

Variation in coral-associated cryptofaunal communities across spatial scales and environmental gradients

Chelsie W. W. Counsell¹  · Megan J. Donahue¹ · Kyle F. Edwards² · Erik C. Franklin¹ · Mark A. Hixon³

Received: 2 March 2018 / Accepted: 20 June 2018 / Published online: 5 July 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract Most of the diversity on coral reefs is in the cryptofauna, the hidden organisms that inhabit the interstitial spaces of corals and other habitat-forming benthos. However, little is known about the patterns and drivers of diversity in cryptofauna. We investigated how the cryptofaunal community associated with the branching coral *Pocillopora meandrina* varies across spatial scales and environmental gradients. We performed nondestructive visual surveys of the cryptofaunal community on 751 *P. meandrina* colonies around the island of O‘ahu (30–73 colonies per site, 3–6 sites per region, five regions). We identified 91 species, including 48 fishes and 43 invertebrates. Most of these species were observed rarely, with only 19 species occurring on greater than 5% of surveyed colonies. Variation in community abundance and species richness was greatest at the scale of the coral colony and lowest at the site scale. Abundance and species richness increased with increasing colony size and maximum wave height, and decreased with increasing surface chlorophyll-a. In an analysis of species-specific responses, colony size,

wave height, and chlorophyll-a were significant drivers of occurrence. Depth and percent live coral tissue were also identified as important correlates for community composition with distinct responses across taxa. Analyzing species-specific responses to environmental gradients documented a unique pattern for the guard crab *Trapezia intermedia*, which had a higher probability of occurring on smaller colonies (in contrast to 18 other common taxa). The results of a principal coordinates analysis on community composition and a co-occurrence analysis further supported *T. intermedia* as having a unique distribution across colonies, even in comparison with four other *Trapezia* species. Overall, these patterns emphasize the importance of host coral characteristics (i.e., colony size and percent live tissue) and physical characteristics of the surrounding habitat (i.e., wave energy, chlorophyll-a, and depth) in structuring cryptofaunal communities and characterize species-specific responses to environmental gradients.

Keywords Community ecology · Cryptofauna · Hawaiian Islands · *Pocillopora meandrina* · Environmental gradients · Spatial scales

Topic Editor Morgan S. Pratchett

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00338-018-1709-7>) contains supplementary material, which is available to authorized users.

✉ Chelsie W. W. Counsell
counsell@hawaii.edu

¹ School of Ocean and Earth Science Technology, Hawai‘i Institute of Marine Biology, University of Hawai‘i at Mānoa, Kāne‘ohe, HI 96744, USA

² Department of Oceanography, University of Hawai‘i at Mānoa, Honolulu, HI 96822, USA

³ Department of Biology, University of Hawai‘i at Mānoa, Honolulu, HI 96822, USA

Introduction

Most of the diversity on coral reefs is in the cryptofauna, the hidden species that inhabit the branches, crevices, and interstitial spaces of corals and other habitat-forming sessile organisms (Reaka-Kudla 1997; Plaisance et al. 2011). Reef-associated cryptofauna constitute 91% of the known species on coral reefs (Stella et al. 2010) and are a critical component of coral reef trophic webs. Cryptofauna capture and recycle nutrients by consuming very small prey items

(e.g., plankton, detritus, coral mucus), and they are a primary food source for many reef fishes, including squirrelfishes, wrasses, triggerfishes, snappers, and groupers (Randall 1967; Enochs 2012; Leray et al. 2015). Despite their abundance and importance, reef cryptofauna are under-represented in traditional reef surveys and, as a result, relatively little is known about the composition of these communities and the ecological processes that structure them.

One of the most diverse coral-associated cryptofaunal communities is associated with living and dead corals in the family Pocilloporidae (Stella et al. 2010). Pocilloporids are structurally complex, reef-building corals that are common, especially on exposed reefs, and widespread throughout much of the Indo-Pacific. The relatively small size and spatial isolation of individual *Pocillopora* colonies facilitates studies of discrete, replicate communities. Given the tractable nature of these communities, *Pocillopora*-associated assemblages have been the focus of research on species interactions and cryptofaunal distribution since the 1960s. Previous research on *Pocillopora*-associated communities has identified cryptofauna-coral host mutualisms for trapeziid crabs, alpheid shrimps, and damselfishes. Crabs in the genus *Trapezia* and shrimps in the genus *Alpheus* help protect their host corals from corallivores, including the predatory gastropod *Drupella cornus* (McKeon and Moore 2014), the cushion star *Calcita novaeguineae* (McKeon et al. 2012; McKeon and Moore 2014), and the crown-of-thorns seastar *Acanthaster planci* (Pratchett 2001; McKeon et al. 2012; McKeon and Moore 2014; Rouzé et al. 2014). These mutualistic decapods also increase the growth and survival of host corals by removing sediments from the coral tissue (Stewart et al. 2006; Stier et al. 2012; Stewart et al. 2013; Rouzé et al. 2014) and reducing negative interactions with vermetid snails (Stier et al. 2010). Damselfishes associated with *Pocillopora* colonies promote coral growth and survival through a variety of mechanisms, including territorial defense that minimizes predation from other reef fishes (Gochfeld 2009; Chase et al. 2014), sleep-swimming behavior that circulates water between coral branches at night (Goldshmid et al. 2004), and excretion-based nutrient enrichment (Holbrook et al. 2008). In addition to describing coral mutualisms, several studies have reported that some species in the coral-associated community can affect the occurrence and survival of other species through predation and territorial behavior (Schmitt et al. 2009; Holbrook et al. 2011; Stier et al. 2013; Stier and Leray 2014; Leray et al. 2015).

Given the focus of *Pocillopora* research on species interactions and the considerable effort required to exhaustively sample cryptofaunal communities, most surveys of *Pocillopora*-associated communities can be fit into

one of two categories: large visual surveys focused on a few key species (e.g., Sin and Lee 2000; Holbrook et al. 2008; Stier and Leray 2014), or thorough, albeit destructive, sampling limited to a small number of colonies (e.g., Austin et al. 1980; Coles 1980; Black and Prince 1983; Gotelli and Abele 1983; Britayev et al. 2017; López-Pérez et al. 2017). Previous surveys have identified host colony size as a strong correlate with the number of individuals and diversity of species in the associated community (Abele and Patton 1976; Stella et al. 2010; Holbrook et al. 2011). In addition, many studies have suggested that a decline in host coral health due to tissue bleaching or mortality shifts the composition of the decapod community from a few obligate species to a more diverse group of facultative species (Coles 1980; Stewart et al. 2006; Enochs and Hockensmith 2008; Plaisance et al. 2009; Stella et al. 2010, 2011; Enochs and Manzello 2012; Leray et al. 2012), although some obligate species, including *Trapezia* crabs, have been observed on dead coral colonies (Preston and Doherty 1990; Stella et al. 2011; Head et al. 2015).

While some research has suggested that *Pocillopora*-associated communities vary over space as a result of environmental drivers (Abele 1976; Austin et al. 1980; Black and Prince 1983; López-Pérez et al. 2017), previous studies were limited in spatial extent (two to four sites) and did not directly investigate the correlation between specific environmental factors and community composition. Environmental factors including depth, reef zone, and wave energy can drive shifts in community composition for corals (Franklin et al. 2013; Gove et al. 2015), non-cryptic reef fishes (Nunes et al. 2013; Jankowski et al. 2015; Darling et al. 2017), and cryptofaunal communities not directly associated with coral hosts (Klumpp et al. 1988; Depczynski and Bellwood 2005).

Here, we examined how environmental factors influence the composition of *Pocillopora*-associated cryptofaunal communities, providing a broader ecological context for the existing experimental work on these communities. We included measures of colony size and percent live tissue, factors highlighted in previous studies, to reflect the habitat quality of the host coral. We hypothesized that species exhibit taxon-specific responses across gradients of depth and wave energy, resulting in unique community compositions. We also considered the effect of host density, hypothesizing that the abundance of obligate species on each colony will decrease with increasing availability of adjacent host colonies. Further, we hypothesized that increasing benthic complexity at the site scale would decrease the species richness of the *Pocillopora*-associated community due to increased habitat structure available for facultative species. Finally, we investigated whether abundance or species richness increased with primary

productivity (measured as satellite-derived surface chlorophyll-*a*). While positive, negative, and unimodal patterns have been observed between productivity and species richness (Mittelbach et al. 2001), here we expected community abundance and species richness to increase with chlorophyll-*a* because some of the species feed directly on plankton and because our study sites are oligotrophic, i.e., where the positive part of a unimodal relationship between productivity and species richness would occur.

To analyze patterns over these environmental gradients, we nondestructively surveyed the communities associated with 751 *Pocillopora meandrina* colonies across 19 sites from five regions around the island of O‘ahu, Hawai‘i. We characterized the fish and invertebrate species found on *P. meandrina*, partitioned variation in the community across spatial scales, and quantified correlations between both community abundance and species richness, and possible environmental drivers. Further, we investigated patterns in community composition by identifying species-specific responses to environmental factors and non-random species co-occurrences.

