabia. Botanica Marina 30:167-176.

80. Kraft, G. T. 2000. Marine and estuarine benthic green algae (Chlorophyta) of Lord Howe Island, south-western Pacific. Australian Systematic Botany 13:509-648.

81. Kraft, G. T., A. J. K. Millar, E. Coppejans, M. Hommersand, L. M. Liao, and F. D. 1999. Marine red algae (Rhodophyta) from Bulusan, Sorsogon Province, Southern Luzon, Philippines. Philippine Scientist 36:1-50.

82. Kuenen, M., and A. O. Debrot. 1995. A quantitative study of the seagrass and algal meadows of the Spaanse Water, Curacao, the Netherlands Antilles. Aquatic Botany 51:291-310.

83. Lee, I. K., and J. W. Kang. 1986. A checklist of marine algae in Korea. Korean Journal of Phycology 1:311-325.

84. Lee, R. K. S. 1990. A catalogue of the marine algae of the Canadian Artic. Natl. Mus. Can. Nat. Sci. Publ. Bot. 9:1-83.

85. Lehman, R. L., and J. W. Tunnell. 1992. Species Composition and Ecology of the Macroalgae of Enmedio Reef, Veracruz, Mexico. Texas Journal of Science 44:445-457.

86. Leliaert, F., O. De Clerck, J. J. Bolton, and E. Coppejans. 2001. New records of the Chlorophyta from South Africa, with the emphasis on the marine benthic flora of KwaZulu-Natal. South African Journal of Marine Science-Suid-Afrikaanse Tydskrif Vir Seewetenskap 67:450-459.

87. Lewis, and Norris. 1987. A history and annotated account of the benthic marine algae of Taiwan. Smithsonian Contribution to the Marine Sciences 29:1-38.

88. Lewis, J. A. 1984. Checklist and bibliography of benthic marine macroalgae recorded from northern Australia I. Rhodophyta. MRL-R-912, Defense Science and Technology Organisation Reports Material Reseach Laboratories, Melbourne.

89. Lewis, J. A. 1985. Checklist and bibliography of benthic marine macroalgae recorded from northern Australia I. Phaeophyta. MRL-R-962, Defense Science and Technology Organisation Reports Material Reseach Laboratories, Melbourne.

90. Lewis, J. A. 1987. Checklist and bibliography of benthic marine macroalgae recorded from northern Australia I. Chlorophyta. MRL-R-1063, Defense Science and Technology Organisation Reports Material Reseach Laboratories, Melbourne.

91. Lipkin, Y., and P. C. Silva. 2002. Marine algae and seagrasses of the Dahlak Archipelago, southern Red Sea. Nova Hedwigia 75:1-90.

92. Littler, D. S., and M. M. Littler. 1997. An illustrated marine flora of the Pelican Cays, Belize. Bulletin of the Biological Society Washington 9:1-149.

93. Littler, D. S., M. M. Littler, and L. Brooks Barrett. 2000. Checklist of marine algae and seagrasses from the ponds of the Pelican Cays, Belize. Atoll Research Bulletin 474:153-206.

94. Lluch, J. R. 2002. Marine benthic algae of Namibia. Scientia Marina 66:5-+.

95. Lobban, C. S., and R. T. Tsuda. 2003. Revised checklist of benthic marine macroalge and seagrasses of Guam and Micronesia. Micronesica 35-36:54-99.

96. Lundberg, B., and N. Popper. 1999. Algal vegetation in the Elat Coral Reef Reserve. Israel Journal of Plant Sciences 47:111-121.

97. Marcotcoqueugniot, J. 1991. A Preliminary List of Marine-Algae from the Banc D-Arguin (Mauritania). Botanica Marina 34:195-199.

98. McDermid, K. J., L. M. Hodgson, and I. A. Abbott. 2002. Marine plants of Pohnpei and Ant Atoll: Rhodophyta with biogeographic comparisons to other Pacific atolls and island groups. Micronesica 34:113-140.

99. Millar, A. J. K. 1999. Marine benthic algae of Norfolk Island, South Pacific. Australian Systematic Botany 12:479-547.

100. Millar, A. J. K. 1999. Marine benthic algae of North East Herald Cay, Coral Sea, South Pacific. Hydrobiologia 399:65-74.

101. Millar, A. J. K. 2004. New records of marine benthic algae from New South Wales, eastern Australia. Phycological Research 52:117-128.

102. Millar, A. J. K., O. De Clerck, E. Coppejans, and L. M. Liao. 1999. Annotated and illustrated survey of the marine macroalgae from Motupore Island and vicinity (Port Moresby area, Papua New Guinea). III. Rhodophyta. Australian Systematic Botany 12:549-591.

103. Millar, A. J. K., and G. T. Kraft. 1993. Catalogue of marine and freshwater red algae (Rhodophyta) of New South Wales, including Lord Howe Is, South-western Pacific. Australian Systematic Botany 6:1-90.

104. Millar, A. J. K., and G. T. Kraft. 1994. Catalogue of marine and brown algae (Phaeophyta) of New South Wales, including Lord Howe Is, South-western Pacific. Australian Systematic Botany 7:1-46.

105. Millar, A. J. K., and G. T. Kraft. 1994. Catalogue of marine and green algae (Chlorophyta) of New South Wales, including Lord Howe Is, South-western Pacific. Australian Systematic Botany 7:419-453.

106. Morton, O. 1994. Marine Algae of Northern Ireland. Ulster Museum, Belfast.

107. Munda, I. M. 2000. Long-term marine floristic changes around Rovinj (Istrian coast, North Adriatic) estimated on the basis of historical data from Paul Kuckuck's field diaries from the end of the 19th century. Nova Hedwigia 71:1-36.

108. Murase, N. T. 1993. Marine algal flora in the eastern waters along the Seto Inland Sea of Yamaguchi Prefecture. Journal of theShimonoseki Unversity of Fisheries 41:237-249.

109. Nabata, S. I. 1991. List of marine algae from the Shiribeshi Coast of Hokkaido Japan. Scientific Reports of Hokkaido Fisheries Experimental Station 36:19-38.

110. Nelson, W. A. 1994. Distribution of Macroalgae in New-Zealand - an Archipelago in-Space and Time. Botanica Marina 37:221-233.

111. Nelson, W. A. 1999. A revised checklist of marine algae naturalised in New Zealand. New Zealand Journal of Botany 37:355-359.

112. Nelson, W. A., and N. M. Adams. 1984. Marine algae of the Kermadec Islands: A list of species. National Museum of New Zealand Miscellaneous Series 10:1- 29.

113. Nelson, W. A., and N. M. Adams. 1985. Marine algae of the Three Kings Islands: A list of species. National Museum of New Zealand Miscellaneous Series 13:1-47.

114. Nelson, W. A., and N. M. Adams. 1987. Marine algae of the Bay of Islands area: A list of species. National Museum of New Zealand Miscellaneous Series 16:1-47.

115. Nelson, W. A., N. M. Adams, and J. M. Fox. 1992. Marine algae of the northern South Island: A list of species. National Museum of New Zealand Miscellaneous Series 26:1-67.

116. Nelson, W. A., N. M. Adams, and C. H. Hay. 1991. Marine algae of the Chatham Islands: A list of species. National Museum of New Zealand Miscellaneous Series 23:1-58.

117. Neto, A. 1994. Checklist of the benthic marine macroalgae of the Azores. Arquipelago Ciencias Da Natureza 0:15-34.

118. Neto, A., D. C. Cravo, and R. J. Haroun. 2001. Checklist of the benthic marine plants of the Madeira Archipeligo. Botanica Marina 44:391-414.

119. Nizamuddin, M., J. A. West, and E. G. Menez. 1979. List of Marine-Algae from Libya. Botanica Marina 22:465-476.

120. Nyeurt, A. D. 1996. A preliminary floristic survey of the benthic marine algae of Rotuma island. Australian Systematic Botany 9:361-490.

121. N'Yeurt, A. D. 2001. Marine algae from the Suva Lagoon and Reef, Fiji. Australian Systematic Botany 14:689-869.

122. Ochoa-Izaguirre, M. J., J. L. Carballo, and F. Paez-Osuna. 2002. Qualitative changes in macroalgal assemblages under two contrasting climatic conditions in a subtropical estuary. Botanica Marina 45:130-138.

123. Payri, C. E., and A. D. R. N'Yeurt. 1997. A revised checklist of Polynesian benthic marine algae. Australian Systematic Botany 10:867-905.

124. Pedrini, A. G., J. E. A. Goncalves, M. C. S. Fonseca, A. S. Zau, and S. Lacorte. 1989. A survey of the marine algae of Tridade Island, Brazil. Botanica Marina 32:97-99.

125. Phang, S. M. 1994. New Records of Malaysian Marine-Algae. Hydrobiologia 285:123-129.

126. Phillips, J. A., C. Conacher, and J. Horrocks. 1999. Marine macroalgae from the Gulf of Carpentaria, tropical northern Australia. Australian Systematic Botany 12:449-478.

127. Phillips, J. A., and I. R. Price. 1997. A catalogue of Phaeophyta (brown algae) from Queensland, Australia. Australian Systematic Botany 10:683-721.

128. Phillips, R. C., R. L. Vadas, and N. Ogden. 1982. The marine algae and seagrasses of the Miskito Bank, Nigaragua. Aquatic Botany 13:187-195.

129. Prud'homme van Reine, W. F., E. Verheij, and E. Coppejans. 1996. Species and ecads of Caulerpa (Ulvophyceae, Chlorophyta) in Malesia (S.E. Asia): taxonomy, biogeography and biodiversity. Netherlands Journal of Aquatic Ecology 30:83-98.

130. Ramirez, M. E., and D. G. Muller. 1991. New Records of Benthic Marine-Algae from Easter Island. Botanica Marina 34:133-137.

131. Ramirez, M. E., and B. Santelices. 1991. Catalogo de las algas marinas bentonicas de la costa temperada del Pacifico de Sudamerica. Facultad de Ciencias Bioloicas Pontificia Univerisidad Catolica de Chile, Santiago.

132. Ribera, M. A., A. G. Garreta, T. Gallardo, M. Cormaci, G. Furnari, and G. Giaccone. 1992. Checklist of Mediterranean Seaweeds .1. Fucophyceae (Warming, 1884). Botanica Marina 35:109-130.

133. Rindi, F., and M. D. Guiry. 2004. A long-term comparison of the benthic marine algal flora of Clare Island, County Mayo, western Ireland. Biodiversity and Conservation 13:471-492.

