

# Hierarchical drivers of reef-fish metacommunity structure

M. AARON MACNEIL,<sup>1,6</sup> NICHOLAS A. J. GRAHAM,<sup>1,7</sup> NICHOLAS V. C. POLUNIN,<sup>1</sup> MICHEL KULBICKI,<sup>2</sup> RENÉ GALZIN,<sup>3</sup>  
MIREILLE HARMELIN-VIVIEN,<sup>4</sup> AND STEVEN P. RUSHTON<sup>5</sup>

<sup>1</sup>*School of Marine Science and Technology, University of Newcastle, Newcastle upon Tyne NE1 7RU United Kingdom*

<sup>2</sup>*Institut de Recherche Pour le Développement, Université de Perpignan, 66860 Perpignan Cedex, France*

<sup>3</sup>*UMR 5244 CNRS-EPHE-UPVD, Université de Perpignan, 66860 Perpignan Cedex, France*

<sup>4</sup>*Centre d'Océanologie de Marseille-UMR CNRS 6540, Station Marine d'Endoume, 13007 Marseille, France*

<sup>5</sup>*School of Biology and Psychology, University of Newcastle, Newcastle upon Tyne NE1 7RU United Kingdom*

**Abstract.** Coral reefs are highly complex ecological systems, where multiple processes interact across scales in space and time to create assemblages of exceptionally high biodiversity. Despite the increasing frequency of hierarchically structured sampling programs used in coral-reef science, little progress has been made in quantifying the relative importance of processes operating across multiple scales. The vast majority of reef studies are conducted, or at least analyzed, at a single spatial scale, ignoring the implicitly hierarchical structure of the overall system in favor of small-scale experiments or large-scale observations. Here we demonstrate how  $\alpha$  (mean local number of species),  $\beta$  diversity (degree of species dissimilarity among local sites), and  $\gamma$  diversity (overall species richness) vary with spatial scale, and using a hierarchical, information-theoretic approach, we evaluate the relative importance of site-, reef-, and atoll-level processes driving the fish metacommunity structure among 10 atolls in French Polynesia. Process-based models, representing well-established hypotheses about drivers of reef-fish community structure, were assembled into a candidate set of 12 hierarchical linear models. Variation in fish abundance, biomass, and species richness were unevenly distributed among transect, reef, and atoll levels, establishing the relative contribution of variation at these spatial scales to the structure of the metacommunity. Reef-fish biomass, species richness, and the abundance of most functional-groups corresponded primarily with transect-level habitat diversity and atoll-lagoon size, whereas detritivore and grazer abundances were largely correlated with potential covariates of larval dispersal. Our findings show that (1) within-transect and among-atoll factors primarily drive the relationship between  $\alpha$  and  $\gamma$  diversity in this reef-fish metacommunity; (2) habitat is the primary correlate with reef-fish metacommunity structure at multiple spatial scales; and (3) inter-atoll connectedness was poorly correlated with the nonrandom clustering of reef-fish species. These results demonstrate the importance of modeling hierarchical data and processes in understanding reef-fish metacommunity structure.

**Key words:**  $\alpha$ ,  $\beta$ , and  $\gamma$  diversity; biodiversity; coral reefs; ecological scales; habitat structure; mixed-effects models; multiple working hypotheses; reef-fish biomass; species abundance; species richness.

## INTRODUCTION

Understanding the significance of ecological processes operating at different spatial scales is a persistent theme in ecology, as patterns observed at small scales are frequently constrained by temporal processes occurring over multiple, larger scales (Allen and Hoekstra 1992). Such hierarchically structured processes generate complex community patterns that are best described explicitly by hierarchical analyses (Allen and Starr 1982). Coral-reef ecosystems are considered to be highly hierarchical in structure, influenced

by processes operating at multiple scales in space and time (Hatcher 1997). However, although many reef-based studies are purported to be hierarchical, frequently the data, rather than the analysis, have been developed at multiple scales (Karlson and Cornell 1999). Ignoring the hierarchical structure of ecosystems and data can result in biased parameter estimates, artificially small standard errors, and a false sense of precision in a poorly fitted analysis (Raudenbush and Byrk 2002).

Confusion between sampling and analysis at multiple scales can also generate debate about which processes dominate control of coral-reef systems (Allen and Hoekstra 1992; Hatcher 1997). Selecting a scale appropriate to the question at hand is critical as it often dictates the method of inference, such as conducting experiments at small scales or making large scale observations (Steele and Forrester 2005). This is particularly true on coral reefs, where the applicability of in situ experimentation has focused on site-level

Manuscript received 20 March 2007; revised 8 May 2008; accepted 14 May 2008. Corresponding Editor: M. H. Carr.

<sup>6</sup> Present address: Australian Institute of Marine Sciences, PMB 3 Townsville MC, Townsville, Queensland 4810 Australia. E-mail: macneil@glau.ca

<sup>7</sup> Present address: ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Queensland, Australia.

processes (e.g., Syms and Jones 2000) whereas the exceptionally high diversity of reef fish species globally has led to broad-scale, regional analyses to understand how such wide-scale diversity developed (e.g., Bellwood and Hughes 2001, Dornelas et al. 2006). The spatial scale at which community patterns are studied will generate support for processes operating at that scale (Allen and Starr 1982, Gust et al. 2001). However, if we are to gain a comprehensive understanding of how reef systems are structured, it is necessary for researchers to quantify the relative effects of processes operating at multiple levels. Among the most relevant concepts for examining spatial linkages at multiple scales is that of the metacommunity, whereby individual communities are linked to other communities by the dispersal of multiple interacting species (Liebold et al. 2004).

Although the metacommunity idea has not received much attention in the coral-reef literature, understanding the nature of reef-fish community structure has been a research focus of reef ecologists for many years (e.g., Sale and Dybdahl 1975), with considerable debate regarding the degree to which regional dispersal dynamics and reef-scale habitat characteristics determine community structure. Larval-dispersal theory suggests that recruitment is the primary process by which reef-fish communities are structured, with mortality and dispersal events in the plankton determining which species are subsequently available for recruitment to a given reef (Sale 1980, Doherty 1981, Sale et al. 1994); low recruitment has been shown to limit the abundance of individual species (Doherty and Fowler 1994, Caley et al. 1996, Doherty 2002).

Within-reef habitat characteristics can influence the abundance and diversity of reef-fish communities (Chabanet et al. 1997, Syms and Jones 2004) where they determine the species-specific responses of community members to their environment (Almany 2004). Complex habitats can mitigate predation by providing shelter (Hixon and Beets 1993), reduce competition through increased niche availability (Jones and Syms 1998, Munday and Jones 1998), and provide specialized settlement habitat for planktonic larvae (Jones et al. 2004). At small scales, complexity, coral species composition, and substrate types have all been shown to be important correlates of community structure, reflecting both species-specific habitat preferences and the interactions among individual fishes (Ault and Johnson 1998, Syms and Jones 2000).

To evaluate the importance of these processes, many studies rely on observational data gathered at a single point in time. Unfortunately, such "snapshot" data have low inferential strength and are all too frequently mined for patterns with which to infer the ecological processes present at the time the observations were made; as a result, these kinds of analyses often lead to poorly supported inferences (Burnham and Anderson 2002). A somewhat stronger approach is to make a priori hypotheses about the patterns we expect to see if specific

processes were operating, encode those ideas in a set of candidate models, and then evaluate their relative support given the data (Hilborn and Mangel 1997).

In this study we investigate the hierarchical structure of coral-reef fish data using hierarchical randomization tests at transect, reef, and atoll scales of (1) reef fish  $\alpha$  (mean local number of species)  $\beta$  (degree of species dissimilarity among local sites), and  $\gamma$  (overall species richness) diversity, and (2) community similarity in terms of abundance, diversity, and biomass. These analyses are developed within a well-defined metacommunity in the Tuamotu Archipelago of French Polynesia. We analyze observational data using hierarchical linear models to compare evidence for three hypothesized mechanisms driving reef-fish metacommunity structure: spatial separation (Bellwood and Hughes 2001), habitat specialization (Gratwicke and Speight 2005), and larval dispersal (Doherty and Fowler 1994). We also discuss the strengths and weaknesses of using patterns in spatial data to make inferences regarding metacommunity structure and dynamic processes.