Materials and methods

Surveys

To characterize the *P. meandrina*-associated community, we surveyed 751 colonies across 19 sites around the island of O‘ahu (Fig. 1) from September 2013 to March 2015. Sites were stratified into five regions (east, south, west, north, and Kāne‘ohe Bay), each characterized by different wave regimes, which is an important variable structuring

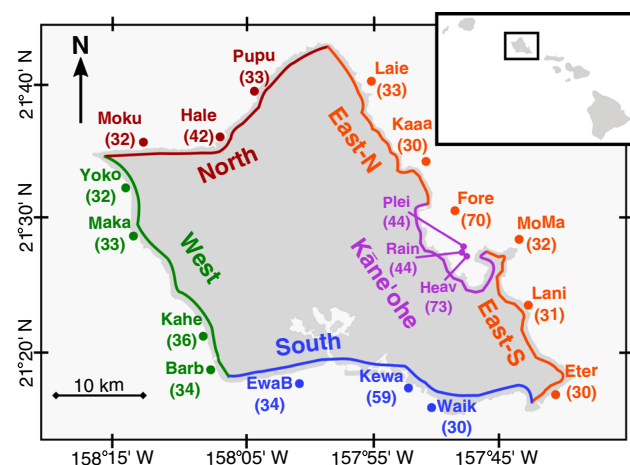


Fig. 1 Location and name codes of 19 study sites around O‘ahu. For each site, the number of colonies surveyed is given in parentheses. Sites are grouped by color into regions. Inset map of the main Hawaiian Islands with O‘ahu outlined. See Table S4 for site names and coordinates

coral communities (Dollar 1982; Franklin et al. 2013; Gove et al. 2015). Prior modeling studies were used to identify potential sites within areas that were predicted to have high *P. meandrina* cover (Franklin et al. 2013), and adequate *P. meandrina* density was verified in situ before each survey.

Surveys were conducted on SCUBA with one diver consistently surveying the cryptic communities and other diver(s) collecting colony-scale environmental characteristics. Focal colonies were selected haphazardly along a compass heading at least two meters from the previous colony and matching a randomly generated size class. To survey the *Pocillopora*-associated communities, a flashlight and side-to-side search pattern were used, and species identities and abundances were recorded for all associated fauna species. Visual surveys provide a conservative estimate of community abundance and species richness particularly for organisms that are either very small (< 3 mm), transparent (e.g., *H. depressa*), or associated with habitat at the base of the colony (e.g., *Trapezia* spp. juveniles, see Preston 1971). We did not observe trapeziid crabs until they approached 5 mm, a size at which they spend more time out on the colony’s branches (Preston 1971) and were consistently identifiable to the species level. For a few other taxa, we could not consistently see distinguishing features and, therefore, we grouped these species to higher taxonomic levels: hermit crabs, *Drupella* snails, vermetid snails, and *Spirobranchus* worms. In addition, “*Sebastapistes* spp.” was used for a set of three visually similar scorpionfish species (*S. fowleri*, *S. galactacma*, and *S. ballieui*). Despite these limitations, a previous study that used a similar technique confirmed that visual surveys were 97% accurate in identifying species and estimating abundance of the cryptofaunal community (Sin and Lee 2000).

For each focal *P. meandrina*, divers recorded the colony size, maximum inter-branch distance, percent live coral tissue, and depth (Table 1). In addition, a photograph was taken of each colony from about 1–2 m above the substratum. These photographs were used to estimate the density of *Pocillopora* spp. colonies in the area immediately surrounding each focal colony (Table 1). Using geospatial software ArcGIS 10.3 (ESRI 2014), maximum significant wave height, bathymetric rugosity (Franklin et al. 2013), and mean surface chlorophyll-*a* (NASA 2018) were estimated at the site scale (Table 1).

Statistical analysis

Community characterization To estimate whether or not the full community was surveyed, a species accumulation curve with a Chao estimate of asymptotic species richness was created (function “specpool” in *R* package *vegan*; Oksanen et al. 2017). A species rank abundance curve was plotted to visualize the balance of common to rare species

Table 1 Environmental driver variables, including the source of data, description of variable, unit, mean value, and range of values. Means and ranges are calculated for all 751 *Pocillopora* colonies surveyed, except for inter-branch distance, which was measured for only 716 colonies, and density of *Pocillopora*, which was available for 743 colonies

Variable	Source	Description	Units	Mean	Range
Colony size	Measured in situ to the nearest cm	Habitat size; colony volume transformed to a linear estimate $(H \times D_1 \times D_2)^{1/3}$ where H is colony height, D_1 is longest diameter, and D_2 is longest orthogonal diameter	cm	21	4–78
Percent live coral tissue	Visually estimated in situ to the nearest 5%	Habitat quality; percent of the <i>Pocillopora</i> colony that is covered in live coral tissue	%	86	0–100
Inter-branch distance	Measured in situ to the nearest mm	Interstitial space of the colony; maximum of five values for the distal distance between two adjacent branches with branch pairs haphazardly selected such that one pair was near the top center and the other four pairs were on the sides of the colony	mm	18.7	1.5–45
Depth	Measured in situ with 0.3 m precision	Depth at the base of the focal colony	m	9.5	0.6–31.1
Density of <i>Pocillopora</i>	Derived from a top down photograph centered on the focal colony	Count of <i>Pocillopora</i> colonies in the adjacent habitat divided by area of substrate in photograph of habitat	<i>Pocillopora</i> m ⁻²	1.0	0.1–7.4
Wave height ^a	SWAN hindcast model (v40.51, 2006) forced with 2000–2009 spectral wave data from WAVEWATCH III (v3.14, Tolman 2009)	Mean model predicted maximum significant wave height validated with comparisons to in situ data from NOAA wave buoys (Franklin et al. 2013)	m	2.6	0.8–5.2
Rugosity ^a	Derived from a synthesis of LIDAR and SONAR data (Hawaii Mapping Research Group 2011)	GIS modeled ratio between the surface area and the planimetric area of a depth grid for central grid cells and their 8 surrounding neighbor cells (Franklin et al. 2013)	ratio	1.003	1.000–1.013
Chlorophyll-a ^a	Moderate Resolution Imaging Spectroradiometer (MODIS) 2008–2015 data from the Aqua satellite (NASA 2018)	Mean near-surface chlorophyll-a concentration for January from measurements of color band ratios (spanning the 440–570 nm spectral regime) based on remotely sensed reflectance data	mg m ⁻³	1.34	0.07–4.13

^aThese factors are estimated at the site scale

and to identify reasonable cutoffs for commonly occurring species to be used in community composition analyses.

Community metrics Community abundance (i.e., total number of individual organisms on the host colony) and species richness values were calculated for each colony focusing on the subset of species that were observed on at least 1% of colonies (i.e., ≥ 8 of 751 colonies). To determine relative levels of variation in abundance and species richness over spatial scales, we ran a Poisson generalized linear mixed model (GLMM, function “glmer” in R package *lme4*; Bates et al. 2015) for each community metric, with site and region as nested random effects to account for the spatial structure of the data and an observation-scale random effect to account for overdispersion (ESM, Model 1). The conditional R^2 , an estimate of variance explained by all factors included in a mixed model (function “r.squaredGLMM” in R package *MuMIn*; Bartoń 2016), was used to estimate how much total variation in abundance and species richness was explained. Random effect

variance estimates were used to determine the relative amount of variation explained at each spatial scale.

We ran additional Poisson GLMMs for community abundance and species richness with colony and site-scale environmental factors (Table 1) as fixed effects, site and region as nested random effects, and an observation-scale random effect to account for overdispersion (ESM, Model 2). Environmental characteristics measured at the colony and site scales (Table 1) were centered and scaled, and correlation coefficients were evaluated with a threshold of ± 0.7 prior to inclusion in models (Table S1). To quantify the relative importance of each environmental factor, all possible models ($n = 256$) were run (function “dredge” in R package *MuMIn*; Bartoń 2016); for the subset of models that contained each variable, the associated model probabilities were summed (Anderson 2008). In addition, a weighted multi-model average was calculated, using models with a $\Delta AIC < 4$, to estimate the effect size of each predictor. These analyses were repeated for the

subset of species that were observed on 5% or more of colonies; the results were qualitatively similar and are not presented.

Species-specific patterns To examine shifts in community composition, we focused on the subset of species that occurred on 5% or more of surveyed colonies (i.e., ≥ 38 of 751 colonies). To evaluate patterns in species occurrence over spatial scales and environmental drivers, we ran a set of binomial GLMMs that had probability of occurrence as the response variable and random effects to account for species identity and survey structure (colony nested in site nested in region) as the base model (Table S2). Three additional model components were compared in the set of GLMMs (Table S2): (1) species-specific patterns over sites and regions, (2) effects of environmental factors averaged across all species, and (3) species-specific responses to the environmental factors (ESM, Model 3). To estimate the variation in occurrence explained by each model component, marginal and conditional R^2 values were calculated for each model (function “r.squaredGLMM” in R package *MuMIn*; Bartoń 2016).