134. Rindi, F., G. Sartoni, and F. Cinelli. 2002. A floristic account of the benthic marine algae of Tuscany (Western Mediterranean Sea). Nova Hedwigia 74:201-250.

135. Rosado, G. A. 1993. Las algas marinas de la Republica Dominicana. Pages 404-414 in Biodiversidad Marina y Costera en la Republica Dominicana. Centro de Investigaciones de Biologia Marina, Fondo Mundial para la Naturaleza, Santo Domingo.

136. Santelices, B. 1989. Algas Marinas de Chile: Distribucion, Ecologia, Utilizacion, Diversidad. Ediciones Universidad Catolica de Chile, Santiago.

137. Santelices, B., and I. A. Abbott. 1987. Geographic and marine isolation: An assessment of the marine algae of Easter Island. Pacific Science 41.

138. Santelices, B., M. E. Ramirez, and I. A. Abbott. 1989. A New Species and New Records of Marine-Algae from Chile. British Phycological Journal 24:73-82.

139. Scagel, R. F., P. W. Gabrielson, D. Garbary, L. Golden, M. W. Hawkes, S. C. Lindstrom, J. C. Oliveira, and T. B. Widdowson. 1988. A Synopsis of the Benthic Marine Algae of British Columbia, Southeast Alaska, Washington and Oregon. Department of Botany, University of British Columbia, Vancouver.

140. Schils, T., and E. Coppejans. 2003. Spatial variation in subtidal plant communities around the Socotra Archipelago and their biogeographic affinities within the Indian Ocean. Marine Ecology-Progress Series 251:103-114.

141. Schils, T., O. De Clerck, F. Leliaert, J. J. Bolton, and E. Coppejans. 2001. The changes in macroalgal assemblages through the Saldanha Bay/Langebaan Lagoon Ecosystem (South Africa). Botanica Marina 44:295-305.

142. Schneider, C. W. 2003. An annotated checklist and bibliography of the marine macroalgae of the Bermuda Islands. Nova Hedwigia 76:275-361.

143. Schneider, C. W., and R. B. Searles. 1991. Seaweeds of the Southeastern United States, Cape Hatteras to Cape Canaveral. Duke University Press, Durham.

144. Seagrief, S. C. 1980. Seaweeds of Maputaland. Pages 35-72 in M. Bruton and K. Cooper, editors. Studies on the Ecology of Maputaland. Durban Rhodes University and the Wildlife Society, Durban.

145. Seagrief, S. C. 1984. A catalogue of the South African green, brown and red marine algae. Mem. Bot. Soc. S. Afr. 47:1-72.

146. Silva, P. C., P. W. Basson, and R. L. Moe. 1996. Catalogue of the Benthic Marine Algae of the Indian Ocean. University of California Press, Berkley.

147. Silva, P. C., E. H. Menez, and R. L. Moe. 1987. Catalogue of the benthic algae of the Philippines. Smithsonian Contribution to the Marine Sciences 27:1-179.

148. Skelton, P. A., and G. R. South. 1999. A Preliminary Checklist of the Benthic Marine Algae of the Samoan Archipelago. 1, Institute of Marine Resources, The University of the South Pacific, Suva, Fiji.

149. Skelton, P. A., and G. R. South. 2002. Annotated catalogue of the benthic marine algae of the Palolo Deep National Marine Reserve of Samoa. Australian Systematic Botany 15:135-179.

150. Soto, R. H., Z. Guzman, J. Pinzon, C. Moreno, H. Gamboa, M. Montoya, and R. Vargas. 1998. Evaluacion ecologica rapida del Parque Nacional Marino Isla Bastimentos y areas adyacentes, provinvia de Bocas del Toro. Tomo 3: Recursos costero-marinos. Asociacion Nacional para la Conservacion de la Naturaleza, Panama.

151. South, G. R. 1992. Contributions to a Catalogue of the Marine Algae of Fiji 1. Halimeda (Chlorophyaceae). 2, Institute of Marine Resources, The University of the South Pacific, Suva, Fiji.

152. South, G. R., P. A. Skelton, and A. Yoshinaga. 2001. Subtidal benthic marine algae of the Phoenix Islands, Republic of Kiribati, Central Pacific. Botanica Marina 44:559-570.

153. South, G. R., and I. Tittley. 1986. A Checklist and Distributional Index of the Benthic Marine Algae of the Northern Atlantic Ocean. Huntsman Marine Laboratories and British Museum (Natural History), St Andrews and London.

154. South, G. R., and S. Yen. 1992. Notes on the benthic marine algae of Nauru, Central Pacific. Micronesica 25:123-131.

155. Stegenga, H., J. J. Bolton, and R. A. Andersen. 1997. Seaweeds of the South African West. Coast Creda Press, Cape Town.

156. Stegenga, H., I. Mol, W. F. Prud'homme van Reine, and G. M. Lokhorst. 1997. Checklist of the marine algae of the Netherlands. Gorteria Suppl. 4:3-57.

157. Stegenga, H., and M. Vroman. 1988. Additions to the marine algal flora of Curacao, Netherlands Antilles. Blumea 33:299-311.

158. Stewart, J. G. 1991. Marine algae and seagrasses of San Diego County. California Sea Grant College, La Jolla.

159. Suarez, A. M. 1973. Catalogo de Algas Cubanas. Centro de Informacion Cientifica y Technica, Universidad de la Habana, Habana, Cuba.

160. Sui, Z. Y., and J. Fu. 1995. Preliminary study of the benthic marine algae on the norther Bohai Sea coast. Acta Botanica Sinica 37:394-400.

161. Taylor, W. R. 1969. Notes on the distribution of West Indian marine algae particularly in the Lesser Antilles. Contributions from the University of Michigan Herbarium 9:125-203.

162. Tittley, I. 1994. Marine algae of the Cayman Islands: a preliminary account. Pages 125-130 in M. A. Brunt and J. E. Davies, editors. The Cayman Islands: Natural History and Biogeography. Kluwer Academic Publishers, Amsterdam.

163. Tittley, I., A. Neto, W. F. Farnham, and M. I. Parente. 2001. Additions to the marine algal (seaweed) flora of the Azores. Botanica Marina 44:215-220.

164. Treat, R., A. R. Keith, and R. T. Wilce. 2003. Preliminary checklist of marine algae from Martha's Vineyard, Massachusetts. Rhodora 105:54-75.

165. Tsuda, R. T. 2002. Checklist and bibliography of the marine benthic algae from the Marshall Islands. Technical Report Technical Report No. 106, University of Guam Marine Laboratory, Managilo, Guam.

166. Tsuda, R. T., and F. O. Wray. 1977. Bibliography of marine benthic algae in Micronesia. Micronesica 13:85-120.

167. Verheij, E., and W. F. P. Vanreine. 1993. Seaweeds of the Spermonde Archipelago, Sw Sulawesi, Indonesia. Blumea 37:385-510.

168. Verlaque, M. 2001. Checklist of the macroalgae of Thau Lagoon (Herault, France), a hot spot of marine species introduction in Europe. Oceanologica Acta 24:29-49.

169. Vinogradova, K. L. 1995. The checklist of the marine algae from Spitzenbergen. Botanicheskia Zhurnal 80:50-61.

170. Vroman, M. 1968. Studies on the Flora of Curacao and Other Caribbean Islands: The Marine Algal Vegetation of St Martin, St Eustatius and Saba. Natuurwetenschappelijke Studiekring voor Suriname en de Nederlandse Antillen, Utrect.

171. Vroman, M., and H. Stegenga. 1988. An Annotated Checklist of the Marine-Algae of the Caribbean Islands Aruba and Bonaire. Nova Hedwigia 46:433-480.

172. Wang, W. L., H. K. Chen, and Y. M. Chiang. 1993. Notes on Marine-Algae New to Taiwan. Botanica Marina 36:545-553.

173. Wang, W. L., and Y. M. Chiang. 1993. Marine algae of Lan Yu (Orchid Island), Taiwan. Acta Oceanographica Taiwanica 0:83-100.

174. Wang, W. L., and Y. M. Chiang. 2001. The marine macroalgae of Lu Tao (Green Island), Taiwan. Taiwania 46:49-61.

175. Wei, T. L., and W. Y. Chin. 1983. Seaweeds of Singapore. Singapore University Press, Singapore.

176. Welten, M. C. M., P. A. J. Audiffred, and W. F. Prud'homme van Reine. 2002. Notes on marine algae collected in Guinea-Bissau, Tropical West Africa. Botanica Marina 45:380-384.

177. Wiencke, C., and M. N. Clayton. 2002. Antartic Seaweeds. ARG Gantner Verlag KG, Ruggell / Lichtenstein.

178. Womersley, H. B. S. 1984. The Benthic Flora of Southern Australia. Part I. Government Printer, Adelaide.

179. Womersley, H. B. S. 1987. The Benthic Flora of Southern Australia. Part II. Government Printer, Adelaide.

180. Womersley, H. B. S. 1994. The Benthic Flora of Southern Australia. Part IIIA. Australian Biological Resources Study, Canberra.

181. Womersley, H. B. S. 1996. The Benthic Flora of Southern Australia. Part IIIB. Australian Biological Resources Study, Canberra.

182. Womersley, H. B. S. 1998. The Benthic Flora of Southern Australia. Part IIIC. Australian Biological Resources Study, Canberra.

183. Womersley, H. B. S. 2003. The Benthic Flora of Southern Australia. Part IIID. Australian Biological Resources Study, Canberra.

184. Womersly, H. B., and A. Bailey. 1970. Marine Algae of Solomon Islands. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 259:257-&.

185. Wynne, M. J. 1986. A Checklist of Benthic Marine-Algae of the Tropical and Subtropical Western Atlantic. Canadian Journal of Botany-Revue Canadienne De Botanique 64:2239-2281.

186. Wynne, M. J. 2001. New records of benthic marine algae from the Sultanate of Oman, northern Arabian Sea. II. Nova Hedwigia 72:347-374.

187. Wynne, M. J., and J. Luong-Van Thinh. 1996. A report on collections of benthic marine algae from Darwin, northern Australia. in Proceedings #79 Northern Territory University. Australian Centre for International Agricultural Research, Canberra.

188. Wysor, B., and O. De Clerck. 2003. An updated and annotated list of marine brown algae (Phaeophyceae) of the Carribean Coast of the Republic of Panama. Botanica Marina 46:151-160.