## METHODS

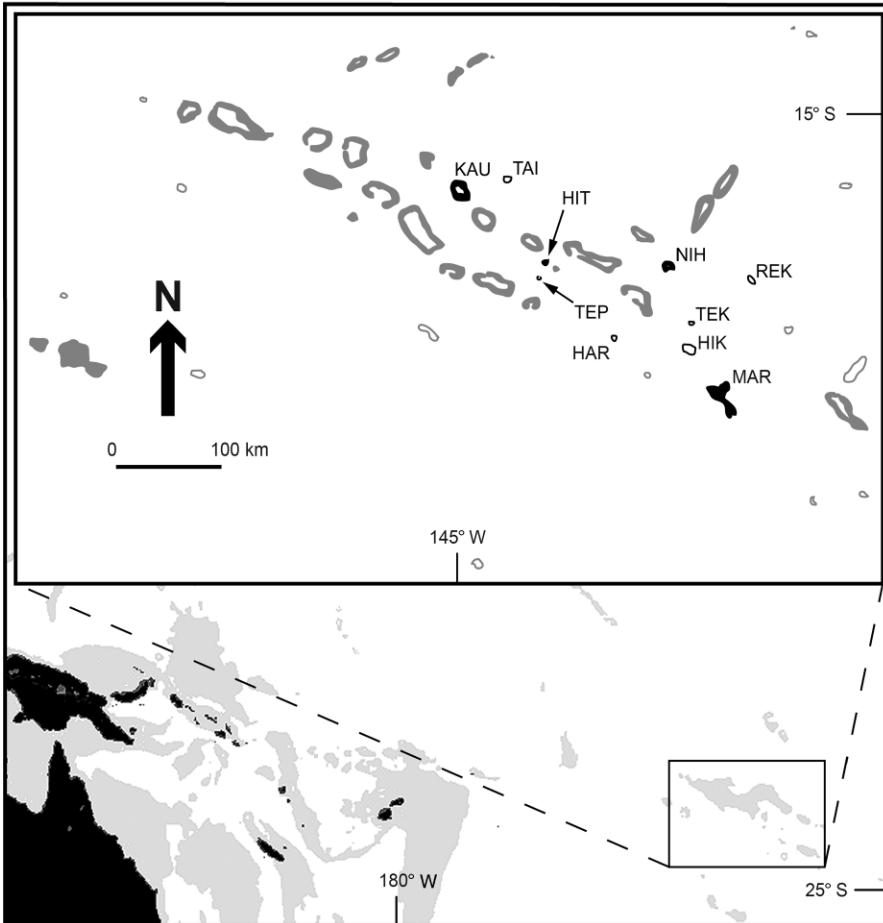
### *Study sites*

Our analysis was conducted using data from the Typatoll programme (Dufour and Harmelin-Vivien 1997) that investigated geomorphological, chemical, and biological characteristics of 10 atolls in the Tuamotu Archipelago, an isolated island chain in French Polynesia. The Archipelago is relatively undisturbed ( $<0.05$  persons/reef km<sup>2</sup>), remote (the nearest islands [Tahiti] are  $\sim 300$  km away), and currents in the region are negligible ( $<0.15$  m/s; NOAA OSCAR database, *available online*).<sup>8</sup> Sampling was conducted from November to December of 1995 and 1996, using underwater visual census (UVC) of six reef types (Appendix B: Table B3) among 10 haphazardly selected atolls (Mellin et al. 2006). UVC sampling was conducted using  $50 \times 5$  m belt transects and their number and locations were set to be proportional to total reef area (Appendix B: Table B4); 509 transects were surveyed in total.

For each transect, one diver recorded all fish detectable within 5 m of the transect line, excluding species  $<3$  cm fork length. Each record included the density (number of fish/250 m<sup>2</sup>) and average size of each species encountered. Fish length estimates by divers are usually within 10% of the true value (see Kulbicki et al. 2005). Fish masses were estimated using length-mass relationships (Kulbicki et al. 2005); biomass was calculated as kilograms per square meter per transect (kg·m<sup>-2</sup>·transect<sup>-1</sup>). The percent cover of 16 habitat variables (Appendix B: Table B2) was also collected along the transect, using the line-intercept technique (English et al. 1994). These values were later used to

<sup>8</sup> (<http://www.oscar.noaa.gov>)

## a) Study locations



## b) Study scale levels

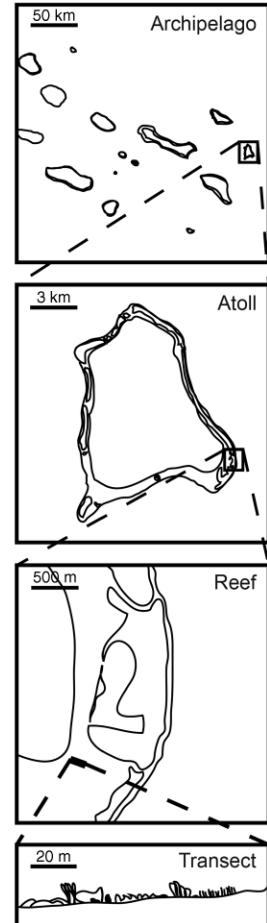


FIG. 1. Study sites and scale levels. (a) Study sampling locations, Tuamotu Archipelago, French Polynesia. Atolls sampled included Haraiki (HAR), Hikueru (HIK), Hiti (HIT), Kauehi (KAU), Marokau (MAR), Nihiru (NIH), RekaReka (REK), Taiaroa (TAI), Tekokota (TEK), and Tepoto (TEP). (b) Spatial scales of analysis.

calculate habitat diversity (see Appendix A). Atolls could be visited only once and individual transects were not sampled repeatedly.

#### *Hierarchical randomization*

Because we assumed the archipelago defined the boundaries of the reef-fish metacommunity, modelling was conducted at ecological units within the archipelago level, according to transect, reef, and atoll groups (Fig. 1). We were specifically interested in the hierarchical structure of the data and how well transect-, reef-, and atoll-level characteristics were correlated with fish community structure. We assessed these relationships using hierarchical randomization tests and hierarchical linear models, which decompose the variation in the data to each scale and improve the accuracy of model parameter estimates (relative to nonhierarchical models).

A critical assumption of our analysis was that atolls, reef habitats within atolls, and transects within reef habitats within atolls defined ecologically meaningful

groups that were hierarchically structured within the data. We began by investigating how community similarity responded to these groupings in hierarchical randomization tests. Such techniques have previously successfully quantified the assumed effects of habitat specialization and spatial separation in determining  $\alpha$  diversity (mean local number of species),  $\beta$  diversity (degree of species dissimilarity among local sites), and  $\gamma$  diversity (overall species richness) of coral species at transect, site (reef), island, and island-group scales (Cornell et al. 2007). Implicit within this, randomized community similarity can be either overdispersed (equally spaced) or underdispersed (clustered). It should be noted, however, that reef communities are nearly always clustered (Karlson et al. 2007) due to processes such as habitat specialization or spatial separation (Freestone and Inouye 2006). Although our data did not include an island-group scale, our methods closely followed those of Cornell et al (2007) and readers should consult their paper for specifics of our approach.

Following Cornell et al. (2007), we calculated mean  $\alpha$  and  $\gamma$  diversity at both the transect within reef (reef, r) and reef within atoll (atoll, a) scales, using the slope of the  $\alpha$  vs.  $\gamma$  diversity relationship ( $p$ ) to estimate the degree of community similarity ( $\beta = 1 - p$ ) at both scales. Slope values at the smaller reef scale ( $p_r$ ) were calculated from the mean  $\alpha$  diversity among transects within each reef type compared to the  $\gamma$  diversity of each reef type (sample sizes are given in Appendix B: Table B4); slopes at the larger, atoll scale ( $p_a$ ) were calculated by pooling transects within reef types and comparing mean  $\alpha$  diversity among reef types within each atoll to the  $\gamma$  diversity of the atoll. These relationships helped to determine the degree of aggregation or clustering among observed reef fish populations at the reef ( $p_{r,ob}$ ) and atoll ( $p_{a,ob}$ ) scales, where lower slope values represented greater levels of aggregation. By comparing mean  $p_r$  with mean  $p_a$ , we could determine the degree to which average reef-fish community similarity changed with a change in scale. We also investigated the effect of unbalanced sampling at the scales of reef and atoll on estimates of gamma diversity to ensure our results were not substantially biased (see Appendix A).

The second step in the hierarchical randomization analysis was to determine how observed slopes changed when the effects of among- and within-scale clustering were removed from the data structure at both the reef habitat and atoll scales. By randomizing species within each scale, the effects of clustering at each scale could be examined relative to a random or neutral distribution of fishes (sensu Cornell et al. 2007: Fig. 3). To remove the effects of within-habitat (wh) clustering from the observed data, individual fishes were shuffled among transects within reefs (each transect retaining the same fish abundance but with species identities randomly assigned from among the species observed in each reef habitat and in each atoll) and a new regression slope was calculated from the randomized data ( $p_{r,wh}$ ). To remove the effects of species clustering among reef habitats (ah) from the observed data, individual fishes were shuffled among transects across habitats within atolls (species identities randomly assigned from among the species observed within each atoll) and an additional regression slope was calculated among reef habitats ( $p_{r,wh+ah}$ ) from the second set of randomized data. Increases in slope values at each step quantified the level of species clustering attributable to the randomized hierarchical level. This process was repeated at the atoll scale, with differences among slopes being tested for statistical significance at each step.