An initial set of GLMMs was run that included all the environmental factors in Table 1. Three of these environmental factors (inter-branch distance, density of *Pocillopora* colonies, and rugosity) had nonsignificant effects and showed minimal variation in species-specific responses. These factors were dropped, and a simpler model with five environmental factors (colony size, percent live coral tissue, depth, wave energy, and chlorophyll-a) was used. Residual plots of all models were visually inspected, and no strong deviations from homoscedasticity or normality were observed.

Species co-occurrence To visualize patterns of species co-occurrence, we ran a principal coordinates analysis on colony-scale community composition for the species observed on 5% or more of colonies. In addition, patterns of species co-occurrence were directly compared (function “cooccur” in R package *cooccur*; Griffith et al. 2016) by classifying species pairs as having positive (i.e., co-occur more often than expected by chance alone), negative (i.e., co-occur less often than expected by chance alone), or random associations (i.e., co-occurrence is not different than expected by chance alone) based on the probabilistic model of species co-occurrence from Veech (2013).

Results

Community characterization For the 751 *P. meandrina* colonies surveyed across 19 sites, the average colony size was 21.1 ± 9.2 cm (mean \pm SD; Table 1). An average of 4 species and a maximum of 13 species per colony were observed. In total, 5887 individuals of 91 different species

(48 fishes and 43 invertebrates; Table S3) were observed in association with *P. meandrina* colonies ($n = 751$). A rarefaction plot indicated that additional surveys would identify more species with an estimated 115.2 ± 13.7 total species (mean \pm SE; Chao estimate) associated with *P. meandrina* (Fig. 2a). Decapods comprised 51% of all individuals and 25% of all species observed. Most species were observed at low colony-scale abundances (1–3 individuals per species per colony); two species were observed at high colony-scale abundances, the gall-forming coral crab *Utinomiella dimorpha* (mean = 7.9 individuals per colony) and the damselfish *Dascyllus albisella* (mean = 5.3 individuals per colony; Fig. 2b).

Of the 91 species found, only one-third (10 fishes and 21 invertebrates) were observed on more than 1% of colonies (Table 2, Fig. 2c), and the 19 species that were observed on $\geq 5\%$ of surveyed colonies (Table 2) accounted for 90% of all observations (4531 of 5037 specimens; Fig. 2c). The 19 most common species included nine known coral mutualists: five species of *Trapezia* crabs, one species of *Alpheus* shrimp, one species of *Harpiliopsis* shrimp and two species of pomacentrid fishes (Table 2).

Community metrics Based on a GLMM with only random effects (ESM, Model 1), variation in cryptofaunal species richness was 10.6% at the region scale, 5.2% at the site scale, and 84.2% at the colony scale (Fig. S1a). The west and north regions were estimated to have an average of ~ 5 species per colony, the south and east regions were estimated to have an average of ~ 3.5 species per colony, and the Kāneʻohe region was estimated to have an average of ~ 2.7 species per colony (Fig. S2a). Variation in community abundance was also predominately at the colony scale (95.1%), with 2.8% at the region scale and 2.1% at the site scale (Fig. S1b). The differences in abundance between regions followed a similar pattern as described for species richness (Fig. S2b).

In the GLMM for species richness with spatial random effects and environmental factors as fixed effects (ESM, Model 2), 42.8% of variation in species richness was explained by environmental factors (Table 1), 0.2% was accounted for at the region scale, 7.3% at the site scale, and 49.6% at the colony scale (Fig. S1a). For community abundance, 48.8% of variation was explained by environmental factors (Table 1), $< 0.1\%$ was accounted for at the region scale, 2.3% at the site scale, and 48.8% at the colony scale (Fig. S1b). Three environmental factors (colony size, wave energy, and chlorophyll-a) had strong model support ($> 85\%$) and average effect sizes that were significantly different from zero (Fig. 3) for both abundance and species richness. For the abundance model, percent live coral tissue and inter-branch distance had average effect sizes that were small but significantly different from zero (Fig. 3). Abundance increased by < 2 individuals per colony with an

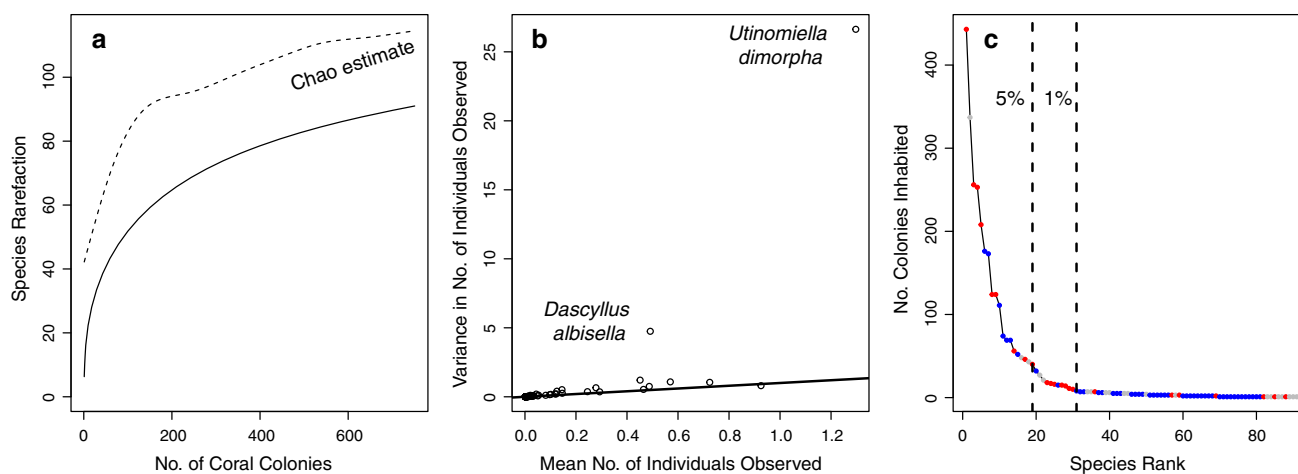


Fig. 2 **a** Species rarefaction curve for the overall pool of species found associated with 751 *P. meandrina*. The dashed line shows the mean Chao estimate for number of species, which reaches 115 species for 751 colonies. **b** For each of 91 species, the mean number of individuals observed and the variance in the number of individuals

increase in percent live tissue from 0 to 100% (Fig. S3m), and decreased by < 3 individuals per colony with increasing inter-branch distance from 1.5 to 45 mm (Fig. S3o). Across the range of colony sizes surveyed (4–78 cm diameter, Table 1), the model predicted an increase from 0 to 45 species and from 0 to 400 individuals per colony with increasing colony size (Fig. S3a and S3i). With increasing maximum wave energy over the range surveyed (Table 1), the model predicted an increase from 2 to 5 species and from 3 to 12 individuals per colony (Fig. S3b and S3j). The model predicted a decrease from 4 to 2.5 species and from 6 to 3 individuals per colony with increasing surface chlorophyll-a levels (Fig. S3c and S3k) over the range surveyed (Table 1). For all other environmental factors included (i.e., depth, percent live coral tissue, density of *Pocillopora*, inter-branch distance, and rugosity), species richness was predicted to change by < 1 species per colony and community abundance was predicted to change by 3 or less individuals per colony (Fig. S3d to S3h and S3l to S3p) over the ranges surveyed (Table 1).

Species-specific variation Of the 91 species observed, five species were observed at all sites (*Alpheus lottini*, *Ophiocoma pica*, *Harpiliopsis depressa*, *Trapezia intermedia*, *T. tigrina*), and 31 species were observed at only one site (Table S3). Relative to other survey sites, Kāneʻohe Bay sites had high proportions of damselfishes (i.e., *Plectroglyphidodon johnstonianus*, *D. albisella*) and low proportions of predatory fishes (e.g., *Paracirrhites arcatus*, *Caracanthus typicus*) (Fig. 4). Three of the five species of *Trapezia* crabs (i.e., *T. digitalis*, *T. bidentata*, *T. flavopunctata*) had occurrence rates approaching zero at all three Kāneʻohe Bay sites, Lanikai, Waikiki, Ewa Beach,

observed across colonies display a relative index of aggregation, where species that fall above the 1:1 line have clumped distributions. **c** Species rank abundance plot with decapods species shown in red, other invertebrates in gray, and fishes in blue. The dashed lines show cutoffs for species observed on 5 and 1% of colonies

and Yokohama (Fig. 4; site names in Table S4). The relatively unique community composition for colonies from sites in Kāneʻohe Bay was documented by a canonical analysis of principal coordinates constrained by survey site (Fig. S4).

Comparisons across binomial GLMMs run with different model components showed that species-specific patterns explained most of the variation in occurrences (Table S2; Fig. S5a and S5b). The base model including random effects for species identity and the hierarchical structure of the surveys explained 30.4% of the variation in species occurrences (Table S2: Base model). Most of this variation, 21.2%, was due to species identity, i.e., some species were more common than others overall. The remaining variation, 9.2%, was allocated to spatial scales reflecting that some locations (i.e., colonies, sites, or regions) had higher probabilities of occurrence for all species. Including species-specific patterns at the site and region scales nearly doubled the explained variation to 57.1% (Table S2: Compositional variation over spatial scales). Most of this variation was at the colony scale, followed by site scale, and then region scale (Fig. S5a).