189. Wysor, B., and W. H. C. F. Kooistra. 2003. An annotated list of marine Chlorophyta from the Caribbean coast of the Republic of Panama. Nova Hedwigia 77:487-523.

190. Yang, H. N., W. L. Wang, and L. M. Liao. 1994. Marine Algal Flora of Pengchia-Yu and Its Special Place in the Marine Phytogeography of Taiwan. Botanica Marina 37:429-432.

191. Yoshida, T., A. Yoshinaga, and Y. Nakajima. 2000. Check list of marine algae of Japan. Japanese Journal of Phycology 48:113-169.

APPENDIX 2: R CODE

CHAPTER 3 CODE

```
# Script: Drivers of Bryopsidales diversity (Chapter 3)
# Software: R2.3.0 (available at http://cran.au.r-project.org/)

# Author: Ailsa Kerswell
# Date: 1st Jan 2006

## CODE to test correlation between climate variables and Bryops
## richness using all sites in the Indo-Pacfic between 30oS-30oN (Chapter 3)

## Run models (likfit) incoporating spatial autocorrelation for all
## variables and for all combos of variables

## For the analyses save a table of the AICs. Employ formulae to
## calculate wAICs and then chose the best model

## For all the best models calculate the fitted values (save betas from
## each of the best models for future reference and use).
## Plot fitted versus observed and calculate % variation explained by
## models


rm(list=ls())

setwd("C:\\Ailsa's Documents\\correlates of diversity\\Bryops analyses")

library(geoR)
options(scipen=4)
options(digits=2)


# Function to calculate wAIC

calc.waic = function(aic.vector){
   waic = rep(NA, length(aic.vector))
   expaic = rep(NA, length(aic.vector))
   for (i in 1:length(aic.vector)){
      minaic = min(aic.vector)
      deltaaic = aic.vector[i]-minaic
      expaic[i] = exp(-0.5*deltaaic)
   }
   sumexp = sum(expaic)
   waic = expaic/sumexp
   return(waic)
}


## INPUT DATA & DEFINE VARIABLES

in.data<-read.table("bryops_data_minus4.txt",header=TRUE,sep="\t")
attach(in.data)
data=in.data[order(in.data$lat),]
```

```
raw.bry=data$bryops
bry=(raw.bry-mean(raw.bry))/sd(raw.bry)
   # normalised richness

lat=data$lat
long=data$gis_long
norm.lat = (lat+30)/60
norm.long = (long+169)/249
   # normalises lat and long to b/w 0-1

raw.temp=data$temp
temp=(raw.temp-mean(raw.temp))/sd(raw.temp)
   # normalised temp

area10=log10(data$area10)
area=(area10-mean(area10))/sd(area10)
   # normalised area

mde2 = (norm.long-0.5)^2+(norm.lat-0.5)^2
   # uses pythag to get sq dist from pts to mid-domain
mde=sqrt(mde2)


raw.light_winter=data$light_winter
light_winter=(raw.light_winter-mean(raw.light_winter))/sd(raw.light_winter)
   # normalised light_winter

raw.light_summer=data$light_summer
light_summer=(raw.light_summer-mean(raw.light_summer))/sd(raw.light_summer)
   # normalised light_summer

raw.light_range=data$light_summer - data$light_winter
light_range=(raw.light_range-mean(raw.light_range))/sd(raw.light_range)
   # normalised light_range

raw.nitrate=log10(data$nitrate)
nitrate=(raw.nitrate-mean(raw.nitrate))/sd(raw.nitrate)
   # normalised nitrate

raw.chla=log10(data$chla)
chla=(raw.chla-mean(raw.chla))/sd(raw.chla)
   # normalised chla


par(mfrow=c(3,2))
par(pty='s')
plot(raw.temp, raw.nitrate)
plot(area10, raw.nitrate)
plot(raw.temp, area10)
plot(area10, mde)
plot(raw.temp, mde)
plot(mde, raw.nitrate)

cor(raw.temp, raw.nitrate)
cor(area10, raw.nitrate)
cor(raw.temp, area10)
```

```
cor(area10, mde)
cor(raw.temp, mde)
cor(mde, raw.nitrate)


# RUN SPATIAL REGRESSION MODELS

## BRYOPSIDALE DATA MODELS WITH ALL VAR COMBOS

long.lat = cbind(data$long, data$lat)

geo.algae = list(coords = long.lat, data=raw.bry, temp = temp,
                 area=area, mde=mde, nitrate=nitrate, chla=chla,
                  light_range=light_range)
   # List with site coords, algae richness data and environmental variables

ini.vals  =  c(var(geo.algae$data),1)
   # Initial conditions for the spatial autocorrelation parameters

## Maximum likelihood fitting of alternative models (including fits for the non-spatial model)
## Spatial covariances are assumed to decrease exponentially with distance.

## 1.  No trend (i.e. y = beta):
no.trend.ml  =  likfit(geo.algae,ini=ini.vals,cov.model="exp",nospatial=TRUE, messages=FALSE)

## 2.  Enviro variables (and their combinations):
temp.ml  =  likfit(geo.algae,trend= ~ temp,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
light_range.ml  =  likfit(geo.algae,trend= ~ light_range,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
nitrate.ml  =  likfit(geo.algae,trend= ~ nitrate,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
chla.ml  =  likfit(geo.algae,trend= ~ chla,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)

area.ml  =  likfit(geo.algae,trend= ~ area,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
mde.ml  =  likfit(geo.algae,trend= ~ mde,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)

TA.ml  =  likfit(geo.algae,trend= ~ temp+area,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
TM.ml  =  likfit(geo.algae,trend= ~ temp+mde,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
TN.ml  =  likfit(geo.algae,trend= ~ temp+nitrate,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
AM.ml  =  likfit(geo.algae,trend= ~ area+mde,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
AN.ml  =  likfit(geo.algae,trend= ~ area+nitrate,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
MN.ml  =  likfit(geo.algae,trend= ~ mde+nitrate,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)

TAM.ml  =  likfit(geo.algae,trend= ~ temp+area+mde,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
TAN.ml  =  likfit(geo.algae,trend= ~ temp+area+nitrate,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
```

```
TMN.ml  =  likfit(geo.algae,trend= ~ temp+mde+nitrate,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
AMN.ml  =  likfit(geo.algae,trend= ~ area+mde+nitrate,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)

all.ml  =  likfit(geo.algae,trend= ~ temp+area+mde+nitrate,ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)


aics_energy = c(no.trend.ml$AIC,
temp.ml$AIC,
light_range.ml$AIC,
nitrate.ml$AIC,
chla.ml$AIC)

calc.waic(aics_energy)

aics = c(no.trend.ml$AIC,
temp.ml$AIC,
nitrate.ml$AIC,
area.ml$AIC,
mde.ml$AIC,
TA.ml$AIC,
TM.ml$AIC,
TN.ml$AIC,
AM.ml$AIC,
AN.ml$AIC,
MN.ml$AIC,
TAM.ml$AIC,
TAN.ml$AIC,
TMN.ml$AIC,
AMN.ml$AIC,
all.ml$AIC)

waics = calc.waic(aics)
   # calculate wAIC from above aics

models = c('no.trend', 'temp', 'nitrate', 'area', 'mde', 'TA', 'TM', 'TN',
           'AM', 'AN', 'MN', 'TAM', 'TAN', 'TMN', 'AMN', 'all')

cbind(models, waics)

## CHECK 2nd ORDER TERMS

temp2.ml  =  likfit(geo.algae,trend= ~ temp+I(temp^2),ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
nitrate2.ml  =  likfit(geo.algae,trend= ~ nitrate+I(nitrate^2),ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
area2.ml  =  likfit(geo.algae,trend= ~ area+I(area^2),ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
mde2.ml  =  likfit(geo.algae,trend= ~ mde+I(mde^2),ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)


temp2alt.ml  =  likfit(geo.algae,trend= ~ I(temp^2),ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
nitrate2alt.ml  =  likfit(geo.algae,trend= ~ I(nitrate^2),ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
```

```
area2alt.ml  =  likfit(geo.algae,trend= ~ I(area^2),ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)
mde2alt.ml  =  likfit(geo.algae,trend= ~ I(mde^2),ini=ini.vals,cov.model="exp",nospatial=TRUE,
messages=FALSE)


temp.aics = c(temp.ml$AIC, temp2.ml$AIC, temp2alt.ml$AIC)
nitrate.aics = c(nitrate.ml$AIC, nitrate2.ml$AIC, nitrate2alt.ml$AIC)
area.aics = c(area.ml$AIC, area2.ml$AIC, area2alt.ml$AIC)
mde.aics = c(mde.ml$AIC, mde2.ml$AIC, mde2alt.ml$AIC)