#### *Hierarchical linear models*

To understand the potential ecological factors underlying the hierarchical structure of the reef-fish metacommunity, we developed hierarchical linear models for three gross metrics of fish metacommunity structure (species richness, total abundance, and total biomass) and one multivariate response (the first axis of a detrended correspondence analysis (DCA) of fish species

abundance; Jongman et al. 1995) for each transect, representing the dominance of associated fish species within the metacommunity. We chose species richness as a direct measure of observed diversity, as opposed to transformed response variables such as evenness; species richness remains an exceptionally common and intuitive measure that is directly informative about the state of a given fish community while evenness combines the observed response into a single value representing multiple quantities (abundance and species richness) that can arise from distinct processes; we wished to explore these measures separately. Both abundance and biomass were log-transformed to fit the normality assumptions of our linear models. Although hierarchical models can readily accommodate nonlinearity and other parametric forms in a Bayesian context, we maintain a frequentist approach here for familiarity with conventional general linear models and to emphasize the theory behind a hierarchical approach. These response variables were further broken down into seven functional groups (detritivores, grazers, scrapevators [the combined “scraper” and “excavator” groups of Bellwood et al. 2004], planktivores, corallivores, invertivores, and piscivores) to understand the average functional-group composition of the observed trends.

We substantiated the hierarchical structure of the four response variables by initially fitting three random-intercept (means-as-outcomes) models that included transect-only, transect-within-atoll, and transect-within-reef-habitat-within-atoll levels of structure (but with no factors present) using the nlme (nonlinear mixed-effects) package (Pinheiro and Bates 2000) in R (R Development Core Team 2005). This allowed us to identify the degree to which response variables were clustered at each hierarchical level by determining whether observations from within groups (reef habitats, atolls) at each level were more closely related to their within-group means or to the group mean at the next, larger scale. We compared the fit of these simple hierarchical models through Akaike’s information criterion (AIC), a dimensionless measure of fit for each model, given the data, that can be used to compare relative empirical support between two or more statistical models; the lowest criterion-valued model is considered to have the most parsimonious fit (for details see Burnham and Anderson 2002).

Having established the structure of the data, we evaluated the relative importance of spatial separation, habitat specialization, and potential covariates of larval dispersal to reef-fish community structure using a candidate set of hierarchical linear models (Table 1; Appendix B: Table B5). For each process (hypothesis) we selected among available factors at the transect (T) and atoll (A) scales that would best provide evidence for the process given that it generated the observed patterns of community structure (see Appendix B). These included a Shannon index measure of habitat diversity that encompassed multiple habitat types along a given

TABLE 1. Candidate set of hierarchical linear models used to quantify patterns of species richness, abundance, and biomass in the reef-fish metacommunity of the Tuamotu Archipelago, French Polynesia.

Model†	Inferred process	Variance structure‡	K§	Equation
Null	none	RI	0	$Y_{ijk} = \beta_{000} + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
Mixed effects				
SS <sub>A,mixed</sub>	spatial separation	FS, RI	1	$Y_{ijk} = \beta_{000,k} + \beta_{001}APO_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
HS <sub>A,mixed</sub>	habitat specialization	FS, RI	1	$Y_{ijk} = \beta_{000,k} + \beta_{001}LSA_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
DP <sub>A,mixed</sub>	dispersal potential	FS, RI	1	$Y_{ijk} = \beta_{000,k} + \beta_{001}CON_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
HS <sub>T,mixed</sub>	habitat specialization	FS, RI	1	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}HSA_{ijk} + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
DP <sub>T,mixed</sub>	dispersal potential	FS, RI	1	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}EXP_{ijk} + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
SS <sub>T,mixed</sub>	spatial separation	FS, RI	2	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}EWT_{ijk} + \beta_{2jk}NST_{ijk} + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
SS <sub>TA,mixed</sub>	spatial separation	FS, RI	3	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}EWT_{ijk} + \beta_{2jk}NST_{ijk} + \beta_{001}APO_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
Random effects				
SSTA <sub>random</sub>	spatial separation	RS, RI	5	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}EWT_{ijk} + \beta_{2jk}NST_{ijk} + \beta_{001}APO_k + \beta_{101,k}EWT_{ijk} \times APO_k + \beta_{201,k}NST_{ijk} \times APO_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
HS <sub>TA,mixed</sub>	habitat specialization	FS, RI	2	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}HSA_{ijk} + \beta_{001}LSA_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
HS <sub>TA,random</sub>	habitat specialization	RS, RI	3	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}HSA_{ijk} + \beta_{001}LSA_k + \beta_{101,k}HSA_{ijk} \times LSA_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
DP <sub>TA,mixed</sub>	dispersal potential	FS, RI	2	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}EXP_{ijk} + \beta_{001}CON_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$
DP <sub>TA,random</sub>	dispersal potential	RS, RI	3	$Y_{ijk} = \beta_{000,k} + \beta_{1jk}EXP_{ijk} + \beta_{001}CON_k + \beta_{101,k}EXP_{ijk} \times CON_k + \varepsilon_{ijk} + \varepsilon_{0jk} + \varepsilon_{00k}$

† A = atoll level, T = transect level.

‡ RI = random intercept, FS = fixed slope, RS = random slope.

§ K = number of parameters.

|| APO = atoll position (latitude–longitude PCA axis score for center of atoll); LSA = lagoon surface area (km<sup>2</sup>); CON = atoll connectedness (km<sup>2</sup>); HAS = habitat Shannon diversity (*H'*); EXP = transect exposure (radians); EWT = transect East–West position (decimal degrees longitude); NST = transect North–South position (decimal degrees latitude).  $\beta$ 's indicate regression slopes by level;  $\varepsilon$ 's indicate normally distributed errors by level; *i* = transect *j* = reef, *k* = atoll. A × B denotes interaction of level-1 and level-2 factors.

transect. Reef habitats were considered an intermediate hierarchical level that was implicitly defined by its observed physical characteristics and we did not assign additional covariates to this level of organization. These multilevel factors were in turn combined in 12 candidate models representing each of the three processes of interest (Table 1).

For each process we developed four models that permitted variation among hierarchical levels to be quantified: (1) a mixed-effects model with fixed-effects for the slope of the transect-level variable and random intercepts; (2) a mixed-effects model with fixed-effects for the slope of the atoll-level variable and random intercepts; (3) a mixed-effects model with fixed transect- and atoll-level slopes but intercepts that could vary among atolls; and (4) a random-effects model in which both the slopes and intercepts of transect-level relationships could vary among atolls (Raudenbush and Byrk 2002). This small set of models allowed us to efficiently evaluate the influence of factors among levels and the level of support for each of the three processes of interest. Importantly, the restricted number of factors per model accurately represented the structure of the data and the degrees of freedom present at each hierarchical level. As there were 10 atolls present in our survey we could only reasonably model one atoll-level variable at a time (Raudenbush and Byrk 2002), constraining our ability to model interactions among factors at the atoll level.

We evaluated the relative levels of support for each hypothesis using AIC with a small-sample bias adjustment (AIC<sub>c</sub>; Hurvich and Tsai 1989). We also calculated an AIC<sub>c</sub>-based relative-importance weight (*w<sub>i</sub>*) that indicated the level of support for each model, given the other models considered and the data (Burnham and Anderson 2002). Because model averaging is not appropriate for mixed-effects models (Rieman et al. 2006), we reported fixed- and random-effect estimates for the top-ranked models of each response, including any models with >15% of the model-based support from the *w<sub>i</sub>* results (Rieman et al. 2006).

To facilitate interpretation of the magnitude of fixed and random effects we plotted the empirical Bayes (shrinkage) estimates from the top-ranked model for each response across the range of local and atoll-level variables observed in our samples (Snijders and Bosker 1999). Model goodness-of-fit (GOF) was evaluated for the likelihood ratio of each model relative to a  $\chi^2$ -distribution in a three-level, intercept-only (null) model. The precision of estimates for each model was estimated from 90% confidence intervals evaluated in the *confit* routine within the *nlme* package in R.