The full GLMM, which included average environmental effects, species-specific variation over spatial scales, and species-specific variation over environmental factors, explained 68.2% of the variance in occurrences (Table S2: Full model). This was 8% more variation than any other model, supporting the complementary explanatory power of species-specific patterns over environmental gradients and species-specific patterns over spatial scales. The average response to environmental factors across all species explained 13.2% of variance in occurrences (Table S2). Three environmental factors had effect sizes

Table 2 Family, genus, species, species codes, % of regions, % of sites, % of colonies, average number of individuals observed on a colony, depth range, and % live coral tissue range listed for fish ($n = 10$) and invertebrate ($n = 21$) species observed on $\geq 1\%$ of surveyed *P. meandrina* ($n = 751$ colonies). Species are listed in descending order of % of colonies inhabited, with species observed

on $\geq 5\%$ of colonies listed in bold. Known coral mutualist species are noted with an *. *Sebastapistes* spp. includes observations of *S. fowleri*, *S. galactacma*, and *S. ballieui* (visually similar species that have been previously observed in Hawai'i in associated with *Pocillopora* corals)

Family	Genus	Species	Spp. code	% of regions	% of sites	% of colonies	Avg. per colony	Depth (m)	% Live coral tissue
*Trapeziidae	<i>Trapezia</i>	<i>intermedia</i>	TRIN	100	100	59.0	1.6	0.6–31.1	20–100
Ophiocomidae	<i>Ophiocoma</i>	<i>pica</i>	OPPI	100	100	45.0	1.6	1.5–30.5	0–100
*Palaemonidae	<i>Harpiliopsis</i>	<i>depressa</i>	HADE	100	100	34.1	1.7	0.9–24.7	0–100
*Alpheidae	<i>Alpheus</i>	<i>lottini</i>	ALLO	100	100	33.7	1.4	0.6–26.8	30–100
*Trapeziidae	<i>Trapezia</i>	<i>tigrina</i>	TRTI	100	100	27.7	1.8	1.5–26.2	30–100
Scorpaenidae	<i>Sebastapistes</i>	spp.	SESP	100	89	23.0	1.9	2.4–30.5	30–100
Cirrhitidae	<i>Paracirrhites</i>	<i>arcatus</i>	PAAR	80	79	23.0	1.3	3.4–31.1	0–100
*Trapeziidae	<i>Trapezia</i>	<i>digitalis</i>	TRDI	100	79	16.5	1.5	2.1–19.5	45–100
Cryptochiridae	<i>Utinomiella</i>	<i>dimorpha</i>	UTDI	80	42	16.5	7.9	3.0–18.3	30–100
Scorpaenidae	<i>Sebastapistes</i>	<i>coniorta</i>	SECO	100	89	15	1.9	2.1–19.5	30–100
Caracanthidae	<i>Caracanthus</i>	<i>typicus</i>	CATY	80	79	9.9	1.5	2.4–16.8	50–100
*Pomacentridae	<i>Plectroglyphididon</i>	<i>johnstonianus</i>	PLJO	100	74	9.2	1.3	2.1–26.2	0–100
*Pomacentridae	<i>Dascyllus</i>	<i>albisella</i>	DAAL	100	58	9.2	5.3	1.5–30.2	0–100
*Trapeziidae	<i>Trapezia</i>	<i>bidentata</i>	TRBI	100	58	7.5	1.6	2.1–18.0	45–100
Labridae	<i>Thalassoma</i>	<i>duperrey</i>	THDU	100	53	6.9	1.4	2.1–20.1	5–100
Ophiocomidae	<i>Ophiocoma</i>	<i>erinaceus</i>	OPER	100	79	6.4	1.3	2.7–20.4	0–100
*Trapeziidae	<i>Trapezia</i>	<i>flavopunctata</i>	TRFL	100	68	6.1	1.7	2.1–16.8	50–100
Sabellidae	<i>Sabellastarte</i>	<i>spectabilis</i>	SASP	60	32	5.7	2.5	2.1–18.3	0–100
Paguroidea	unidentified		DIOG	60	53	5.3	2.4	2.4–19.8	0–100
Cirrhitidae	<i>Amblycirrhitus</i>	<i>bimacula</i>	AMBI	100	42	4.3	1.2	3.0–14.3	0–100
Echinometridae	<i>Echinometra</i>	<i>mathaei</i>	ECMA	80	53	3.6	1.5	4.3–29.6	0–100
Muricidae	<i>Quoyula</i>	<i>monodonta</i>	QUMO	80	42	2.8	1.6	2.7–17.7	50–100
Domeciidae	<i>Domecia</i>	<i>hispida</i>	DOHI	80	47	2.4	1.4	3.4–19.8	40–100
Xanthidae	<i>Pseudoliomera</i>	<i>speciosa</i>	PSSP	100	68	2.3	1.4	2.1–14.3	75–100
Grapsidae	<i>Percnon</i>	<i>planissimum</i>	PEPL	100	37	2.1	1.3	2.1–19.2	0–98
Labridae	<i>Pseudocheilinus</i>	<i>tetrataenia</i>	PSTE	80	47	2.0	1.2	4.9–20.3	45–100
Hippolytidae	<i>Saron</i>	<i>marmoratus</i>	SAMA	80	37	2.0	1.6	2.1–15.5	50–100
Portunidae	<i>Charybdis</i>	<i>hawaiensis</i>	CHHA	80	47	1.9	1.1	2.1–25.3	50–98
Palaemonidae	<i>Palaemon</i>	<i>pacificus</i>	PAPA	60	32	1.5	1.1	2.1–13.4	20–100
Hippolytidae	<i>Saron</i>	<i>neglectus</i>	SANE	60	26	1.3	2.1	3.7–10.4	0–80
Cirrhitidae	<i>Cirrhitops</i>	<i>fasciatus</i>	CIFA	100	32	1.1	1	2.1–16.8	55–98

that were significantly different from zero: colony size (0.99 ± 0.12 SE; Fig. 5a), wave height (0.79 ± 0.24 SE; Fig. 5b), and chlorophyll-a (-0.71 ± 0.30 SE; Fig. 5c). Including species-specific responses for each of five environmental factors significantly improved the model fit (Chi-squared likelihood ratio tests, function “anova” in *R* base *stats* package) supporting significant variation in community composition across these environmental gradients. The standard deviation for species-specific responses to environmental factors was greatest for chlorophyll-a (0.99), followed by depth (0.85), wave height (0.64),

colony size (0.48), and percent live coral tissue (0.37) (Fig. 5f). Species-specific responses to depth and percent live coral tissue depicted distinct shifts in the community over these gradients with some species having higher probability of occurrence at low values and other species having higher probability of occurrence at high values (Fig. 5d, e). For example, guard crab *T. digitalis*, coral gall crab *U. dimorpha*, flattened coral shrimp *H. depressa*, and coral croucher *C. typicus* had higher probabilities of occurrence on shallower corals, while guard crab *T. tigrina*, brittlestar *O. pica*, damselfish *D. albisella*, and hawkfish *P.*