## COMPARE AICs from spatial and non-spatial models

no.trend.aic.ns = -2*(no.trend.ml$nospatial$loglik.ns)+2*no.trend.ml$nospatial$npars.ns
temp.aic.ns = -2*(temp.ml$nospatial$loglik.ns)+2*temp.ml$nospatial$npars.ns
nitrate.aic.ns = -2*(nitrate.ml$nospatial$loglik.ns)+2*nitrate.ml$nospatial$npars.ns
area.aic.ns = -2*(area.ml$nospatial$loglik.ns)+2*area.ml$nospatial$npars.ns
mde.aic.ns = -2*(mde.ml$nospatial$loglik.ns)+2*mde.ml$nospatial$npars.ns

TA.aic.ns = -2*(TA.ml$nospatial$loglik.ns)+2*TA.ml$nospatial$npars.ns
TM.aic.ns = -2*(TM.ml$nospatial$loglik.ns)+2*TM.ml$nospatial$npars.ns
TN.aic.ns = -2*(TN.ml$nospatial$loglik.ns)+2*TN.ml$nospatial$npars.ns
AM.aic.ns = -2*(AM.ml$nospatial$loglik.ns)+2*AM.ml$nospatial$npars.ns
AN.aic.ns = -2*(AN.ml$nospatial$loglik.ns)+2*AN.ml$nospatial$npars.ns
MN.aic.ns = -2*(MN.ml$nospatial$loglik.ns)+2*MN.ml$nospatial$npars.ns

TAM.aic.ns = -2*(TAM.ml$nospatial$loglik.ns)+2*TAM.ml$nospatial$npars.ns
TAN.aic.ns = -2*(TAN.ml$nospatial$loglik.ns)+2*TAN.ml$nospatial$npars.ns
TMN.aic.ns = -2*(TMN.ml$nospatial$loglik.ns)+2*TMN.ml$nospatial$npars.ns
AMN.aic.ns = -2*(AMN.ml$nospatial$loglik.ns)+2*AMN.ml$nospatial$npars.ns

all.aic.ns = -2*(all.ml$nospatial$loglik.ns)+2*all.ml$nospatial$npars.ns

AIC.nospat = c(no.trend.aic.ns, temp.aic.ns, nitrate.aic.ns, area.aic.ns, mde.aic.ns,
                  TA.aic.ns, TM.aic.ns, TN.aic.ns, AM.aic.ns, AN.aic.ns, MN.aic.ns,
                 TAM.aic.ns, TAN.aic.ns, TMN.aic.ns, AMN.aic.ns, all.aic.ns)
   # AICs for models without spatial component

compare.aics = cbind(models, AIC.nospat, aics)
   # spatial models always have much lower AICs



## PREDICTED VALUES FOR SOME MODELS

temp.pred = temp.ml$beta[1]+temp.ml$beta[2]*temp

area.pred = area.ml$beta[1]+area.ml$beta[2]*area

mde.pred = mde.ml$beta[1]+mde.ml$beta[2]*mde

nitrate.pred = nitrate.ml$beta[1]+nitrate.ml$beta[2]*nitrate

TN.pred = TN.ml$beta[1]+TN.ml$beta[2]*temp+TN.ml$beta[3]*nitrate
```

```
TMN.pred = TMN.ml$beta[1]+TMN.ml$beta[2]*temp+TMN.ml$beta[3]*mde+TMN.ml$beta[4]*nitrate

all.pred =
all.ml$beta[1]+all.ml$beta[2]*temp+all.ml$beta[3]*area+all.ml$beta[4]*mde+all.ml$beta[5]*nitrate


## Plot observed versus predicted from the best model
## (in this case best model is TMN.ml)

par(mfrow=c(1,2))
par(pty='s')

plot(TMN.pred, raw.bry, ylim=c(0,110), xlim=c(0,110), ylab='observed',
     xlab='predicted', cex=1.5)
abline(0,1)

TMN.resid = raw.bry-TMN.pred

plot(TMN.pred, TMN.resid, ylab='residuals', xlab='fitted')
abline(0,0)


# Plot some of the results without MDEs included

par(mfrow=c(2,2))
par(pty='s')

plot(TMN.pred, raw.bry, ylim=c(0,110), xlim=c(0,110))
plot(TN.pred, raw.bry, ylim=c(0,110), xlim=c(20,60))
plot(temp.pred, raw.bry, ylim=c(0,110), xlim=c(20,60))
plot(nitrate.pred, raw.bry, ylim=c(0,110), xlim=c(20,60))



## MDE ANALYSES
# Data here are constrained to be within the domain
# i.e. lats and longs only between 30-30o lat

ends.data<-read.table("bryops_ip_endpoints.txt",header=TRUE,sep=",")
attach(ends.data)

min.lat = ends.data$min.lat
max.lat = ends.data$max.lat
min.long = ends.data$min.long
max.long = ends.data$max.long
   # range endpoints

lat.range.deg = max.lat-min.lat
long.range.deg = max.long-min.long
   # range extents in degrees

zero=lat.range.deg==0
lat.range.deg[zero]=1
zero=long.range.deg==0
long.range.deg[zero]=1
   # give endemics range size of 1 degree lat and long

dom.lat = max(max.lat)-min(min.lat)
```

```
dom.long = max(max.long)-min(min.long)
   # max size of domain

lat.range = lat.range.deg/dom.lat
long.range = long.range.deg/dom.long
   # range extent normalised to b/w 0-1


# Function to calculate prob of overlap of site x by spp
# with range size r from SPREADING DYE model
# Based on Connolly 2005 eqn 2

spread.dye = function(x, r){

   oneminr = 1-r

   if(oneminr<=x & x<=r){
      Pr = 1
      }else{

      if(r<=x & x<=oneminr){
         Pr = r
         }else{

         if(x<min(r, oneminr)){
            Pr = x+(r/2)
            }else{

            if(x>max(r, oneminr)){
               Pr = 1-x+(r/2)
               }
            }
         }
      }
   return(Pr)
}



# Need to run spread.dye for all spp over all domain locations

N = length(min.lat)
   # number of species in the domain

L = 100
   # number of locations on domain
locs = seq(0,1,1/L)
   # locations on the domain

## For latitude:

lat.SD.mat = matrix(NA, nrow=N, ncol=(L+1))

for(l in 1:(L+1)){
   for(n in 1:N){
      prob = spread.dye(locs[l], lat.range[n])
      lat.SD.mat[n,l] = prob
   }
```

```
}
lat.SD.rich = colSums(lat.SD.mat)


## For longitude:

long.SD.mat = matrix(NA, nrow=N, ncol=(L+1))

for(l in 1:(L+1)){
   for(n in 1:N){
      prob = spread.dye(locs[l], long.range[n])
      long.SD.mat[n,l] = prob
   }
}
long.SD.rich = colSums(long.SD.mat)



par(mfrow=c(1,2))
par(pty='s')

plot(locs, lat.SD.rich, type='l', lwd=2)
plot(locs, long.SD.rich, type='l', lwd=2)


par(mfrow=c(1,2))
par(pty='s')
plot(norm.lat, mde.pred, cex=1.25,ylim=c(0,140))
lines(locs, lat.SD.rich, lwd=2)

plot(norm.long, mde.pred, cex=1.25,ylim=c(0,120))
lines(locs, long.SD.rich, lwd=2)


mde.seq = abs(locs-0.5)
mde.my.est = mde.ml$beta[1]+mde.ml$beta[2]*mde.seq
plot(locs, mde.my.est)
   # calculates 1-d mde estimate based on my equations
   # developed in 2-d



# Boekma 2-d fully stochastic

p = norm.lat
q = 1-norm.lat

r = norm.long
t = 1-norm.long

S = 290
   # richness of total species pool

exp.rich = 4*p*q*r*t*N

plot(norm.lat, exp.rich)
plot(norm.long, exp.rich)
```

```
# Want to get mde estimate from Poisson model
# using Connolly 2005 eqn 9
# therefore need to first estimate lambda from mde predictions
# versus norm lat and long data using least squares

library(nlme)

lat.gls=gls(mde.pred~norm.lat+I(norm.lat^2))
coeffs.lat = lat.gls$coeff

long.gls=gls(mde.pred~norm.long+I(norm.long^2))
coeffs.long = long.gls$coeff


poisson = function(x, coeffs){
   lambda =
   Pr = (2-exp(-lambda*x)-exp(-lambda*(1-x)))/lambda
   return(Pr)
}


lat.poisson = rep(NA, length(locs))

for(l in 1:length(locs)){
   prob = poisson(locs[l], coeffs.lat)
   lat.poisson[l] = prob
}


long.poisson = rep(NA, length(locs))

for(l in 1:length(locs)){
   prob = poisson(locs[l], lambda.long)
   long.poisson[l] = prob
}


par(mfrow=c(1,2))
par(pty='s')
plot(locs, lat.poisson*N, type='l')
plot(locs, long.poisson*N, type='l')



windows()

par(mfrow=c(1,2))
par(pty='s')
plot(norm.lat, mde.pred, cex=1.25,pch=19, ylim=c(0,100))
points(norm.lat, exp.rich)
lines(locs, lat.SD.rich, lwd=2)
lines(locs, lat.poisson*N, col='red', lwd=2)

plot(norm.long, mde.pred, cex=1.25, pch=19, ylim=c(0,120))
points(norm.long, exp.rich)
lines(locs, long.SD.rich, lwd=2)
```

```
lines(locs, long.poisson*N, col='red', lwd=2)



par(mfrow=c(1,3))
par(pty='s')
plot(norm.lat, mde.pred, cex=1.25, ylim=c(0,100))
lines(norm.lat, lat.gls$fitted, lwd=2)

plot(norm.long, mde.pred, cex=1.25, ylim=c(0,100))
plot(norm.long, long.gls$fitted, cex=1.25, ylim=c(0,100))



## CHECK AUTOCORRELATION CAPTURE

## To check how well spatial autocorrelation captured plot variogram
## of residuals as well as fitted line from predicted varigram

par(mfrow=c(2,2))
par(pty='s')

temp.ml.resid = geo.algae$data - temp.pred
   # Residuals of trend part of model
plot(variog(coords=geo.algae$coords,data=temp.ml.resid), pch=19, cex=1.25)
   # Variogram of residuals of trend part
lines(temp.ml, lwd=2)
   # Predicted (fitted) variogram


area.ml.resid = geo.algae$data - area.pred
plot(variog(coords=geo.algae$coords,data=area.ml.resid), pch=19, cex=1.25)
lines(area.ml, lwd=2)

mde.ml.resid = geo.algae$data - mde.pred
plot(variog(coords=geo.algae$coords,data=mde.ml.resid), pch=19, cex=1.25)
lines(mde.ml, lwd=2)

TM.ml.resid = geo.algae$data - TM.pred
plot(variog(coords=geo.algae$coords,data=TM.ml.resid), cex=1.25)
lines(TM.ml, lwd=2)
```