## RESULTS

### *Hierarchical data structure*

Observed reef-fish community similarity increased with spatial scale (Fig. 2); the atoll-level regression slope ( $p_a = 0.57$ ) was significantly higher than the reef-level

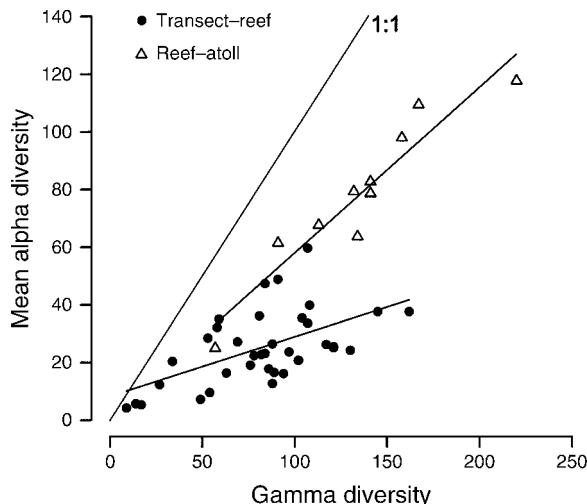


FIG. 2. Linear regressions of average  $\alpha$  and  $\gamma$  diversity at the transect-reef, and reef-atoll scales from 10 atolls within the Tuamotu Archipelago.

slope ( $p_r = 0.21$ ;  $t = 3.580$ ,  $n = 10$  data points;  $P = 0.005$ ) showing that a large proportion of among-atoll ( $\gamma$ ) diversity was subsumed by within-transect ( $\alpha$ ) diversity at the larger scale. The effects of patchiness on community similarity were significant at both scales but in different ways. At the reef-habitat scale, the majority of the species clustering present was attributable to within-habitat (transect-level) clustering, with  $p_{r,wh}$  and  $p_{r,wh+ah}$  being  $>3$  times greater than  $p_{r,ob}$  (Table 2). At the atoll scale, the effect of within-atoll (reef habitat-level) clustering was negligible and the slope attributable to among-atoll clustering was 61% greater than the observed slope values, showing a high level of clustering occurred at the atoll level. These results indicated that most of the nonrandom meta-community structure resided at the transect and atoll scales and that species appeared to be randomly dispersed within reef habitats in a given atoll.

The hierarchical linear model results indicated that 99.9% of the total support from the data favored three levels of hierarchical structure (Appendix B: Table B6). In addition, by hierarchically structuring the data, the three-level, intercept-only model “explained”  $\sim 55\%$  of the total variation in each response, given the level of reduced model deviance. The proportions of total variance per hierarchical level varied by response type (Table 3); species richness and log(abundance) were more evenly distributed among levels than log(biomass) or DCA of fish abundance, which were dominated by transect and inter-atoll variability. There was considerable uncertainty in the DCA estimates however, suggesting that the ordination had somewhat skewed the hierarchical structure present in the raw variables.

Average functional-group contributions to the response variables varied greatly by response type (Fig. 3). Relative species richness was dominated by invertivores at low diversities, but this declined with an increase in the proportion of planktivorous species up to about 30 species per transect; relative functional-group proportions stabilized beyond 30 species per transect, suggesting high functional redundancy at the most diverse locations (Fig. 3a). A dominance of invertivores evident at mid-log(abundance) levels was offset by a sharp increase in planktivores at the upper end of the log(abundance) range (Fig. 3b). Log(biomass) had relatively constant functional-group proportions across the range of values (Fig. 3c), whereas the DCA ordination clearly represented a shift from planktivore-dominated to invertivore-dominated transects (Fig. 3d).

*Hierarchical models*

Model-selection results for species richness strongly favored the random-intercept habitat-specialization model ( $HS_{TA,mixed}$ , Table 1) that incorporated both lagoon size and habitat diversity, with a high level ( $>97\%$ ) of statistical support (Table 4). A normal probability plot showed adequate model fit. Although the relationship between species richness and habitat

TABLE 2. The effects of (1) within-habitat specialization and (2) habitat specialization, and (3) the combined effects of within-habitat spatial separation and habitat specialization on community similarity at two scales within the Tuamotu reef-fish metacommunity.

Model†	Data structure	$\Delta \text{sim}‡$	SE	t	P
<b>Reef scale</b>					
$p_{r,ob} - 0$	observed vs. 1:1 line	0.207	0.051	4.093	<0.001
$p_{r,wh} - p_{r,ob}$	(1) randomized within reef vs. observed [transects within reefs]	0.660	0.078	8.473	<0.001
$p_{r,wh} - p_{r,wh+ah}$	(2) randomized within reefs vs. randomized within and among reefs	0.024	0.109	0.216	0.830
$p_{r,wh+ah} - p_{r,ob}$	(3) randomized within and among reefs vs. observed	0.683	0.105	6.527	<0.001
<b>Atoll scale</b>					
$p_{a,ob} - 0$	observed vs. 1:1 line	0.575	0.069	8.369	<0.001
$p_{a,wh} - p_{a,ob}$	(1) randomized within atolls vs. observed [reefs within atolls]	0.002	0.097	0.015	0.988
$p_{a,wh} - p_{a,wh+ah}$	(2) randomized within atolls vs. randomized within and among atolls	0.352	0.082	4.276	0.003
$p_{a,wh+ah} - p_{a,ob}$	(3) randomized within and among atolls vs. observed	0.353	0.082	4.309	0.003

Note: The use of (1)–(3) in the table caption and the “Data structure” column is intended to convey the link between the descriptions in the table caption and the data-structure descriptions in the body of the table.

† Key:  $p_r$  = reef-level regression slope,  $p_a$  = atoll-level regression slope; ob = observed, wh = within habitat, ah = among habitats.

‡  $\Delta \text{sim}$  is the difference between regression slopes defined by data structure.

TABLE 3. Hierarchical levels and proportion of variation in response variables explained by the hierarchical structure of transect, reef, and atoll spatial scales from the Tuamotu Archipelago.

Scale	Richness	log(abundance)	log(biomass)	DCA
Transect	0.38 (0.35, 0.40)	0.38 (0.15, 0.62)	0.38 (0.24, 0.53)	0.41 (0.00, 0.83)
Reef	0.40 (0.38, 0.42)	0.24 (0.00, 0.49)	0.09 (0.00, 0.27)	0.10 (0.00, 0.60)
Atoll	0.22 (0.20, 0.25)	0.37 (0.14, 0.61)	0.51 (0.38, 0.68)	0.49 (0.08, 0.90)

Notes: Variance proportions and 95% confidence bounds (in parentheses) are estimated from intercept-only, three-level hierarchical models of species richness, log(abundance), log(biomass), and DCA of fish abundance.

diversity was constant among atolls (fixed-effects model parallel slopes; Appendix B: Table B7), mean (within-atoll) richness varied strongly by lagoon size (Fig. 4a). The deviance reductions from the three-level intercept-only model suggested lagoon size explained 99% of the inter-atoll variation whereas habitat diversity accounted for only 10% of the variation at the transect level. Yet lagoon size was not the sole factor supporting the habitat specialization models as, when modeled alone

( $HS_{A,mixed}$ ), it had a low level of support from both information criteria, relative to other candidate models.

The distribution of log(biomass) showed the highest level of  $AIC_C$  support for the random-intercept habitat specialization model ( $HS_{TA,mixed}$ ), but also some support for the importance of transect-level habitat diversity alone ( $HS_{T,mixed}$ ). From these results we considered both models as credible, but with  $HS_{TA,mixed}$  having twice the level of  $AIC_C$ -based support in observed log(biomass).