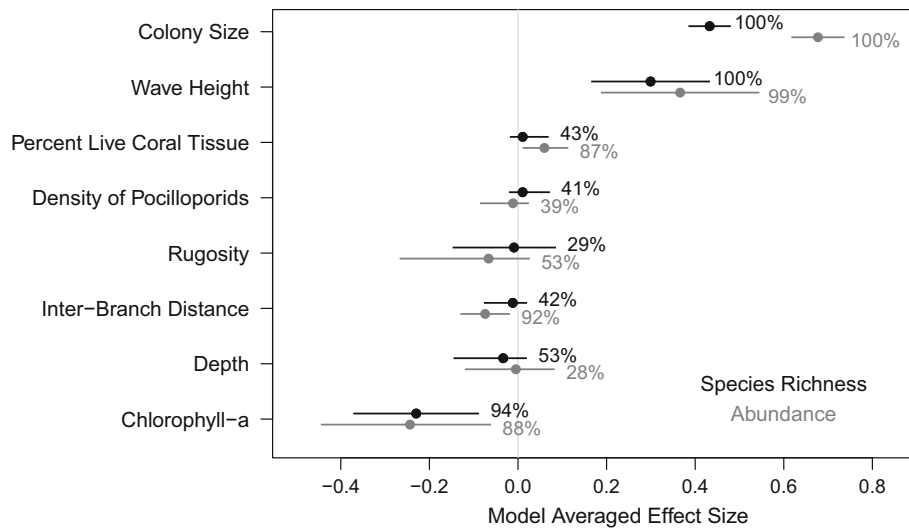
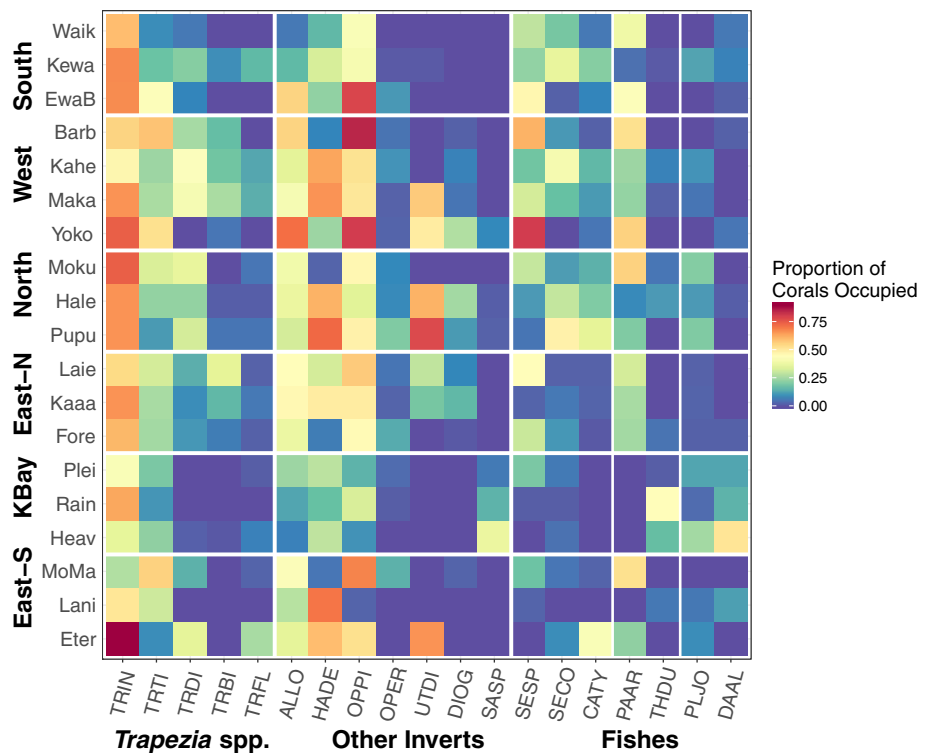


Fig. 3 Multi-model-averaged parameter coefficients with 95% confidence intervals from the subset of models with delta Δ AIC < 4 for each community metric, i.e., species richness (black, 31 models) and community abundance (gray, 10 models). For the effect of a variable to be significantly different from zero, the error bars cannot overlap zero (thin vertical line), e.g., chlorophyll-a had a significant negative

effect size for both abundance and species richness; however, percent live coral tissue only had a significant effect size in relation to abundance. As a metric of relative model support, the summed model probabilities for the subset of models containing each parameter are listed

Fig. 4 Heat map showing the proportion of colonies at a site (see Table S4 for site details) inhabited by each species (see Table 2 for species names), calculated for the species observed on 5% or more of colonies ($n = 751$). White horizontal lines divide the sites into regions. White vertical lines break the species into functional groups with fishes being further divided into: predatory fishes that live tucked in between the branches (left), other predatory fishes (middle), and planktivorous fishes (right)



arcatus had higher probabilities of occurrence on deeper corals, and guard crab *T. intermedia* and snapping shrimp *A. lottini* had no change in probability of occurrence over the depth range surveyed (Fig. 5e).

Species co-occurrence The guard crab, *T. intermedia*, had the highest probability of occurrence and showed a

distinct negative response to colony size (Fig. 5a). *T. intermedia* was also separated from the other species (including four other *Trapeziid* crab species) in a principal coordinates analysis of the community (Fig. 6a). A co-occurrence analysis of 171 pairs of the commonly observed species found that 55.6% were non-random: 65 species

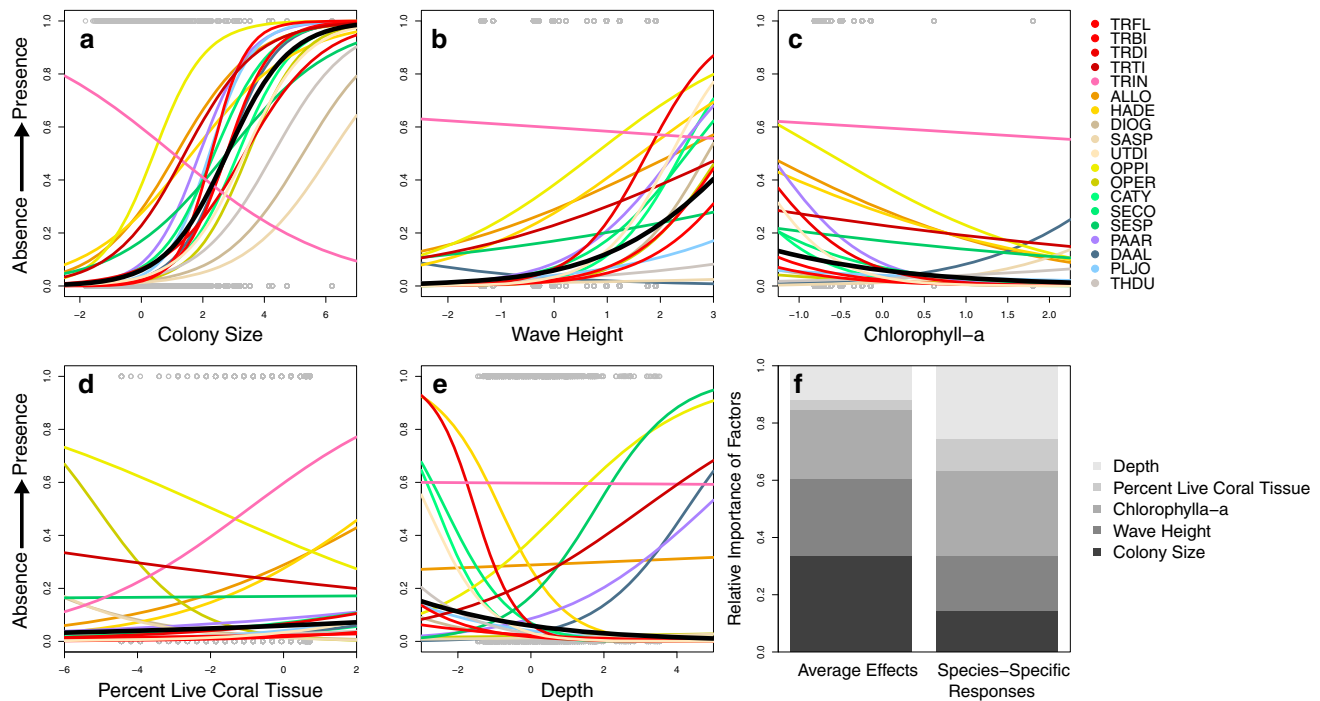


Fig. 5 For a GLMM with species-specific responses to environmental gradients, sites, and regions (ESM, Model 3), the species-specific probability of occurrence trends over environmental factors: **a** colony size, **b** wave height, **c** chlorophyll-a, **d** percent live coral tissue, and **e** depth are plotted with the average trend (across species) in black

and data points in gray. See Table 2 for species codes. **f** The relative importance of each environmental factor based on average effects (fixed effect estimates) and species-specific responses (standard deviation of random effects)

pairs occurred more frequently than expected (i.e., positive co-occurrences), and 30 species pairs occurred less frequently than expected (i.e., negative co-occurrences). *T. intermedia* was in 26.7% of all negative co-occurrences (Fig. 6b), including negative co-occurrences with three of the other *Trapezia* species (*T. flavopunctata*, *T. tigrina*, and *T. bidentata*). Of the remaining *Trapezia* pairs, only one other was a negative co-occurrence (*T. tigrina*–*T. digitalis*), three were positive (*T. digitalis*–*T. flavopunctata*, *T. flavopunctata*–*T. bidentata*, and *T. digitalis*–*T. bidentata*), and three were random (*T. intermedia*–*T. digitalis*, *T. flavopunctata*–*T. tigrina*, and *T. tigrina*–*T. bidentata*).

Discussion

Our results documented spatial scales of variation in cryptofaunal communities associated with *P. meandrina* colonies and the importance of environmental factors in structuring community composition. Patterns in community abundance and species richness were similar with the largest amount of spatial variation observed at the colony scale (Fig. S1) and strong correlations for both community metrics with colony size, maximum wave height, and surface chlorophyll-a (Fig. 3). Species-specific responses to environmental factors revealed additional shifts in the

community across a depth gradient and a range of percent live coral tissue, and presented a unique occurrence pattern for the most commonly observed trapeziid crab.

High variation in the cryptofaunal community at the colony scale indicates the importance of colony quality factors in structuring the associated community. Colony size, one metric of colony quality, had the largest correlation with abundance and species richness (Fig. 3), with shifts in predicted values an order of magnitude larger than across the range of any other factor (Fig. S3). The increase in species richness with increasing colony size is consistent with species–area relationships (Arrhenius 1921) and previous studies of *Pocillopora*-associated communities (Stella et al. 2010; Holbrook et al. 2011; Head et al. 2015; Britayev et al. 2017). Our results suggest that there are additional, unmeasured colony quality factors because colony-scale variation remained relatively high in the model that included colony size, percent live coral tissue, and inter-branch distance (Fig. S1). Colony quality factors to which species may be responding could include colony age, symbiont clade, or complexity of interstitial microhabitats.