CHAPTER 4 CODE

```
# Script: Developing a new guild proportionality test
# Software: R2.3.0 (available at http://cran.au.r-project.org/)

# Author: Ailsa Kerswell
# Date: 9th Sept 2006


# This script is used to develop a new test for gp in
# aggregated communities.  Using hypothetical datasets
# generated to have either random, homogeneous or
# heterogeneous structure, I explore 5 alternative methods
# to test for guild proportionality in entire communities.
# 1. Schluter chi-squared test
# 2. Hybrid of Wilson randomisation and Schluter (and with mean-correction)
# 3. Stone et al method (with same and different range size for guilds)
# 4. Bernoulli trials

# To check for consistencies between Wilson and Schluter
# do randomisations with replace=T and replace=F and
# compare frequency distributions.

# Guild propotionality statistics are also calculated as per
# Wilson (1989) and variance comparisons plotted



library(MASS)
options(digits=2)

rm(list=ls())
nspp <- 20
nsites <- 6
ngroups <- 4
numsims <- 1000
grp.rich = c(3,4,6,7)

# Hypothetical occurrence matrices

## 1. Random structure matrix
sppmat <- matrix(NA,nrow=nsites,ncol=nspp)
sppmat[1,] <- c(0,1,0,0,0,0,1      ,0,0,0,0,0,1,0,0,0,1,0,0,1)
sppmat[2,] <- c(1,0,0,1,0,0,0      ,1,0,1,1,0,1,0,0,0,0,1,0,0)
sppmat[3,] <- c(0,1,0,0,0,1,0      ,1,0,0,0,0,1,0,0,0,0,1,1,1)
sppmat[4,] <- c(1,0,1,1,1,1,1      ,1,1,1,1,1,1,1,0,0,1,1,1,1)
sppmat[5,] <- c(1,1,1,1,1,0,0      ,1,1,1,1,1,1,1,1,1,1,1,0,1)
sppmat[6,] <- c(0,0,1,1,1,1,1      ,1,1,1,1,1,1,1,1,1,1,1,1,1)


# species' probabilities of occurrence
sppprobs <- colSums(sppmat)/sum(colSums(sppmat))

# cumulative number of species (used to facilitate making the group x site matrix)
richvec <- cumsum(c(3,4,6,7))

# Richness of sites
```

```
siterich <- rowSums(sppmat)
siterich

# Group x site matrix for the "real" hypothetical data, and chi-squared test
grpmat.real = matrix(NA,nrow=nsites,ncol=ngroups)
colnames(grpmat.real) = c('a', 'b', 'c', 'd')
grpmat.real[1,] = c(sum(sppmat[1,1:3]), sum(sppmat[1,4:7]), sum(sppmat[1,8:13]),
sum(sppmat[1,14:20]))
grpmat.real[2,] = c(sum(sppmat[2,1:3]), sum(sppmat[2,4:7]), sum(sppmat[2,8:13]),
sum(sppmat[2,14:20]))
grpmat.real[3,] = c(sum(sppmat[3,1:3]), sum(sppmat[3,4:7]), sum(sppmat[3,8:13]),
sum(sppmat[3,14:20]))
grpmat.real[4,] = c(sum(sppmat[4,1:3]), sum(sppmat[4,4:7]), sum(sppmat[4,8:13]),
sum(sppmat[4,14:20]))
grpmat.real[5,] = c(sum(sppmat[5,1:3]), sum(sppmat[5,4:7]), sum(sppmat[5,8:13]),
sum(sppmat[5,14:20]))
grpmat.real[6,] = c(sum(sppmat[6,1:3]), sum(sppmat[6,4:7]), sum(sppmat[6,8:13]),
sum(sppmat[6,14:20]))



# SPECIES MATCHING METHOD PLOT

chisq.test(grpmat.real)
real.chi=chisq.test(grpmat.real)$statistic
chi.expect = chisq.test(grpmat.real)$expected

plot(dchisq(1:30, 15), type='l', lwd=2)
points(real.chi, 0, cex=1.5)



# SAMLPING SPECIES WITH REPLACEMENT

chival <- rep(NA,numsims)
chiprob <- rep(NA,numsims)
chidf <- rep(NA,numsims)
siterich.sim <- matrix(NA,nrow=numsims,ncol=nsites)
grprich.sim <- matrix(NA,nrow=numsims,ncol=ngroups)
grpmat <- array(NA,dim=c(nsites,ngroups,numsims))
max.guild = matrix(NA, nrow=numsims, ncol=ngroups)
for (sim in 1:numsims) {
   for (site in 1:nsites) {
      # Sample species with replacement; prob of occurrence weighted by
      # observed frequency of occurrence
      samplevec <- sample(c(1:nspp),siterich[site],replace=T,prob=sppprobs)

      # Calculate number of species per functional group
      grpmat[site,1,sim] <- sum(samplevec<=richvec[1])
      grpmat[site,2,sim] <- sum(samplevec<=richvec[2])-grpmat[site,1,sim]
      grpmat[site,3,sim] <- sum(samplevec<=richvec[3])-grpmat[site,1,sim]-grpmat[site,2,sim]
      grpmat[site,4,sim] <- length(samplevec)-sum(grpmat[site,c(1:3),sim])

      max.guild[sim,] = apply(grpmat[,,sim], 2, max)
   }
   # Conduct chi-squared test
```

```
   chitest <- chisq.test(grpmat[,,sim])
   chival[sim] <- chitest$stat
   chiprob[sim] <- chitest$p.value
   chidf[sim] <- chitest$par
}

exceed.max = rep(NA, ngroups)
for(i in 1:ngroups){
   exceed.max[i] = sum(max.guild[,i]>grp.rich[i])
}
exceed.max/numsims

truehist(chival, col='grey', xlim=c(0, 30))
lines(dchisq(1:30, 15), lwd=2)
points(real.chi, 0, cex=1.5)


# SAMLPING SPECIES WITHOUT REPLACEMENT - MODIFIED SPECIES MATCHING

siterich.sim <- matrix(NA,nrow=numsims,ncol=nsites)
grprich.sim <- matrix(NA,nrow=numsims,ncol=ngroups)
grpmat <- array(NA,dim=c(nsites,ngroups,numsims))
max.guild = matrix(NA, nrow=numsims, ncol=ngroups)
for (sim in 1:numsims) {
   for (site in 1:nsites) {
      # Sample species with replacement; prob of occurrence weighted by
      # observed frequency of occurrence
      samplevec <- sample(c(1:nspp),siterich[site],replace=F,prob=sppprobs)

      # Calculate number of species per functional group
      grpmat[site,1,sim] <- sum(samplevec<=richvec[1])
      grpmat[site,2,sim] <- sum(samplevec<=richvec[2])-grpmat[site,1,sim]
      grpmat[site,3,sim] <- sum(samplevec<=richvec[3])-grpmat[site,1,sim]-grpmat[site,2,sim]
      grpmat[site,4,sim] <- length(samplevec)-sum(grpmat[site,c(1:3),sim])

      max.guild[sim,] = apply(grpmat[,,sim], 2, max)
   }
}

exceed.max = rep(NA, ngroups)
for(i in 1:ngroups){
   exceed.max[i] = sum(max.guild[,i]>grp.rich[i])
}
exceed.max/numsims

expt.freqs = apply(grpmat,c(1,2),mean)
   # mean grp frequencies at each site from randomisations

pseudochi.expt = rep(NA, numsims)
for(sim in 1:numsims){
   X2 = sum((grpmat[,,sim] - expt.freqs)^2/expt.freqs)
   pseudochi.expt[sim] = X2
}

pseudochi.obs = sum((grpmat.real - expt.freqs)^2/expt.freqs)


truehist(pseudochi.expt, col='grey', xlim=c(0, 25))
```

```
points(pseudochi.obs, 0, cex=1.5, lwd=2)
   # modified species matching plot



## WILSON GUILD-BY-GUILD METHOD

## Get guild prop variance for each group (across sites
## within the region in question)
## Do this for each randomisation and plot histogram
## of random community variances

rand.var = matrix(NA, nrow=numsims, ncol=ngroups)
colnames(rand.var) = c('A','B','C','D')

for(mat in 1:numsims){
   propmat = grpmat[,,mat]/siterich
      # converts number of gen in group to proportion
   rand.var[mat, ] = apply(propmat, 2, var)
}

real.prop = grpmat.real/siterich
real.var = apply(real.prop, 2, var)
   # real variance values

gp.index = matrix(NA, nrow=numsims, ncol=ngroups)
colnames(gp.index) = c('A','B','C','D')
   # guild prop index (Wilson 1989) Vobs/Vnull
   # if this is >1 then no guild prop if <1 guild prop

for(mat in 1:numsims){
   gp.index[mat,] = rand.var[mat,]/real.var
}

gp.sig = rep(NA, ngroups)
names(gp.sig) = c('A','B','C','D')
   # significance level for gp index.  Sig gp at alpha =0.05
   # if gp.sig<=0.025 (2-tailed test)

for(grp in 1:ngroups){
   TF = gp.index[,grp]<1
   gp.sig[grp] = sum(TF)/numsims
}


randvar.stats = matrix(NA, nrow=3, ncol=ngroups)
colnames(randvar.stats) = c('A','B','C','D')
rownames(randvar.stats) = c('mean', 'upper', 'lower')

for(i in 1:ngroups){
   sorted = sort(rand.var[,i])
   randvar.stats[1,i] = mean(sorted)
   randvar.stats[2,i] = sorted[975]
   randvar.stats[3,i] = sorted[25]
}

windows()
```

```
x = seq(1:4)
plot(x, randvar.stats[1,], ylim=c(-0.005,0.045),
     xlab='grps (alphabetical)', ylab='random comm variance +/- 95%CI')
arrows(x, randvar.stats[2,], x, randvar.stats[3,], length=0.1, code=3, angle = 90)
points(x, real.var, pch=19, cex=1.5)




## BERNOULLI APPROACH

spprich <- rep(1,nspp)