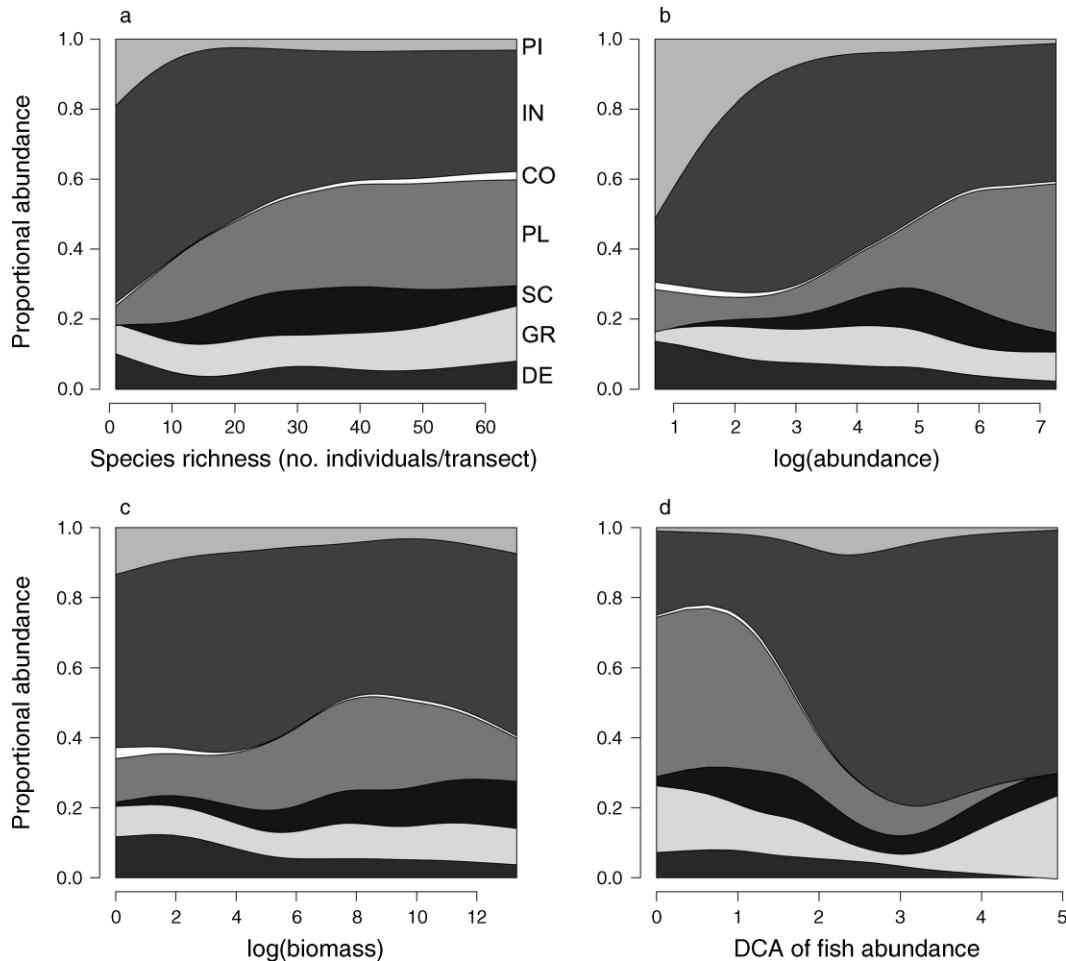


FIG. 3. Proportional contributions of reef-fish functional groups to (a) species richness, (b) log(abundance), (c) log(biomass), and (d) DCA (detrended correspondence analysis) of fish abundance. Plots are spline trends ( $df=6$ ) of average relative proportions (by abundance) of detritivores (DE), grazers (GR), scrapevators (SC; i.e., scrapers + excavators), planktivores (PL), corallivores (CO), invertivores (IN), and piscivores (PI). The order of the groups plotted is the same in each panel.

Both models demonstrated adequate model fit in normal probability plots. Habitat diversity and lagoon size accounted for ~20% of the reef habitat-level and atoll-level variation, respectively. The mean log(biomass) responses to transect-level habitat diversity were substantially different among atolls, with exceptionally high biomass in large, high-diversity lagoons (Fig. 4b).

The DCA response also supported the  $HS_{TA,mixed}$  and  $HS_{T,mixed}$  models, making both credible explanations for the planktivore–invertivore axis evident in the response (Fig. 3d, Table 4). For both models, there was a negative relationship between transect-level habitat diversity and DCA, indicating that planktivore-dominated transects were observed on more diverse habitats. Once again, model fit was adequate, but little (<7%) of the variation at the reef habitat or atoll scales was accounted for by either model. This was likely related to the uncertainty in the hierarchical structure of the ordination-derived response.

Patterns of log(abundance) were less well represented by all of the candidate models, with GOF values near the  $\alpha = 0.05$  level (Table 4). In particular,  $AIC_c$  values scarcely differed among the highest-ranked models, showing that the candidate models did not strongly favor one process over another. The lack of agreement between species richness and abundance model-selection results (Table 4) indicated that the patterns of species richness were not a sampling effect of abundance but were generated by distinct processes. To explore the abundance patterns further, we ran the log(abundance) of functional groups against the 12 models in the candidate model set.

Functional-group results fell into three process-based groups. The detritivores and grazers were correlated with larval dispersal covariates, with dispersal-potential (DP)-based models having >98% of the total support in the data (Table 5); relationships were consistently positive for both groups (Appendix B: Table B8). Scrapevators, planktivores, corallivores, and piscivores were all correlated with habitat specialization, where  $HS_{TA,mixed}$ ,  $HS_{T,mixed}$ , and  $HS_{A,mixed}$  were the highest-ranked models. Each functional group was positively correlated with transect-level habitat diversity and lagoon size (Appendix B: Table B8). This also supported the functional difference between grazer and scrapevator groups. Finally, invertivores were alone in being most strongly correlated with spatial position in model  $SS_{T,mixed}$ .

## DISCUSSION

Our hierarchical approach to understanding metacommunity structure demonstrates that transect- and atoll-level factors determine the level of nonrandom clustering within a marine metacommunity and that habitat is the prominent covariate associated with community characteristics at both scales. Importantly, these results demonstrate how modeling hierarchical

TABLE 4. Results of model selection for top-ranked hierarchical models of species richness, log(abundance), and log(biomass) on transects within the Tuamotu Archipelago.

Model	Summary statistics					
	GOF <sup>†</sup>	log $L$ <sup>‡</sup>	K <sup>§</sup>	$AIC_c$	$\Delta i$ <sup>  </sup>	$w_i$
Richness						
$HS_{TA,mixed}$	0.000	1876.0	2	3764.0	0.00	0.974
$DP_{TA,mixed}$	0.000	1880.4	2	3772.8	8.83	0.012
log(abundance)						
$SS_{TA,mixed}$	0.02	649.6	2	1311.2	0.00	0.179
$DP_{T,mixed}$	0.05	650.6	1	1311.2	0.05	0.175
$HS_{TA,mixed}$	0.02	649.7	2	1311.3	0.06	0.174
$DP_{TA,mixed}$	0.02	649.7	2	1311.4	0.12	0.169
log(biomass)						
$HS_{TA,mixed}$	0.000	977.4	2	1966.8	0.00	0.599
$HS_{T,mixed}$	0.001	978.8	1	1967.7	0.89	0.384
$SS_{TA,mixed}$	0.029	982.3	2	1976.5	9.78	0.005
DCA#						
$HS_{T,mixed}$	0.011	456.6	1	923.3	0.00	0.463
$HS_{TA,mixed}$	0.010	456.5	2	925.0	1.78	0.191
$DP_{T,mixed}$	0.04	458.0	1	926.0	2.71	0.054

Note: Model equations are given in Table 1.

<sup>†</sup> A likelihood ratio test goodness of fit ( $P$  values) relative to three-level intercept-only model (GOF). Models with GOF  $P > 0.05$  are not shown.

<sup>‡</sup> Model log likelihood.

<sup>§</sup> Number of parameters.

<sup>||</sup> Relative  $AIC_c$  differences.

<sup>¶</sup> Aikake weights.

<sup>#</sup> Detrended correspondence analysis of fish species abundance.

data and processes can aid in understanding reef-fish metacommunity structure within a well-defined system.

### Metacommunity structure

Evidence for a strong relationship between small- and large-scale processes has been elusive, although often suggested, in metacommunity studies. For instance, coral community structure has been deemed more variable than predicted by neutral theory due to “spatio-temporal environmental stochasticity” (Dornelas et al. 2006) and nonrandom clustering (Cornell et al. 2007). Yet these explanations are only the first step towards understanding metacommunity processes, as observed patterns are undoubtedly the result of multiple processes interacting at multiple scales through space and time (Clark et al. 2007).

We observed patterns of reef ( $\alpha$ ) and atoll ( $\gamma$ ) diversity similar to those reported for corals (Cornell et al. 2007), where community similarity increased with spatial scale and randomized community similarity was higher than observed similarity at both scales. Similarly, we also found differences between within- and among-scale diversity, with greater community similarity at larger scales and both high transect-level and high atoll-level clustering. But, contrary to the previous work, we did not assume a single process for each level of randomization—randomization tests alone do not address why species occur in nonrandom clusters. We attempted to