At the regional scale, cryptofaunal communities had higher abundance and species richness along the northern and western shorelines, average values along the southern and eastern shorelines, and relatively low values within

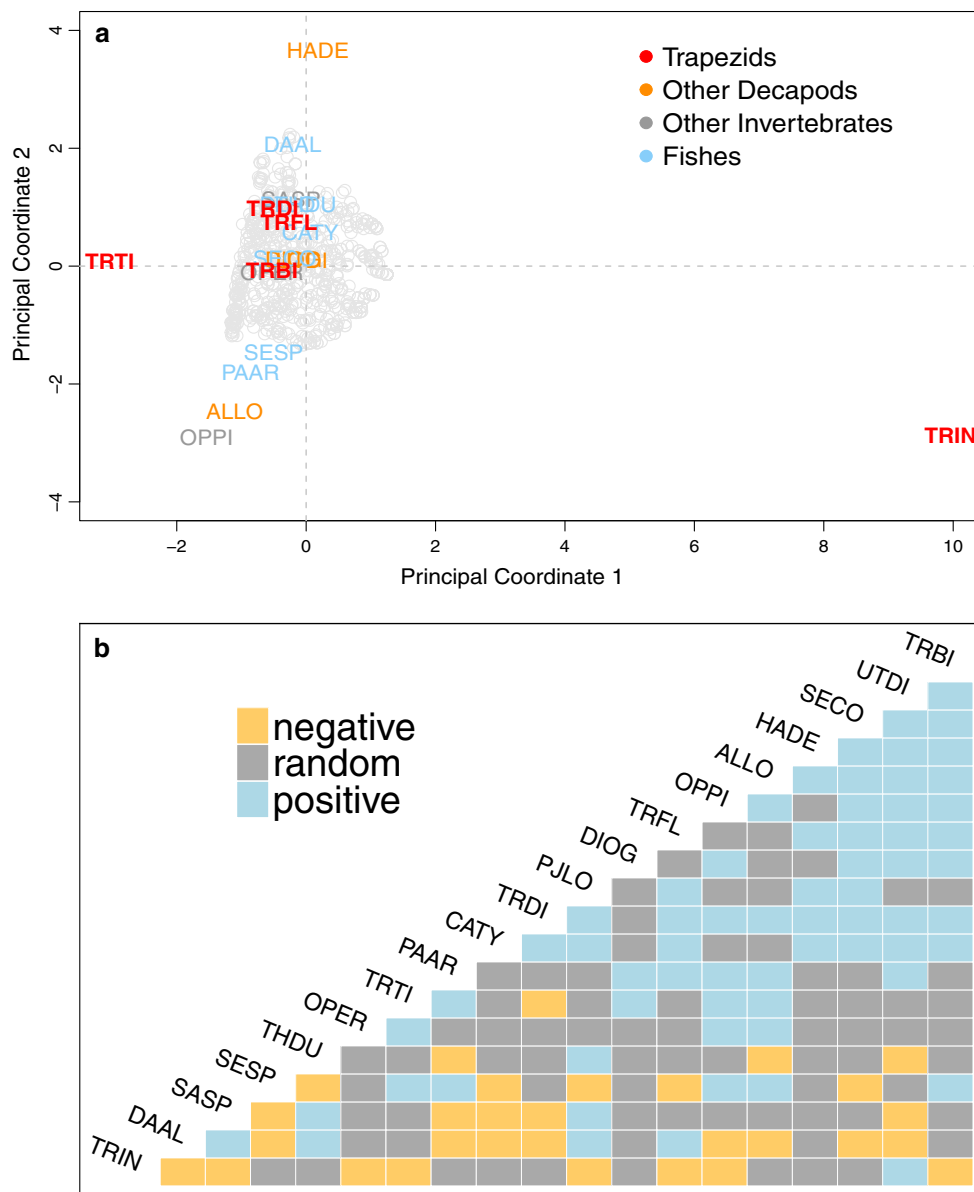


Fig. 6 a The first two coordinates of a PCoA on community composition data for species observed on 5% or more of colonies (19 species, 751 colonies). *Trapezia intermedia* (TRIN) is the only species in the lower right quadrant. **b** Co-occurrence analysis depicts

TRIN as being in 27% of negative co-occurrences (i.e., co-occurrences that were less frequent than expected given a random distribution of species). See Table 2 for species codes

Kāneʻohe Bay (Fig. S2). Environmental factors including wave height and surface chlorophyll explained most of this regional scale variation (Fig. S1). Community abundance and species richness increased with increasing wave height (Fig. S3b and S3j). This trend aligns with observed regional weather patterns, such as strong winter storms create exceptionally large waves along the NW coasts, the SE coasts have a more consistent level of wave energy with occasional storm driven peaks that are generally smaller than the NW storms, and Kāneʻohe Bay is the most sheltered region surveyed (Fletcher et al. 2008). Previous work has quantified the importance of wave energy and water

movement for structuring benthic cover (Franklin et al. 2013; Gove et al. 2015), and the community composition of both non-cryptic and cryptic reef fishes (Nunes et al. 2013; Depczynski and Bellwood 2005). Maximum wave height can be a metric of disturbance for coral reef communities. The increase in species richness observed in this study from colonies at sheltered sites to colonies at sites with large seasonal waves corresponds with the expected shift in diversity among habitats with small disturbances to habitats with an intermediate level of disturbances (Connell 1978).

Cryptofaunal community abundance and species richness decreased with increasing surface chlorophyll-a, a

measure of productivity (Fig. S3c and S3k). Previous work has found that there is often a correlation between diversity and productivity, but the direction of this relationship shows high variation over different communities and spatial scales (Cornell and Karlson 2000; Mittelbach et al. 2001; Chase and Leibold 2002). For this study, surface chlorophyll-a was estimated from satellite data at the site scale. At this resolution, chlorophyll-a had approximately the same value for all three sites in Kāneʻohe Bay, 4.13 mg m^{-3} , almost double the next highest chlorophyll-a estimate (Table 1). Thus, in the context of this study, chlorophyll-a may represent the distinctiveness of Kāneʻohe Bay, an estuary which receives high freshwater and sediment input from various streams and frequent orographic rainfall on the adjacent mountain range (Jokiel 1991), rather than variability in productivity across all sites. The relatively low abundance and species richness observed in Kāneʻohe Bay may reflect that habitat characteristics within the bay are not preferred habitat for some species, that restricted exchange into Kāneʻohe Bay limits dispersal and prevents less common species from becoming established, or a combination of both.

Percent live coral tissue, depth, inter-branch distance, density of *Pocillopora* colonies, and rugosity did not have a significant effect on cryptofaunal species richness (Fig. 3). While previous work has demonstrated species-specific preferences based on inter-branch distances, e.g., *P. arcatus* prefers larger distances (Kane et al. 2009), our study did not support a correlation between species richness and inter-branch distance or species-specific patterns across colonies with differing inter-branch distances. We did observe higher abundances on colonies with smaller inter-branch distances (Fig. 3). This pattern has also been recorded for communities associated with acroporid corals, with greater abundances of cryptofauna found on tightly branched *Acropora* spp. than on arborescent *Acropora* spp. (Vytopil and Willis 2001). Habitat complexity is known to be positively correlated with abundance and diversity of non-cryptic reef-associated fishes (Gratwicke and Speight 2005; Darling et al. 2017), yet our results showed that these factors (i.e., density of *Pocillopora* colonies and rugosity) were not significantly correlated with species richness nor were there substantial species-specific patterns relative to these factors. These results suggest either that these factors do not matter or that the scales at which these factors were measured (Table 1) were not appropriate for the *P. meandrina*-associated community. Species richness was not related to depth or percent live coral tissue, although species-specific trends were observed across these two factors, suggesting that species may filter in and out over these environmental gradients, thereby concealing a significant pattern when considering species richness. This outcome elucidates the benefit of considering species-

specific responses to identify shifts in communities over environmental gradients that are not associated with significant changes in community-scale metrics.

Similar to variation in community metrics, most species-specific variation in community composition was at the colony scale (Fig. S5). Environmental factors and species-specific responses to these factors accounted for some of the variation in species-specific occurrences at the colony and region scales, yet site-scale variation remained relatively high (Fig. S5b) suggesting that species responded to unmeasured factors at the site scale. Variation at the site scale could be due to constraints in local dispersal or habitat quality of the area surrounding the host corals, e.g., coral cover or adjacent habitat complexity (with a finer resolution than was considered here). The full model accounted for 68% of the variation in species-specific probability of occurrence, with the remaining 32% of variation in occurrence due to species-specific patterns at the colony scale (ESM, Model 3). Species-specific patterns at the colony scale that were not explained by the environmental factors or associated with colony identity are likely the result of species-specific responses to unmeasured colony-scale metrics of habitat quality. This could include order-of-arrival community assembly dynamics, such as priority effects (Shulman et al. 1983; Almany 2003) with species avoiding or preferring colonies based on community composition, or the complexity of inter-branch microhabitats.