# nspp x numsims matrix with group richnesses (for calculating expected values)
spprichmat <- matrix(rep(spprich,nsites),nrow=nsites,byrow=T)


# frequencies (as proportion) of spp
sppfreq <- colSums(sppmat)/sum(sppmat)

# calculate the expected richness of spp in each site
fgfreqmat <- matrix(rep(sppfreq,nsites),nrow=nsites,byrow=T)
richmat <- matrix(rep(siterich,nspp),nrow=nsites,byrow=F)
expmat <- fgfreqmat*richmat

# and the probability of occurrence for each
# site & ssp
probmat <- expmat/spprichmat

# simulated incidence matrix
sppsim <- array(NA,dim=c(nsites,nspp,numsims))
for (sim in 1:numsims) {
   for (site in 1:nsites) {
      for (spp in 1:nspp) {
         # generate a random number for each species in group, determine
         # stochastically whether species present or not
         #if(probmat[site,spp]>1){
          #   x=1
          #   }else{
            x=probmat[site,spp]
         sppsim[site,spp,sim] <- sum(runif(spprich[spp])<x)
      }
   }
}



# summarise incidence matrix into group matrix

grpmat.binom = array(NA,dim=c(nsites,ngroups,numsims))
colnames(grpmat.binom) = c('a', 'b', 'c', 'd')

for(sim in 1:numsims){
   grpmat.binom[1,,sim] = c(sum(sppsim[1,1:3,sim]), sum(sppsim[1,4:7,sim]), sum(sppsim[1,8:13,sim]),
sum(sppsim[1,14:20,sim]))
   grpmat.binom[2,,sim] = c(sum(sppsim[2,1:3,sim]), sum(sppsim[2,4:7,sim]), sum(sppsim[2,8:13,sim]),
sum(sppsim[2,14:20,sim]))
```

```
   grpmat.binom[3,,sim] = c(sum(sppsim[3,1:3,sim]), sum(sppsim[3,4:7,sim]), sum(sppsim[3,8:13,sim]),
sum(sppsim[3,14:20,sim]))
   grpmat.binom[4,,sim] = c(sum(sppsim[4,1:3,sim]), sum(sppsim[4,4:7,sim]), sum(sppsim[4,8:13,sim]),
sum(sppsim[4,14:20,sim]))
   grpmat.binom[5,,sim] = c(sum(sppsim[5,1:3,sim]), sum(sppsim[5,4:7,sim]), sum(sppsim[5,8:13,sim]),
sum(sppsim[5,14:20,sim]))
   grpmat.binom[6,,sim] = c(sum(sppsim[6,1:3,sim]), sum(sppsim[6,4:7,sim]), sum(sppsim[6,8:13,sim]),
sum(sppsim[6,14:20,sim]))
}


# do chi-squared test on all randomised matrices

chi.binom = rep(NA, numsims)

for(sim in 1:numsims){
   chi.test = chisq.test(grpmat.binom[,,sim])
   chi.binom[sim] = chi.test$statistic
}


truehist(chi.binom, col='grey')
points(real.chi,0,cex=2)

lines(density(chi.binom, na.rm=T), col='red', lwd=2)



## make binomial expected matrix using the means
## of each cell over all randomisations

binom.expmat = matrix(NA, nrow=nsites, ncol=ngroups)

for(grp in 1:ngroups){
   for(site in 1:nsites){
      binom.expmat[site, grp] = mean(grpmat.binom[site, grp,])
   }
}


ttest.mat.bern = matrix(NA, nrow=nsites, ncol=ngroups)
colnames(ttest.mat.bern) = c('a', 'b', 'c', 'd')

for(site in 1:nsites){
   for(grp in 1:ngroups){
      data.string = grpmat.binom[site, grp,]
      ttest.mat.bern[site,grp] = t.test(data.string, mu=chi.expect[site, grp])$p.value
   }
}


## check row and column totals under bionomial approx

row.total = matrix(NA, nrow=numsims, ncol=nsites)
col.total =  matrix(NA, nrow=numsims, ncol=ngroups)

for(sim in 1:numsims){
   row.total[sim,] = rowSums(grpmat.binom[,,sim])
```

```
   col.total[sim,] = colSums(grpmat.binom[,,sim])
}

sort.row = apply(row.total,2,sort)
sort.col = apply(col.total,2,sort)

row.mean = colMeans(sort.row)
row.sd = apply(sort.row,2,sd)
row.upp = sort.row[975,]
row.low = sort.row[25,]

col.mean = colMeans(sort.col)
col.sd = apply(sort.col,2,sd)
col.upp = sort.col[975,]
col.low = sort.col[25,]




## STONE ET AL METHOD (RAND GROUPS SPP BELONG TO)

nrow=length(sppmat[,1])
ncol=length(sppmat[1,])
reps=1000

spp = c(rep('A',3), rep('B',4), rep('C',6), rep('D',7))
   # species occurring in the pool (in groups A-D)
spp_sample=as.data.frame(spp)

grp.list = c('A', 'B' , 'C', 'D')
n.grps = 4

spp.prob=colSums(sppmat)/sum(colSums(sppmat))
   # gives each spp same prob of being selected


Amat = matrix(NA, nrow=nrow, ncol=reps)
Bmat = matrix(NA, nrow=nrow, ncol=reps)
Cmat = matrix(NA, nrow=nrow, ncol=reps)
Dmat = matrix(NA, nrow=nrow, ncol=reps)
   # matrices to record how many of each group occur at each site
   # across reps number of randomisations

rand.mat.array = array(NA, dim=c(nsites, ngroups, reps))
colnames(rand.mat.array) = grp.list

max.guild = matrix(NA, nrow=reps, ncol=ngroups)

chi.vec.stone = rep(NA, reps)

for(r in 1:reps){

   rand.name=sample(spp_sample$factor, 20, replace=F, prob=spp.prob)
      # randomly chose group for each genus (i.e. colname for site.spp mat)
```

```
    summ.group=matrix(NA, nrow, n.grps)
    colnames(summ.group)=c('A', 'B', 'C', 'D')

    for(i in 1:nrow){
       TF = sppmat[i,]==1
          # records true if genus present
       replace.row = rand.name[TF]
          # replaces true with name of group that genus belongs to
       summ.group[i,] = summary(replace.row)
          # for that row, summarises how many genera in each group
    }

    chi.vec.stone[r] = chisq.test(summ.group)$statistic
       # chisq stat from test on randomised matrix

    Amat[,r] = summ.group[,1]
    Bmat[,r] = summ.group[,2]
    Cmat[,r] = summ.group[,3]
    Dmat[,r] = summ.group[,4]
       # number of genera in each group at each site for each randomisation

    rand.mat.array[,,r] = cbind(Amat[,r], Bmat[,r], Cmat[,r], Dmat[,r])

    max.guild[r,] = apply(rand.mat.array[,,r], 2, max)
}

exceed.max = rep(NA, ngroups)
for(i in 1:ngroups){
    exceed.max[i] = sum(max.guild[,i]>grp.rich[i])
}
exceed.max/reps

## Check if freq of occur for each group (i.e. mean col totals)
## maintained at observed values in random mats


expt = apply(rand.mat.array, c(1,2), mean)
    # mean freq of occur of each group at each site


sdev = apply(rand.mat.array, c(1,2), sd)
    # st dev of freq of occur of each group at each site

total.grp.occur = apply(rand.mat.array, 2, colSums)
    # total occurrence of each group in the random guild mat

mean.grp.occur = colMeans(total.grp.occur)
    # mean occurrence of each group in random guild mats
    # i.e. expected row totals

sdev.grp.occur = apply(total.grp.occur, 2, sd)
    # st dev of occurrence of each group in random guild mats
    # i.e. times 2 for ~95% CI on row totals

rbind(mean.grp.occur+(2*sdev.grp.occur), mean.grp.occur, mean.grp.occur-(2*sdev.grp.occur))

colSums(grpmat.real)
```

```
grpmat.real

expt

ttest.mat.stone = matrix(NA, nrow=nsites, ncol=ngroups)
colnames(ttest.mat.stone) = c('a', 'b', 'c', 'd')

for(site in 1:nsites){
   for(grp in 1:ngroups){
      data.string = rand.mat.array[site, grp,]
      ttest.mat.stone[site,grp] = t.test(data.string, mu=chi.expect[site, grp])$p.value
   }
}

pseudochi.expt.stone = rep(NA, numsims)
for(sim in 1:numsims){
   X2 = sum((rand.mat.array[,,sim] - expt)^2/expt)
   pseudochi.expt.stone[sim] = X2
}

pseudochi.obs.stone = sum((grpmat.real - expt)^2/expt)


truehist(pseudochi.expt.stone, col='grey', xlim=c(0, 25))
points(pseudochi.obs.stone, 0, cex=1.5, lwd=2)
   # modified species matching plot



plot(colSums(grpmat.real), mean.grp.occur, xlim=c(5,31), ylim=c(5,31), cex=1.25)
abline(0,1)
arrows(colSums(grpmat.real),mean.grp.occur+(2*sdev.grp.occur), colSums(grpmat.real),
        mean.grp.occur-(2*sdev.grp.occur), length=0)



# GETTING INITIAL MATRIX

sppmat <- matrix(NA,nrow=nsites,ncol=nspp)
sppmat[1,] <- c(rbinom(20, 1, p=0.3))
sppmat[2,] <- c(rbinom(20, 1, p=0.3))
sppmat[3,] <- c(rbinom(20, 1, p=0.3))
sppmat[4,] <- c(rbinom(20, 1, p=0.8))
sppmat[5,] <- c(rbinom(20, 1, p=0.8))
sppmat[6,] <- c(rbinom(20, 1, p=0.8))

colSums(sppmat)
```