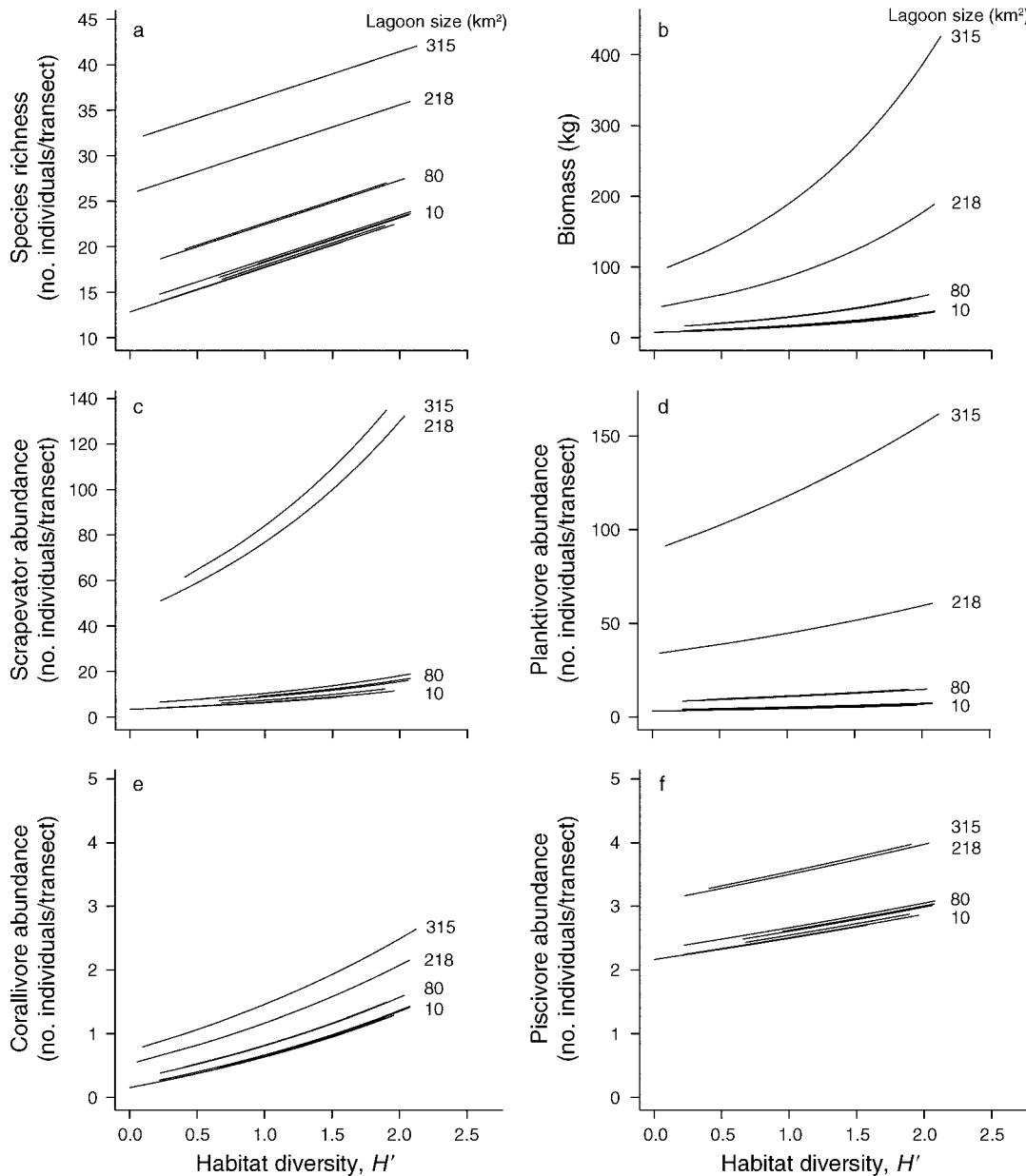


FIG. 4. Empirical Bayes estimates from the mixed-effect habitat-specialization model ( $HS_{TA,mixed}$ ) of the relationship between lagoon size and habitat Shannon diversity ( $H'$ ) with (a) species richness, (b) total biomass, (c) scrapevator abundance, (d) planktivore abundance, (e) corallivore abundance, and (f) piscivore abundance for 10 atolls in the Tuamotu Archipelago. Model equations and parameter estimates are given in Table 1 and Appendix B: Tables B3 and B6.

explicitly connect the observed patterns to potential mechanisms.

The importance of habitat availability and diversity in determining metacommunity structure supports a number of recent analyses identifying habitat structure as a dominant driver of reef-fish diversity at both reef (Graham et al. 2006) and regional (Resetarits 2005) scales. Although local species richness has been shown to correlate with both site environmental quality and the broad-scale distribution of climate conditions determining overall ranges (Connolly 2005), variation in reef-fish

biodiversity at the ocean scale can also depend on the availability of appropriate shallow-water habitat (Bellwood and Hughes 2001). Our analysis supports such connections among scales by identifying the role of transect-scale habitat characteristics in patterns of atoll-scale metacommunity structure.

The importance of habitat structure has been observed in other aquatic systems, where freshwater studies have repeatedly found that the small-scale diversity and structure of metacommunities are driven by habitat factors rather than regional limits on dispersal (Cottenie

et al. 2003, Urban 2004). Such agreement between freshwater and marine systems is not surprising because, although dispersal is the process that defines the boundaries of a given metacommunity, it is not necessarily the single large-scale process that is most likely to lead to a nonrandom distribution of species. Indeed, experimental metacommunity studies in sea-grass systems have shown that dispersal can increase community similarity among local communities (lower  $\beta$  diversity) and within regions (lower  $\gamma$  diversity) even when small-scale diversity ( $\alpha$  diversity) is constant (France and Duffy 2006). Similar results have been observed in pitcher-plant communities, where high dispersal levels were correlated with higher levels of community similarity but with the caveat that regional diversity ( $\gamma$  diversity) also increased (Kneitel and Miller 2003). Regardless of its effects on regional diversity, both studies suggest a limited role for dispersal in generating nonrandom clustering among individual communities, a conclusion supported by our analysis.

Large-scale patterns in reef-fish community structure have been linked to lagoon size in French Polynesian atolls previously (Galzin et al. 1994), but our analysis suggests further that these trends are driven primarily by increases in the number and abundance of planktivores, where larger lagoons support a larger and more diverse planktivore assemblage, potentially due to higher overall productivity (Charpy and Blanchot 1998). Such increases are likely to have additional effects that drive the positive correlation between piscivore abundance and lagoon size (Appendix B: Table B8), with increases in prey species paralleled by an increase in predators. In short, lagoon size appears to promote higher mean levels of fish biomass, abundance, and diversity throughout the archipelago, setting a baseline for differences attributable to site-level characteristics or other potential processes.

Our random-intercept models indicate that transect-level habitat diversity is correlated with species richness and abundance for several functional groups at the transect level within the context of larger-scale factors (Fig. 4). Corallivores were most strongly associated with increases in transect habitat diversity (Appendix B: Table B8), reflecting the dependence of these groups on hard corals that were prevalent on the more diverse transects, whereas the poor relationship between grazers and detritivores reflects independence from diverse habitats, a characteristic observed elsewhere by their relative success following habitat losses due to coral bleaching (Wilson et al. 2006). The bulk of our analysis suggests habitat composition is the primary determinant of nonrandom clustering within the Tuamotu fish metacommunity.

In terms of alternative hypotheses, we found little support for spatial separation affecting metacommunity structure. Only invertivores were spatially segregated, demonstrating their independence from specific habitat types or potential dispersal covariates. The resources

TABLE 5. Results of model selection for top-ranked hierarchical models of functional-group  $\log(\text{abundance})$  among transects within the Tuamotu Archipelago.

Model	Summary statistics					
	GOF	$\log L$	$K$	AICc	$\Delta i$	$w_i$
<b>Detritivores</b>						
DP <sub>T,mixed</sub>	0.000	801.3	1	1612.6	0.00	0.655
DP <sub>TA,mixed</sub>	0.000	801.2	2	1614.5	1.92	0.250
DP <sub>TA,random</sub>	0.000	801.2	2	1616.5	3.92	0.092
<b>Grazers</b>						
DP <sub>T,mixed</sub>	0.000	810.3	1	1630.7	0.00	0.667
DP <sub>TA,mixed</sub>	0.000	810.3	2	1632.6	1.99	0.243
DP <sub>TA,random</sub>	0.000	810.3	2	1634.6	3.99	0.089
<b>Scrapevators</b>						
HS <sub>TA,mixed</sub>	0.000	801.4	2	1614.8	0.00	0.470
HS <sub>T,mixed</sub>	0.001	802.7	1	1615.5	0.68	0.335
HS <sub>TA,random</sub>	0.000	801.4	2	1616.8	2.00	0.173
DP <sub>T,mixed</sub>	0.022	806.1	1	1622.4	7.65	0.001
<b>Planktivores</b>						
HS <sub>TA,mixed</sub>	0.010	903.2	2	1818.3	0.00	0.298
HS <sub>A,mixed</sub>	0.040	904.6	1	1819.1	0.819	0.198
HS <sub>T,mixed</sub>	0.010	903.1	2	1820.3	2.00	0.110
DP <sub>TA,mixed</sub>	0.030	904.3	2	1820.5	2.227	0.036
<b>Corallivores</b>						
HS <sub>T,mixed</sub>	0.000	546.7	1	1103.8	0.00	0.327
HS <sub>TA,mixed</sub>	0.000	546.0	2	1103.9	0.07	0.316
HS <sub>TA,random</sub>	0.000	546.0	2	1105.9	2.07	0.116
DP <sub>TA,mixed</sub>	0.000	547.1	2	1106.1	2.26	0.106
<b>Invertivores</b>						
SS <sub>T,mixed</sub>	0.024	678.4	2	1368.7	0.00	0.263
SS <sub>TA,mixed</sub>	0.012	677.7	3	1369.4	0.65	0.190
SS <sub>TA,random</sub>	0.011	677.6	5	1371.3	2.47	0.075
<b>Piscivores</b>						
HS <sub>A,mixed</sub>	0.008	609.6	1	1229.2	0.000	0.482
HS <sub>TA,mixed</sub>	0.004	609.0	2	1230.0	0.806	0.322
HS <sub>TA,random</sub>	0.004	609.0	2	1232.0	2.806	0.185

Notes: Model abbreviations and summary statistics are as defined in Table 4. Model equations are given in Table 1.

used by invertivores are not strongly linked to habitat complexity and even simple detritic rubble substrates can provide sufficient prey for many species (Ferreira et al. 2004).