The environmental drivers emphasized as strong correlates for species richness, i.e., colony size, wave height, and surface chlorophyll-a, had limited variation in species-specific responses. Almost all commonly observed species (with the exception of guard crab *T. intermedia*) had a higher probability of being observed on larger colonies than smaller ones (Fig. 5a). Most of the variation in species-specific responses to colony size was due to differences in the smallest size on which each species had a high probability of occurrence. For example, the hawkfish *P. arcatus*, damselfish *P. johnstonianus*, and guard crab *T. flavopunctata* were observed with low probability until the colony was relatively large (Fig. 5a). Species-specific responses to wave height were fairly consistent, with the average trend of increased probability of occurrence with increasing wave height. Two exceptions, the damselfish *D. albisella* and the guard crab *T. intermedia*, had a modest decrease in their probability of occurrence with increasing wave height (Fig. 5b). For most species, the probability of occurrence decreased with increasing surface chlorophyll-a, although the opposite trend was observed for a few species, including the damselfish *D. albisella* and the wrasse *T. duperrey* (Fig. 5c).

There was no change in overall probability of occurrence with varying percent live coral tissue, yet distinct

species-specific patterns were observed (Fig. 5d). Some known coral-obligate species, including guard crab *T. intermedia*, snapping shrimp *A. lottini*, and flattened coral shrimp *H. depressa*, had higher probabilities of occurrence with higher percent live host coral tissue (Fig. 5d). In contrast, two species of brittle stars (*Ophiocoma pica* and *O. erinaceus*) were observed more often on colonies with lower proportions of live tissue (Fig. 5d). Previous studies have found different responses to the percent of live coral across different functional groups. For example, a higher proportion of live coral is associated with a higher diversity of reef fishes (Rasher et al. 2013) and a lower diversity of cryptic motile invertebrates (Coles 1980; Enochs and Hockensmith 2008; Enochs and Manzello 2012; Leray et al. 2012). Prior studies have shown that the invertebrate communities associated with dead corals are mainly, although not exclusively (Head et al. 2015), composed of facultative species with higher diversity per colony and higher variability across corals (Coles 1980). Our study confirms a shift from obligate to facultative species as the percent of live coral tissue declines, including a small increase in community abundance but no overall change in species richness.

Although depth had no overall effect on the probability of occurrence, there were strong species-specific responses from species across the depth gradient (Fig. 5e). These species-specific patterns are likely due to variation in recruitment and survival rates for each species across the depth gradient, which have been shown to structure the depth range of a *Pocillopora*-associated goby, *Paragobiodon xanthosoma* (Smallhorn-West et al. 2017). Shifts in species composition over depth gradients have been previously shown for both cryptic reef fish communities (Depczynski and Bellwood 2005) and non-cryptic reef fishes assemblages (Nunes et al. 2013; Jankowski et al. 2015; Darling et al. 2017).

In addition to revealing shifts in community composition across environmental factors, our species-specific GLMM depicted an intriguingly unique response to coral colony size for the most commonly observed species, *T. intermedia*. Unlike all other common species, the probability of occurrence of *T. intermedia* was higher for smaller colonies than for larger colonies (Fig. 5a) despite being observed across a broad range of colony sizes (7–65 cm). Of the five trapeziid species observed, *T. intermedia* was of a similar body size to all other species except *T. flavo-unctata* which was distinctively larger than the other species, suggesting that the unique relationship between *T. intermedia* occurrence and colony size was not driven by body size differences. For the other environmental factors, the probability of *T. intermedia* occurrence was either largely unaffected (i.e., wave height, surface chlorophyll-a, and depth) or followed a similar pattern to other obligate

species (i.e., percent live coral tissue). The observation that *T. intermedia* had a high probability of occurring on smaller colonies suggests that *T. intermedia* is one of the first species to colonize *P. meandrina*. Previous studies focused on the decapod communities associated with *P. meandrina* also noted that *T. intermedia* was the predominant trapeziid on small colonies (Barry 1965; Preston 1971; Huber and Coles 1986). A PCoA further supported unique characteristics of the distribution of *T. intermedia*, which was separated from other species in multidimensional space (Fig. 6a). A co-occurrence analysis found that *T. intermedia* occurred less often than expected by chance with three of the four other *Trapezia* crab species (Fig. 6b) likely due to competitive behavior. The patterns we observed for *T. intermedia* are consistent with the patterns expected for a species that is a good colonizer (first to arrive to small colonies) but a poor competitor (not often observed with congeneric species).

While our analyses focused on the most common species, our surveys also provided information regarding the rare cryptofaunal species inhabiting *P. meandrina*. Our results were consistent with the hypothesis that most of the species richness in reef cryptofauna is due to rare species (Austin et al. 1980; Plaisance et al. 2009; Stella et al. 2010; Plaisance et al. 2011), with 60 of 91 species observed on < 1% of colonies, and 22 species observed on only one colony (Fig. 2c). Interestingly, some of the species that were rare in our surveys of *P. meandrina* colonies are relatively common in the larger reef ecosystem (e.g., the surgeonfish *Acanthurus triostegus* and the urchin *Echinothrix diadema*), suggesting that these species are transient in the context of *P. meandrina* communities, temporarily associated with the colony despite it not being their primary habitat (Sgarbi and Melo 2017).

Previous work has highlighted the importance of certain *Pocillopora*-associated species and their species interactions for the host coral's health and survival as well as the structure of the cryptic community. This study provides context for this existing body of literature by characterizing naturally occurring patterns in the community relative to environmental factors and partitioning variation in the community across spatial scales. While some trends in community composition emerged at site and regional scales, the highest level of variation was at the colony scale. Our results emphasize the importance of colony size, wave height, and surface chlorophyll-a for driving the composition of cryptofaunal communities associated with *P. meandrina*. In addition, our study documented a shift in community composition over both depth and percent live coral tissue largely driven by species-specific patterns. Unique species-specific patterns for *T. intermedia* were identified, and we recommend further examination of the role of this species in community assembly processes. This

study strengthens our understanding of how cryptofaunal reef communities, where most of the diversity on coral reefs is hidden, vary across environmental gradients.