CHAPTER 5 CODE

```
# Script: Applying new gp test to algae data (Chapter 5)
# Software: R2.3.0 (available at http://cran.au.r-project.org/)


# Author: Ailsa Kerswell
# Date: 15th September 2006


# This script applies the Bernoulli trial method to the algae
# data.  Starting at a regional level and scaling up to biome
# ocean and global.

# Results are plotted as frequency distributions of null model
# statistics with observed statistic marked as point.
# Statistics can be combined within same spatial scale due to
# additive nature of chi-squared.
# Significance determined by sorting null models values and
# comparing observed to the 95% CIs of null in whichever
# direction (<2.5th for homog and >97.5th for hetero).

# This code is for taxonomic clasifications but use exactly the
# same code for functional groups, just change the lables and
# the input data to fxn groups


rm(list=ls())

setwd("C:\\Ailsa's Documents\\Research\\R_data\\Community structure")
library(MASS)


fxngrps=c('cbr', 'crust', 'fbr', 'fila', 'gel', 'jcalc', 'leath', 'other.1', 'sheet', 'tuft')
fxngrps_sample=as.data.frame(fxngrps)
   # fxngrps in question

reps=1000
   # number of randomisations
ngroups=length(fxngrps)
   # number of fxngrps



# Observed group incidence matrices with chi-squared tests

robs<-read.table("Regions_all.txt",header=TRUE,sep="\t")
attach(robs)

robsF<-robs[,16:25]

r1F<-robsF[1:20,]
r2F<-robsF[21:50,]
r3F<-robsF[51:60,]
r4F<-robsF[61:74,]
r5F<-robsF[75:95,]
r6F<-robsF[96:120,]
```

```
  # indFividual matrices for each regions observed number of genera

trop.biomeF<-rbind(r1F, r2F, r5F)
temp.biomeF<-rbind(r3F, r4F, r6F)
ipac.oceanF<-rbind(r1F, r2F, r3F, r4F)
atl.oceanF<-rbind(r5F, r6F)
  # individual matrices for each biome/ocean observed number of genera


chi.indF=chisq.test(r1F)$statistic
chi.tpacF=chisq.test(r2F)$statistic
chi.npacF=chisq.test(r3F)$statistic
chi.spacF=chisq.test(r4F)$statistic
chi.tatlF=chisq.test(r5F)$statistic
chi.natlF=chisq.test(r6F)$statistic

chi.tropF=chisq.test(trop.biomeF)$statistic
chi.tempF=chisq.test(temp.biomeF)$statistic

chi.ipacF=chisq.test(ipac.oceanF)$statistic
chi.atlF=chisq.test(atl.oceanF)$statistic

chi.allF=chisq.test(robsF)$statistic



# Observed species - incidence matrices
incid.mat<-read.table("incidence_allsites.txt",header=TRUE,sep="\t")
attach(incid.mat)

# Break up data by region etc

indF.cols=incid.mat[,1:23]
tempF.indF=rowSums(indF.cols[,4:23])>0
indian=indF.cols[tempF.indF,]

tpacF.cols=cbind(incid.mat[,1:3], incid.mat[,24:53])
tempF.tpacF=rowSums(tpacF.cols[,4:33])>0
tpacF=tpacF.cols[tempF.tpacF,]

npacF.cols=cbind(incid.mat[,1:3], incid.mat[,54:63])
tempF.npacF=rowSums(npacF.cols[,4:13])>0
npacF=npacF.cols[tempF.npacF,]

spacF.cols=cbind(incid.mat[,1:3], incid.mat[,64:77])
tempF.spacF=rowSums(spacF.cols[,4:17])>0
spacF=spacF.cols[tempF.spacF,]

tatlF.cols=cbind(incid.mat[,1:3], incid.mat[,78:98])
tempF.tatlF=rowSums(tatlF.cols[,4:24])>0
tatlF=tatlF.cols[tempF.tatlF,]

natlF.cols=cbind(incid.mat[,1:3], incid.mat[,99:123])
tempF.natlF=rowSums(natlF.cols[,4:28])>0
natlF=natlF.cols[tempF.natlF,]
```

```
tropF.cols=cbind(incid.mat[,1:3], incid.mat[,4:53], incid.mat[,78:98])
tempF.tropF=rowSums(tropF.cols[,4:74])>0
tropF=tropF.cols[tempF.tropF,]

tempF.cols=cbind(incid.mat[,1:3], incid.mat[,54:77], incid.mat[,98:123])
tempF.tempF=rowSums(tempF.cols[,4:52])>0
tempF=tempF.cols[tempF.tempF,]



ipacF.cols=cbind(incid.mat[,1:3], incid.mat[,4:77])
tempF.ipacF=rowSums(ipacF.cols[,4:77])>0
ipacF=ipacF.cols[tempF.ipacF,]

atlF.cols=cbind(incid.mat[,1:3], incid.mat[,78:123])
tempF.atlF=rowSums(atlF.cols[,4:49])>0
atlF=atlF.cols[tempF.atlF,]


allF = incid.mat




# Function to make randomised spp-incidence matrices
# using binomial method

rand.spp.mat = function(ncols, reg.incid, numsims){
   # ncols:number of columns in indFcidence mat for that region
   # reg.incid: incidence matrix (to delete first 3 colunms)
   # numsims: number of simulations to run

   minus.cols=reg.incid[,4:ncols]
   sppmat=t(minus.cols)
      # observed spp incidence matrix

   nsites=length(sppmat[,1])
      # number of sites in incidence matrix
   siterich=rowSums(sppmat)
      # spp richness at each site

   nspp=length(sppmat[1,])
   spprich=rep(1,nspp)
   sppfreq=colSums(sppmat)/sum(sppmat)
      # matrix of expected frequency of spp for each cell
      # of observed incidence matrix
   spprichmat=matrix(rep(spprich,nsites),nrow=nsites,byrow=T)
      # nspp x numsims matrix with spp richnesses
      # (for calculating expected values)

   freqmat=matrix(rep(sppfreq,nsites),nrow=nsites,byrow=T)
   richmat=matrix(rep(siterich,nspp),nrow=nsites,byrow=F)
   expmat=freqmat*richmat
      # calculates the expected richness of spp in each site

   #probmat=expmat/spprichmat
```

```
      # the probability of occurrence for each spp @ each site

   # simulated incidence matrix
   sppsim =array(NA,dim=c(nsites,nspp,numsims))
   for (sim in 1:numsims) {
      for (site in 1:nsites) {
         for (spp in 1:nspp) {
            # generate a random number for each species in group, determine
            # stochasticallFy whether species present or not
            x=expmat[site,spp]
            sppsim[site,spp,sim] = sum(runif(spprich[spp])<x)
         }
      }
   }
   return(sppsim)
}



# Make random spp-incidence matrices for allF regions
randspp.indF=rand.spp.mat(23, indian, reps)
randspp.tpacF=rand.spp.mat(33, tpacF, reps)
randspp.npacF=rand.spp.mat(13, npacF, reps)
randspp.spacF=rand.spp.mat(17, spacF, reps)
randspp.tatlF=rand.spp.mat(24, tatlF, reps)
randspp.natlF=rand.spp.mat(28, natlF, reps)



## Function to make spp-incidence matrix into grp-incidence matrix

make.grp.matF = function(real.incidence, randspp.array, nsites, ngroups, numsims){

   table.mat=table(real.incidence[,2])
   table.mat.cum=cumsum(table.mat)

   randgrp.mat = array(NA,dim=c(nsites,ngroups,numsims))
   colnames(randgrp.mat) = fxngrps

   for(site in 1:nsites){
      for(sim in 1:numsims){
         randgrp.mat[site,,sim] = c(sum(randspp.array[site, 1:table.mat.cum[1], sim]),
                                    sum(randspp.array[site, (table.mat.cum[1]+1):table.mat.cum[2], sim]),
                                    sum(randspp.array[site, (table.mat.cum[2]+1):table.mat.cum[3], sim]),
                                    sum(randspp.array[site, (table.mat.cum[3]+1):table.mat.cum[4], sim]),
                                    sum(randspp.array[site, (table.mat.cum[4]+1):table.mat.cum[5], sim]),
                                    sum(randspp.array[site, (table.mat.cum[5]+1):table.mat.cum[6], sim]),
                                    sum(randspp.array[site, (table.mat.cum[6]+1):table.mat.cum[7], sim]),
                                    sum(randspp.array[site, (table.mat.cum[7]+1):table.mat.cum[8], sim]),
                                    sum(randspp.array[site, (table.mat.cum[8]+1):table.mat.cum[9], sim]),
                                    sum(randspp.array[site, (table.mat.cum[9]+1):table.mat.cum[10], sim]))
      }
   }
   return(randgrp.mat)
}
```

```
randgrp.indF=make.grp.matF(indian, randspp.indF, 20, 10, reps)
randgrp.tpacF=make.grp.matF(tpacF, randspp.tpacF, 30, 10, reps)
randgrp.npacF=make.grp.matF(npacF, randspp.npacF, 10, 10, reps)
randgrp.spacF=make.grp.matF(spacF, randspp.spacF, 14, 10, reps)
randgrp.tatlF=make.grp.matF(tatlF, randspp.tatlF, 21, 10, reps)
randgrp.natlF=make.grp.matF(natlF, randspp.natlF, 25, 10, reps)



# Function to do chi-squared test on allF randomised matrices
# different ones for tropFical and tempFerate (b/c laminariales)

get.chirand.trop = function(randgrp.array, numsims){
   chirand = rep(NA, numsims)

   for(sim in 1:numsims){
      chi.test = chisq.test(randgrp.array[,-8,sim])
      chirand[sim] = chi.test$statistic
   }
   return(chirand)
}


get.chirand.temp = function(randgrp.array, numsims){
   chirand = rep(NA, numsims)

   for(sim in 1:numsims){
      chi.test = chisq.test(randgrp.array[,,sim])
      chirand[sim] = chi.test$statistic
   }
   return(chirand)
}


chirand.indF=get.chirand.temp(randgrp.indF,reps)
chirand.tpacF=get.chirand.temp(randgrp.tpacF,reps)
chirand.npacF=get.chirand.temp(randgrp.npacF,reps)
chirand.spacF=get.chirand.temp(randgrp.spacF,reps)
chirand.tatlF=get.chirand.temp(randgrp.tatlF,reps)
chirand.natlF=get.chirand.temp(randgrp.natlF,reps)


par(mfrow=c(2,3))
par(pty='s')


truehist(chirand.indF, col='grey', xlim=c(min(chirand.indF), max(chirand.indF)))
points(chi.indF, 0, cex=1.5)
title("indian")

truehist(chirand.tpacF, col='grey', xlim=c(min(chirand.tpacF), max(chirand.tpacF)))
points(chi.tpacF, 0, cex=1.5)
title("tropF pac")

truehist(chirand.tatlF, col='grey', xlim=c(min(chirand.tatlF), max(chirand.tatlF)))
points(chi.tatlF, 0, cex=1.5)
title("tropF atlF")
```

```
truehist(chirand.npacF, col='grey', xlim=c(min(chirand.npacF), max(chi.npacF)))
points(chi.npacF, 0, cex=1.5)
title("north pac")

truehist(chirand.spacF, col='grey', xlim=c(min(chirand.spacF), max(chi.spacF)))
points(chi.spacF, 0, cex=1.5)
title("south pac")

truehist(chirand.natlF, col='grey', xlim=c(min(chirand.natlF), max(chi.natlF)))
points(chi.natlF, 0, cex=1.5)
title("north atlF")