Our exposure-based dispersal models of recruitment potential among atolls were only correlated with detritivore and grazer abundance patterns. Like invertivores, the resources exploited by these functional groups would not be expected to be strongly linked to habitat complexity or diversity (unless wide-scale collapse of structural complexity occurs; e.g., Graham et al. 2007), making potential dispersal processes relatively more important than for other functional groups. Although previous studies have asserted that the supply of both coral and fish recruits are expected to diminish through isolation (Ayre and Hughes 2004, Cowen et al. 2006), the metacommunity we studied is likely too well connected (Planes et al. 1996) to detect such trends in other functional groups or may be driven by a substantial level of atoll-level self-recruitment (Jones et al. 1999) that was unaffected by prevailing

currents. Although larval dispersal is undoubtedly a critical component of any reef-fish community, we found no spatial evidence that exposure leads to nonrandom clustering in this system, largely agreeing with population-genetics data from within the region (Planes et al. 1996, Planes and Fauvelot 2002).

It is important to note, however, that if stochastic (nonspatial) recruitment is the primary determinant of small-scale species abundances through time, species-specific dynamics may become obscured in an observational study from a single point in time, providing little to no signal for dispersal. This also suggests that the size structuring of the metacommunity is decoupled among size groups, with small, potentially fast-turnover species dominating abundance and likely governed by variable recruitment, and larger species being controlled more by atoll lagoon size. Although we did not measure individual fishes, our functional-group results support this conclusion, where planktivores drove trends in total abundance and piscivores were most correlated with lagoon size. With this in mind, isolation of individual atolls may not be great enough to be limiting community structure; however, a disturbance of sufficient spatial scale to impact the entire archipelago could have long-lasting effects due to its isolation (Graham et al. 2006).

#### *Hierarchical analysis*

Many reef studies focus on one scale of analysis to simplify and strengthen the conclusions made, whereas others report results at multiple scales when the analysis has been conducted only at one (Karlson and Cornell 1999). However, explicitly hierarchical analyses help to quantify the role of multiple processes operating at different ecological scales, by properly specifying uncertainty surrounding the quantities of greatest ecological interest—the effect sizes. Previously, hierarchical analysis has effectively evaluated the relative evidence for recruitment limitation, density-independent mortality, and density-dependent mortality for Polynesian labrid recruits (Shima 1999). The relative importance of processes varied across all scales; at some times and locations density dependence was important, while at others, recruits were dominated by recruitment limitation (Shima 1999). Therefore, the hierarchical approach led to a clear and dynamic description of multiple processes interacting through time.

Although the low number of atolls sampled limited our ability to explore atoll-level interactions, had we used a general linear model, with no hierarchical structure (e.g., the same factors run with common slopes and variance among reefs and atolls) to quantify the full habitat-specialization model, the effect of habitat diversity would have been inflated by 160%—primarily due to the larger atolls having higher average species richness and habitat diversity. Ignoring the hierarchical structure of ecosystems and data can frequently lead to gross underestimates of standard errors that give an inflated level of precision to poorly structured analyses

(Raudenbush and Byrk 2002). This is one potential reason why improper fits between models and data are prevalent; researchers develop a false sense of confidence in their results once they have achieved an acceptable level of explained variance, often when  $R^2$  values are  $>0.50$ . But by this criterion our analysis was able to “explain”  $>55\%$  of the variation in three response metrics using intercept-only models with no factors present—passing the generally acceptable level of precision for large-scale observational studies.

A major temptation in regression-based statistics is to maximize the variance explained in a given analysis once the data are familiar and to ‘improve’ interpretation through data mining. However, the more ecologically interesting questions are about the processes operating in a system and the effect sizes of factors that represent them. We agree with a growing number of ecologists that a focus on “variance explained” criteria should be redirected into ecologically meaningful model comparisons where the data are used to arbitrate among hypotheses (Hilborn and Mangel 1997, Johnson 1999). This approach requires a substantive knowledge of the system under study and careful consideration of multiple processes at multiple scales, but also tends to promote greater care prior to analysis. Although snapshot ecological data discount the temporal variability present in ecological systems and potentially misrepresent the importance of dynamic processes (Shima 1999), static pattern data are often the best information available, especially at large scales. Under such conditions using multiple working hypotheses is often the best available approach. Ecologists who know something about the system under study can, and should, use that knowledge to generate a priori hypotheses about patterns of static data to see which hypotheses have any degree of empirical support, thereby maximizing the inferential power of pattern-based analyses.

#### ACKNOWLEDGMENTS

We thank Jim Peterson and Wade Blanchard for guidance on hierarchical modeling. We also thank two anonymous reviewers who made considerable efforts to improve the manuscript. This research was funded through doctoral studentships from the National Science and Engineering Research Council of Canada and the Fisheries Society of the British Isles. The Tyatoll research programme was supported financially by the Institut de Recherches en Développement, the French National Programme on Coral Reefs, and the French Polynesian Territory. Thanks are expressed to G. Mou-Tham, and the officers and crew of RV *Alis* for their contributions to data collection. This paper is dedicated to the memory of our colleague Ram Myers, who many of us were proud to call our friend.

#### LITERATURE CITED

- Allen, T., and T. Hoekstra. 1992. *Toward a unified ecology. Complexity in ecological systems.* Columbia University Press, New York, New York, USA.
- Allen, T., and T. Starr. 1982. *Hierarchy: perspectives for ecological complexity.* University of Chicago Press, Chicago, Illinois, USA.