# Combine the obs stats from each region into one combined regional statistic
# Do the same for the null distributions (sort chirand vectors first, then add)

region.matF = rbind(sort(chirand.indF), sort(chirand.tpacF), sort(chirand.tatlF),
                sort(chirand.npacF), sort(chirand.spacF), sort(chirand.natlF))

chirand.regionF = colSums(region.matF)
chi.regionF = sum(chi.indF, chi.tpacF, chi.tatlF, chi.npacF, chi.spacF, chi.natlF)

truehist(chirand.regionF, xlim=c(min(chirand.regionF)-20, chi.regionF+20), col='grey')
points(chi.regionF, 0, cex=1.5)


# plot tropF regions combined and temperate regions combined

chirand.region.tropF = colSums(region.matF[1:3,])
chi.region.tropF = sum(chi.indF, chi.tpacF, chi.tatlF)

par(mfrow=c(1,2))
par(pty='s')
truehist(chirand.region.tropF, xlim=c(min(chirand.region.tropF)-20, max(chirand.region.tropF)+20),
col='grey')
points(chi.region.tropF, 0, cex=1.5)

chirand.region.tempF = colSums(region.matF[4:6,])
chi.region.tempF = sum(chi.npacF, chi.spacF, chi.natlF)

truehist(chirand.region.tempF, xlim=c(min(chirand.region.tempF)-20, chi.region.tempF+20), col='grey')
points(chi.region.tempF, 0, cex=1.5)



## Do the same procedure for biomes/oceans/globe

randspp.tropF=rand.spp.mat(74, tropF, reps)
randgrp.tropF=make.grp.matF(tropF, randspp.tropF, 71, 10, reps)
chirand.tropF=get.chirand.temp(randgrp.tropF,reps)


randspp.tempF=rand.spp.mat(52, tempF, reps)
randgrp.tempF=make.grp.matF(tempF, randspp.tempF, 49, 10, reps)
chirand.tempF=get.chirand.temp(randgrp.tempF,reps)
```

```
randspp.ipacF=rand.spp.mat(77, ipacF, reps)
randgrp.ipacF=make.grp.matF(ipacF, randspp.ipacF, 74, 10, reps)
chirand.ipacF=get.chirand.temp(randgrp.ipacF,reps)


randspp.atlF=rand.spp.mat(49, atlF, reps)
randgrp.atlF=make.grp.matF(atlF, randspp.atlF, 46, 10, reps)
chirand.atlF=get.chirand.temp(randgrp.atlF,reps)


randspp.allF=rand.spp.mat(123, allF, reps)
randgrp.allF=make.grp.matF(allF, randspp.allF, 120, 10, reps)
chirand.allF=get.chirand.temp(randgrp.allF,reps)



par(mfrow=c(2,3))
par(pty='s')

truehist(chirand.tropF, col='grey', xlim=c(chi.tropF-20, max(chirand.tropF)))
points(chi.tropF, 0, cex=1.5)
title("tropF")

truehist(chirand.tempF, col='grey', xlim=c(min(chirand.tempF), chi.tempF+20))
points(chi.tempF, 0, cex=1.5)
title("tempF")

truehist(chirand.ipacF, col='grey', xlim=c(min(chirand.ipacF), chi.ipacF+20))
points(chi.ipacF, 0, cex=1.5)
title("ipacF")

truehist(chirand.atlF, col='grey', xlim=c(min(chirand.atlF), chi.atlF+20))
points(chi.atlF, 0, cex=1.5)
title("atlF")

truehist(chirand.allF, col='grey', xlim=c(min(chirand.allF), chi.allF+20))
points(chi.allF, 0, cex=1.5)
title("allF")



# Combine the obs stats from each biome/ocean into one combined statistic
# Do the same for the null distributions (sort chirand vectors first, then add)

par(mfrow=c(1,2))
par(pty='s')

biome.matF = rbind(sort(chirand.tropF), sort(chirand.tempF))
chirand.biomeF = colSums(biome.matF)
chi.biomeF = sum(chi.tropF, chi.tempF)
truehist(chirand.biomeF, xlim=c(min(chirand.biomeF)-20, chi.biomeF+20), col='grey')
points(chi.biomeF, 0, cex=1.5)


ocean.matF = rbind(sort(chirand.ipacF), sort(chirand.atlF))
chirand.oceanF = colSums(ocean.matF)
chi.oceanF = sum(chi.ipacF, chi.atlF)
```

```
truehist(chirand.oceanF, xlim=c(min(chirand.oceanF)-20, chi.oceanF+20), col='grey')
points(chi.oceanF, 0, cex=1.5)



# Plot each full model (region,biome,ocean,global)

par(mfrow=c(2,2))
par(pty='s')

truehist(chirand.regionF, xlim=c(min(chirand.regionF)-20, chi.regionF+20), col='grey')
points(chi.regionF, 0, cex=1.5)
title('regions model F')

truehist(chirand.biomeF, xlim=c(min(chirand.biomeF)-20, chi.biomeF+20), col='grey')
points(chi.biomeF, 0, cex=1.5)
title('biomes model F')

truehist(chirand.oceanF, xlim=c(min(chirand.oceanF)-20, chi.oceanF+20), col='grey')
points(chi.oceanF, 0, cex=1.5)
title('oceans model F')

truehist(chirand.allF, col='grey', xlim=c(min(chirand.allF), chi.allF+20))
points(chi.allF, 0, cex=1.5)
title("global model F")



save(randspp.indF, randspp.tpacF, randspp.npacF, randspp.spacF, randspp.tatlF, randspp.natlF,
file="binomial_randomisation_functional_region_spp_arrays.Rdata")


save(randgrp.indF, randgrp.tpacF, randgrp.npacF, randgrp.spacF, randgrp.tatlF, randgrp.natlF,
file="binomial_randomisation_functional_region_grp_arrays.Rdata")


save(randspp.tropF, randspp.tempF, randspp.ipacF, randspp.atlF, randspp.allF,
file="binomial_randomisation_functional_over_region_spp_arrays.Rdata")

save(randgrp.tropF, randgrp.tempF, randgrp.ipacF, randgrp.atlF, randgrp.allF,
file="binomial_randomisation_functional_over_region_grp_arrays.Rdata")


save(chirand.indF, chirand.tpacF, chirand.npacF, chirand.spacF,chirand.tatlF,chirand.natlF,
      chi.indF, chi.tpacF, chi.npacF, chi.spacF,chi.tatlF,chi.natlF,
      chirand.tropF, chirand.tempF, chirand.ipacF, chirand.atlF, chirand.allF,
      chi.tropF, chi.tempF, chi.ipacF, chi.atlF, chi.allF,
file="binomial_randomisation_functional_chis.Rdata")



## LIKELIHOOD METHODS TO DISTINGUISH BETWEEN BEST
## MODEL (TROP BIOME OR REGION)


## To get Expected value for each group at each site for the
## tropical biomes model
```

```
ngroups=10
nsites=71
siterich=colSums(tropF[,4:74])

Emat.tropB = matrix(NA, nrow=nsites, ncol=ngroups)
colnames(Emat.tropB)=fxngrps

for(site in 1:nsites){
   for(grp in 1:ngroups){
      prop.mat = randgrp.tropF/siterich
      Emat.tropB[site,grp] = mean(prop.mat[site,grp,])
   }
}



## For indian, tpacF, tatlF - then combine these into
## Expected value for each group at each site for the
## tropical regions model

nsites=20
siterich=colSums(indian[,4:23])

Emat.ind = matrix(NA, nrow=nsites, ncol=ngroups)
colnames(Emat.ind)=fxngrps

for(site in 1:nsites){
   for(grp in 1:ngroups){
      prop.mat = randgrp.indF/siterich
      Emat.ind[site,grp] = mean(prop.mat[site,grp,])
   }
}


nsites=30
siterich=colSums(tpacF[,4:33])

Emat.tpac = matrix(NA, nrow=nsites, ncol=ngroups)
colnames(Emat.tpac)=fxngrps

for(site in 1:nsites){
   for(grp in 1:ngroups){
      prop.mat = randgrp.tpacF/siterich
      Emat.tpac[site,grp] = mean(prop.mat[site,grp,])
   }
}


nsites=21
siterich=colSums(tatlF[,4:24])

Emat.tatl = matrix(NA, nrow=nsites, ncol=ngroups)
colnames(Emat.tatl)=fxngrps

for(site in 1:nsites){
   for(grp in 1:ngroups){
      prop.mat = randgrp.tatlF/siterich
```

```
        Emat.tatl[site,grp] = mean(prop.mat[site,grp,])
    }
}


Emat.tropR = rbind(Emat.ind, Emat.tpac, Emat.tatl)


## Calculate log likes for each model
## See paper for formula explanation

LL.region = sum(trop.biomeF*log(Emat.tropR))
LL.biome = sum(trop.biomeF*log(Emat.tropB))


aicR = -2*LL.region + 2*3*9
aicB = -2*LL.biome + 2*9
    # AIC for each model


## TRADITIONAL GP TESTS
## Get guild prop variance for each group (across sites
## within the region in question)
## Do this for each randomisation and plot histogram
## of random community variances


wilson.gp = function(numsims, grpmat, grpmat.real){

    siterich = rowSums(grpmat.real)

    rand.var = matrix(NA, nrow=numsims, ncol=ngroups)
    colnames(rand.var) = fxngrps

    for(mat in 1:numsims){
        propmat = grpmat[,,mat]/siterich
            # converts number of gen in group to proportion
        rand.var[mat, ] = apply(propmat, 2, var)
    }

    real.prop = grpmat.real/siterich
    real.var = apply(real.prop, 2, var)
        # real variance values

    gp.index = matrix(NA, nrow=numsims, ncol=ngroups)
    colnames(gp.index) = fxngrps
        # guild prop index (Wilson 1989) Vobs/Vnull
        # if this is >1 then no guild prop if <1 guild prop

    for(mat in 1:numsims){
        gp.index[mat,] = rand.var[mat,]/real.var
    }

    gp.sig = rep(NA, ngroups)
    names(gp.sig) = fxngrps
        # significance level for gp index.  Sig gp at alpha =0.05
        # if gp.sig<=0.025 (2-tailed test)
```

```
    for(grp in 1:ngroups){
        TF = gp.index[,grp]<1
        gp.sig[grp] = sum(TF)/numsims
    }

    randvar.stats = matrix(NA, nrow=3, ncol=ngroups)
    colnames(randvar.stats) = fxngrps
    rownames(randvar.stats) = c('mean', 'upper', 'lower')

    for(i in 1:ngroups){
        sorted = sort(rand.var[,i])
        randvar.stats[1,i] = mean(sorted)
        randvar.stats[2,i] = sorted[9]
        randvar.stats[3,i] = sorted[2]
    }

    return(list(var.rand=rand.var, sig=gp.sig, plot.stats=randvar.stats, var.obs=real.var))
}

trop.wilson.gpF=wilson.gp(reps, randgrp.tropF, trop.biomeF)
temp.wilson.gpF=wilson.gp(reps, randgrp.tempF, temp.biomeF)


save(trop.wilson.gpF, temp.wilson.gpF,
file="binomial_randomisation_functional_biome_wilson.Rdata")


par(mfrow=c(1,2))
par(pty='s')

x=seq(1,10)

trop.o = order(trop.wilson.gpF$plot.stats[1,], decreasing=F)
new.trop = trop.wilson.gpF$plot.stats[,trop.o]

plot(x, new.trop[1,], ylim=c(0,max(new.trop)))
arrows(x, new.trop[2,], x, new.trop[3,], length=0.1, code=3, angle=90)
points(trop.wilson.gpF$var.obs[trop.o], pch=19)


temp.o = order(temp.wilson.gpF$plot.stats[1,], decreasing=F)
new.temp = temp.wilson.gpF$plot.stats[,temp.o]

plot(x, new.temp[1,], ylim=c(0,max(new.temp)))
arrows(x, new.temp[2,], x, new.temp[3,], length=0.1, code=3, angle=90)
points(temp.wilson.gpF$var.obs[temp.o], pch=19)
```