- Almany, G. 2004. Does increased habitat complexity reduce predation and competition in coral reef fish assemblages? *Oikos* 106:275–284.
- Ault, T., and C. Johnson. 1998. Spatially and temporally predictable fish communities on coral reefs. *Ecological Monographs* 68:25–50.
- Ayre, D., and T. Hughes. 2004. Climate change, genotypic diversity and gene flow in reef-building corals. *Ecology Letters* 7:273–278.
- Bellwood, D., and T. Hughes. 2001. Regional-scale assembly rules and the biodiversity of coral reefs. *Science* 292:1532–1534.
- Bellwood, D., T. Hughes, C. Folke, and M. Nyström. 2004. Confronting the coral reef crisis. *Nature* 429:827–833.
- Burnham, K., and D. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer-Verlag, New York, New York, USA.
- Caley, M. J., M. Carr, M. Hixon, T. Hughes, G. Jones, and B. Menge. 1996. Recruitment and the local dynamics of open marine populations. *Annual Reviews in Ecology and Systematics* 27:477–500.
- Chabanet, P., H. Ralambondrainy, M. Amanieu, G. Faure, and R. Galzin. 1997. Relationships between coral reef substrata and fish. *Coral Reefs* 16:93–102.
- Charpy, L., and J. Blanchot. 1998. Photosynthetic picoplankton in French Polynesian atoll lagoons: estimation of taxa contribution to biomass and production by flow cytometry. *Marine Ecology Progress Series* 162:57–70.
- Clark, J., M. Dietze, S. Chakarabarty, P. Agarwal, I. Ibanez, S. LaDeau, and M. Wolosin. 2007. Resolving the biodiversity paradox. *Ecology Letters* 10:1–13.
- Connolly, S. 2005. Process-based models of species distributions and the mid-domain effect. *American Naturalist* 166:1–11.
- Cornell, H., R. Karlson, and T. Hughes. 2007. Scale-dependent variation in coral community similarity across sites, islands, and island groups. *Ecology* 88:1707–1715.
- Cottenie, K., E. Michels, N. Nuytten, and L. De Meester. 2003. Zooplankton metacommunity structure: regional vs. local processes in highly interconnected ponds. *Ecology* 84:991–1000.
- Cowen, R., C. Paris, and A. Srinivasan. 2006. Scaling of connectivity in marine populations. *Science* 311:522–527.
- Doherty, P. 1981. Coral reef fishes: recruitment-limited assemblages? Pages 465–479 in E. D. Gomes, C. E. Birke-land, R. W. Buddemeier, R. E. Johnson, J. A. March, Jr., and R. T. Tsuda, Proceedings of the Fourth International Coral Reef Symposium, 18–22 May 1981, volume 2. Marine Sciences Center, University of the Philippines, Quezon City, The Philippines.
- Doherty, P. 2002. Variable replenishment and the dynamics of reef fish populations. Pages 327–355 in P. F. Sale, editor. *Coral reef fishes: dynamics and diversity in a complex ecosystem*. Academic Press, San Diego, California, USA.
- Doherty, P., and T. Fowler. 1994. An empirical test of recruitment limitation in a coral reef fish. *Science* 263:935–939.
- Dornelas, M., S. Connolly, and T. Hughes. 2006. Coral reef diversity refutes the neutral theory of biodiversity. *Nature* 440:80–83.
- Dufour, P., and M.-L. Harmelin-Vivien. 1997. A research program for a typology of atoll lagoons: strategy and first results. Pages 843–848 in H. A. Lessios and I. G. Macintyre, editors. *Proceedings of the Eighth International Coral Reef Symposium*, volume 1. Smithsonian Tropical Research Institute, Balboa, Panama.
- English, S., C. Wilkinson, and V. Barker. 1994. Survey manual for tropical marine resources. Technical report, ASEAN-Australia Science Project. Australian Institute of Marine Science, Townsville, Australia.
- Ferreira, C. E. L., S. R. Floeter, J. L. Gasparini, B. P. Ferreira, and J. C. Joyeux. 2004. Trophic structure patterns of Brazilian reef fishes: a latitudinal comparison. *Journal of Biogeography* 31:1093–1106.
- France, K. E., and J. E. Duffy. 2006. Diversity and dispersal interactively affect predictability of ecosystem function. *Nature* 441:1139–1143.
- Freestone, A. L., and B. D. Inouye. 2006. Dispersal limitation and environmental heterogeneity shape scale-dependent diversity patterns in plant communities. *Ecology* 87:2425–2432.
- Galzin, R., S. Planes, P. Dufour, and B. Salvat. 1994. Variation in diversity of coral reef fish between French Polynesian atolls. *Coral Reefs* 13:175–180.
- Graham, N. A. J., S. K. Wilson, S. Jennings, N. V. C. Polunin, J. P. Bijoux, and J. Robinson. 2006. Dynamic fragility of oceanic coral reef ecosystems. *Proceedings of the National Academy of Sciences (USA)* 103:8425–8429.
- Graham, N. A. J., S. K. Wilson, S. Jennings, N. V. C. Polunin, J. Robinson, J. P. Bijoux, and T. M. Daw. 2007. Lag effects in the impacts of mass coral bleaching on coral reef fish, fisheries, and ecosystems. *Conservation Biology* 21:1291–1300.
- Gratwicke, B., and M. Speight. 2005. Effects of habitat complexity on Caribbean marine fish assemblages. *Marine Ecology Progress Series* 292:310–310.
- Gust, N., J. Choat, and M. McCormick. 2001. Spatial variability in reef fish distribution, abundance, size and biomass: a multi-scale analysis. *Marine Ecology Progress Series* 214:237–251.
- Hatcher, B. 1997. Coral reef ecosystems: How much greater is the whole than the sum of the parts? *Coral Reefs* 16:S77–S91.
- Hilborn, R., and M. Mangel. 1997. The ecological detective. Number 28 in *Monographs in Population Biology*. Princeton University Press, Princeton, New Jersey, USA.
- Hixon, M., and J. P. Beets. 1993. Predation, prey refuges, and the structure of coral-reef fish assemblages. *Ecological Monographs* 63:77–101.
- Hurvich, C., and C. Tsai. 1989. Regression and time series model selection in small samples. *Biometrika* 76:297–307.
- Johnson, D. 1999. The insignificance of statistical significance testing. *Journal of Wildlife Management* 63:763–772.
- Jones, G., M. McCormick, M. Srinivasan, and J. Eagle. 2004. Coral decline threatens fish biodiversity in marine reserves. *Proceedings of the National Academy of Sciences (USA)* 101:8251–8253.
- Jones, G., M. Milicich, M. Emslie, and C. Lunow. 1999. Self-recruitment in a coral reef fish population. *Nature* 402:802–804.
- Jones, G., and C. Syms. 1998. Disturbance, habitat structure and the ecology of fishes on coral reefs. *Australian Journal of Ecology* 23:287–297.
- Jongman, R., J. Ter Braak, and O. Van Tongeren. 1995. *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge, UK.
- Karlson, R., and H. Cornell. 1999. Integration of local and regional perspectives on the species richness of coral assemblages. *American Zoologist* 39:104–112.
- Karlson, R. H., H. V. Cornell, and T. P. Hughes. 2007. Aggregation influences coral species richness at multiple spatial scales. *Ecology* 88:170–177.
- Kneitel, J. M., and T. E. Miller. 2003. Dispersal rates affect species composition in metacommunities of *Sarracenia purpurea* inquilines. *American Naturalist* 162:165–171.
- Kulbicki, M., N. Guillemot, and M. Amand. 2005. A general approach to length–weight relationships for Pacific lagoon fishes. *Cybius* 29:235–252.
- Liebold, M., M. Holyoak, N. Mouquet, P. Amraskare, J. Chase, M. Hoopes, R. Holt, J. Shurin, R. Law, D. Tillman, M. Loreau, and A. Gonzalez. 2004. The metacommunity concept: a framework for multi-scale ecology. *Ecology Letters* 7:601–6013.

- Mellin, C., J. Ferraris, R. Galzin, M. Kulbicki, and D. Ponton. 2006. Diversity of coral reef fish assemblages: modelling of the species richness spectra from multi-scale environmental variables in the Tuamotu Archipelago (French Polynesia). *Ecological Modelling* 198:409–425.
- Munday, P., and G. Jones. 1998. The ecological implications of small body size among coral-reef fishes. *Oceanography and Marine Biology Annual Review* 36:373–411.
- Pinheiro, J., and D. Bates. 2000. . Mixed-effects models in S and S-PLUS. Springer-Verlag, New York, New York, USA.
- Planes, S., and C. Fauvelot. 2002. Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific Ocean. *Evolution* 56:378–399.
- Planes, S., R. Galzin, and F. Bonhomme. 1996. A genetic metapopulation model for reef fishes in oceanic islands: the case of the surgeonfish, *Acanthurus trigostegus*. *Journal of Evolutionary Biology* 9:103–117.
- R Development Core Team. 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (<http://www.R-project.org>)
- Raudenbush, S., and A. Byrk. 2002. Hierarchical linear models: applications and data analysis methods. Advanced quantitative techniques in the social sciences. Sage, Thousand Oaks, California, USA.
- Resetarits, W., Jr. 2005. Habitat selection behaviour links local and regional scales in aquatic systems. *Ecology Letters* 8: 480–486.
- Rieman, B., J. Peterson, and D. Myers. 2006. Have brook trout (*Salvelinus fontinalis*) displaced bull trout (*Salvelinus confluentus*) along longitudinal gradients in central Idaho streams? *Canadian Journal of Fisheries and Aquatic Sciences* 63:63–78.
- Sale, P. 1980. Assemblages of fish on patch reefs—predictable or unpredictable? *Environmental Biology of Fishes* 5:243–249.
- Sale, P., and R. Dybdahl. 1975. Determinants of community structure for coral reef fishes in an experimental habitat. *Ecology* 56:1343–1355.
- Sale, P., J. Guy, and W. Steel. 1994. Ecological structure of assemblages of coral reef fishes on isolated patch reefs. *Oecologia* 98:83–99.
- Shima, J. 1999. Variability in relative importance of determinants of reef fish recruitment. *Ecology Letters* 2:304–310.
- Snijders, T., and R. Bosker. 1999. Multilevel analysis: an introduction to basic and advanced multilevel modelling. Sage, Thousand Oaks, California, USA.
- Steele, M., and G. Forrester. 2005. Small-scale field experiments accurately scale up to predict density dependence in reef fish populations at larger scales. *Proceedings of the National Academy of Sciences (USA)* 102:13513–13516.
- Syms, C., and G. Jones. 2000. Disturbance, habitat structure, and the dynamics of a coral-reef fish community. *Ecology* 81: 2714–2729.
- Syms, C., and G. Jones. 2004. Habitat structure, disturbance and the composition of sand-dwelling goby assemblages in a coral reef lagoon. *Marine Ecology Progress Series* 268:221–230.
- Urban, M. 2004. Disturbance heterogeneity determines freshwater metacommunity structure. *Ecology* 85:2971–2978.
- Wilson, S., N. Graham, M. Pratchett, G. Jones, and N. V. C. Polunin. 2006. Multiple disturbances and the global degradation of coral reefs: Are reef fishes at risk or resilient? *Global Change Biology* 12:2220–2234.

#### APPENDIX A

Descriptions of covariate development and associated statistical methods (*Ecological Archives* E090-016-A1).

#### APPENDIX B

Eight tables presenting proxy diversity estimates, data and model characteristics, and statistical results (*Ecological Archives* E090-016-A2).

#### APPENDIX C

A figure illustrating the observed relationship between percentage live coral cover and habitat Shannon diversity (*Ecological Archives* E090-016-A3).