Acknowledgements C Counsell, M Donahue, E Franklin, and M Hixon developed the design for this study. M Donahue, the Hawai'i Institute of Marine Biology, and the UH Department of Biology provided equipment for field surveys. J Jones and J Kuwabara supported this project with boat and dive safety support. Dive assistants included J Caldwell, IR Caldwell, R Coleman, C Couch, M Donahue, E Franklin, N Gutlay, M Hixon, C Jerolmon, K Lubarsky, A Moran, E Nalley, M Ross, N Silbiger, T Wester, and J Zill. C Counsell organized and conducted the surveys, managed the database, and with M Donahue's guidance analyzed the data and wrote the manuscript. K Edwards assisted in developing and interpreting the species-specific community analysis. All authors reviewed the manuscript. C Counsell's participation in this study was supported by a National Science Foundation Graduate Research Fellowship (Grant No. 2012103208). Additional funding support was provided by M Donahue and the Castle Foundation (Grant 3846, M Hixon PI). This is HIMB contribution no. 1731 and SOEST contribution no. 10409.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- Abele LG (1976) Comparative species richness in fluctuating and constant environments: coral-associated decapod crustaceans. *Science* 192:461–463
- Abele LG, Patton WK (1976) The size of coral heads and the community biology of associated decapod crustaceans. *J Biogeogr* 3:35–47
- Anderson DR (2008) Model based inference in the life sciences: A primer on evidence. Fort Collins, Springer Science
- Almany GR (2003) Priority effects in coral reef fish communities. *Ecology* 84:1920–1935
- Arrhenius O (1921) Species and area. *J Ecol* 9:95–99
- Austin AD, Austin SA, Sale PF (1980) Community structure of the fauna associated with the coral *Pocillopora damicornis* (L.) on the Great Barrier Reef. *Aust J Mar Freshw Res* 31:163–174
- Barry CK (1965) Ecological study of the decapod crustaceans commensal with the branching coral *Pocillopora meandrina* var. *Nobilis* Verrill. MS thesis, Univ Hawaii
- Bartoń K (2016) MuMIn: Multi-model inference. R Packag version 1156
- Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48
- Black R, Prince J (1983) Fauna associated with the coral *Pocillopora damicornis* at the southern limit of its distribution in Western Australia. *J Biogeogr* 10:135–152
- Britayev TA, Spiridonov VA, Deart YV, El-Sherbiny M (2017) Biodiversity of the community associated with *Pocillopora verrucosa* (Scleractinia: Pocilloporidae) in the Red Sea. *Mar Biodivers* 47:1093–1109
- Chase JM, Leibold MA (2002) Spatial scale dictates the productivity-biodiversity relationship. *Nature* 416:427–430
- Chase TJ, Pratchett MS, Walker SPW, Hoogenboom MO (2014) Small-scale environmental variation influences whether coral-dwelling fish promote or impede coral growth. *Oecologia* 176:1009–1022
- Coles SL (1980) Species diversity of decapods associated with living and dead reef coral *Pocillopora meandrina*. *Mar Ecol Prog Ser* 2:281–291
- Connell JH (1978) Diversity in tropical rain forests and coral reefs. *Science* 199:1302–1310
- Cornell HV, Karlson RH (2000) Coral species richness: ecological versus biogeographical influences. *Coral Reefs* 19:37–49
- Darling ES, Graham NAJ, Januchowski-Hartley FA, Nash KL, Pratchett MS, Wilson SK (2017) Relationships between structural complexity, coral traits, and reef fish assemblages. *Coral Reefs* 36:1–15
- Depczynski M, Bellwood D (2005) Wave energy and spatial variability in community structure of small cryptic coral reef fishes. *Mar Ecol Prog Ser* 303:283–293
- Dollar SJ (1982) Wave stress and coral community structure in Hawaii. *Coral Reefs* 1:71–81
- Enochs IC (2012) Motile cryptofauna associated with live and dead coral substrates: Implications for coral mortality and framework erosion. *Mar Biol* 159:709–722
- Enochs IC, Hockensmith G (2008) Effects of coral mortality on the community composition of cryptic metazoans associated with *Pocillopora damicornis*. *Proc 11th Int Coral Reef Symp* 26:1368–1372
- Enochs IC, Manzello DP (2012) Responses of cryptofaunal species richness and trophic potential to coral reef habitat degradation. *Diversity* 4:94–104
- Fletcher CH, Bochicchio C, Conger CL, Engels MS, Feirstein EJ, Frazer N, Glenn CR, Grigg RW, Grossman EE, Harney JN, Isoun E, Murray-Wallace CV, Rooney JJ, Rubin KH, Sherman CE, Vitousek S (2008) Geology of Hawaii reefs. In: Riegl B, Dodge R (eds) *Coral reefs of the USA*. Springer, New York, NY, pp 435–487
- Franklin EC, Jokiel PL, Donahue MJ (2013) Predictive modeling of coral distribution and abundance in the Hawaiian Islands. *Mar Ecol Prog Ser* 481:121–132
- Gochfeld DJ (2009) Territorial damselfishes facilitate survival of corals by providing an associational defense against predators. *Mar Ecol Prog Ser* 398:137–148
- Goldshmid R, Holzman R, Weihs D, Genin A (2004) Aeration of corals by sleep-swimming fish. *Limnol Oceanogr* 49:1832–1839
- Gotelli NJ, Abele LG (1983) Community patterns of coral-associated decapods. *Mar Ecol Prog Ser* 13:131–139
- Gove JM, Williams GJ, McManus MA, Clark SJ, Ehes JS, Wedding LM (2015) Coral reef benthic regimes exhibit non-linear threshold responses to natural physical drivers. *Mar Ecol Prog Ser* 522:33–48
- Gratwicke B, Speight MR (2005) The relationship between fish species richness, abundance and habitat complexity in a range of shallow tropical marine habitats. *J Fish Biol* 66:650–667
- Griffith DM, Veech JA, Marsh CJ (2016) cooccur: Probabilistic species co-occurrence analysis in R. *J Stat Softw* 69:1–17
- Head CEI, Bonsall MB, Koldewey H, Pratchett MS, Speight M, Rogers AD (2015) High prevalence of obligate coral-dwelling decapods on dead corals in the Chagos Archipelago, central Indian Ocean. *Coral Reefs* 34:905–915
- Holbrook SJ, Brooks AJ, Schmitt RJ, Stewart HL (2008) Effects of sheltering fish on growth of their host corals. *Mar Biol* 155:521–530
- Holbrook SJ, Schmitt RJ, Brooks AJ (2011) Indirect effects of species interactions on habitat provisioning. *Oecologia* 166:739–749
- Huber ME, Coles SL (1986) Resource utilization and competition among the five Hawaiian species of Trapezia Crustacea, Brachyura. *Mar Ecol Prog Ser* 1:21–31

- Jankowski MW, Gardiner NR, Jones GP (2015) Depth and reef profile: Effects on the distribution and abundance of coral reef fishes. *Environ Biol Fishes* 98:1373–1386
- Jokiel PL (1991) *Jokiel's Illustrated Scientific Guide To Kane'Ohe Bay, Oahu*
- Kane CN, Brooks AJ, Holbrook SJ, Schmitt RJ (2009) The role of microhabitat preference and social organization in determining the spatial distribution of a coral reef fish. *Environ Biol Fishes* 84:1–10
- Klump DW, McKinnon AD, Mundy CN (1988) Motile cryptofauna of a coral reef: abundance, distribution and trophic potential. *Mar Ecol Prog Ser* 45:95–108
- Leray M, Béraud M, Anker A, Chancerelle Y, Mills SC (2012) *Acanthaster planci* outbreak: Decline in coral health, coral size structure modification and consequences for obligate decapod assemblages. *PLoS One* 7:1–10
- Leray M, Meyer CP, Mills SC (2015) Metabarcoding dietary analysis of coral dwelling predatory fish demonstrates the minor contribution of coral mutualists to their highly partitioned, generalist diet. *PeerJ* 3:e1047
- López-Pérez A, Granja-Fernández R, Benítez-Villalobos F, Jiménez-Antonio O (2017) *Pocillopora damicornis*-associated echinoderm fauna: Richness and community structure across the southern Mexican Pacific. *Mar Biodivers* 47:481–490
- McKeon CS, Moore JM (2014) Species and size diversity in protective services offered by coral guard-crabs. *PeerJ* 2:e574
- McKeon CS, Stier AC, McLroy SE, Bolker BM (2012) Multiple defender effects: Synergistic coral defense by mutualist crustaceans. *Oecologia* 169:1095–1103
- Mittelbach GG, Steiner CF, Scheiner SM, Gross KL, Reynolds HL, Waide RB, Willig MR, Dodson SI, Gough L (2001) What is the observed relationship between species richness and productivity? *Ecology* 82:2381–2396
- de Nunes JACC, Sampaio CLS, Barros F (2013) How wave exposure, group size and habitat complexity influence foraging and population densities in fishes of the genus *Halichoeres* (Perciformes: Labridae) on tropical rocky shores. *Mar Biol* 160:2383–2394
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlenn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2017) *vegan: Community Ecology Package*. R Packag version 2.4-2
- Plaisance L, Caley MJ, Brainard RE, Knowlton N (2011) The diversity of coral reefs: What are we missing? *PLoS One* 6:1–7
- Plaisance L, Knowlton AN, Paulay AG, Meyer AC (2009) Reef-associated crustacean fauna: biodiversity estimates using semi-quantitative sampling and DNA barcoding. *Coral Reefs* 28:977–986
- Pratchett MS (2001) Influence of coral symbionts on feeding preferences of crown-of-thorns starfish *Acanthaster planci* in the western Pacific. *Mar Ecol Prog Ser* 214:111–119
- Preston EM (1971) Niche overlap and competition among five sympatric congeneric species of xanthid crabs. PhD thesis, Univ Hawaii
- Preston NP, Doherty PJ (1990) Cross-shelf patterns in the community structure of coral-dwelling Crustacea in the central region of the Great Barrier Reef. I. Agile shrimps. *Mar Ecol Prog Ser* 66:47–61
- Randall JE (1967) Food habits of reef fishes of the West Indies. *Stud Trop Ocean* 5:665–847
- Rasher DB, Hoey AS, Hay ME (2013) Consumer diversity interacts with prey defenses to drive ecosystem function. *Ecology* 94:1347–1358
- Reaka-Kudla ML (1997) The global biodiversity of coral reefs: a comparison with rainforests. In: Reaka-Kudla ML, Wilson DE, Wilson EO (eds) *Biodiversity II: Understanding and protecting our natural resources*. National Academy Press, Washington, DC, pp 83–108
- Rouzé H, Lecellier G, Mills SC, Planes S, Berteaux-Lecellier V, Stewart H (2014) Juvenile *Trapezia* spp. crabs can increase juvenile host coral survival by protection from predation. *Mar Ecol Prog Ser* 515:151–159
- Schmitt RJ, Holbrook SJ, Brooks AJ, Lape JCP, Schmitt J, Brooks J, Lape P (2009) Intraguild predation in a structured habitat: distinguishing multiple-predator effects from competitor effects. *Ecology* 90:2434–2443
- Sgarbi L, Melo A (2017) You don't belong here: explaining the excess of rare species in terms of habitat, space and time. *Oikos* 127:497–506
- Shulman MJ, Ogden JC, Ebersole JP, McFarland WN, Miller SL, Wolf NG (1983) Priority effects in the recruitment of juvenile coral reef fishes. *Ecology* 64:1508–1513
- Sin TM, Lee AC (2000) Host specialisation in trapeziid crabs: Consequences for rarity at local scales. *Proc 9th Int Coral Reef Symp* 23–27
- Smallhorn-West PF, Bridge TCL, Munday PL, Jones GP (2017) Depth distribution and abundance of a coral-associated reef fish: roles of recruitment and post-recruitment processes. *Coral Reefs* 36:157–166
- Stella JS, Jones GP, Pratchett MS (2010) Variation in the structure of epifaunal invertebrate assemblages among coral hosts. *Coral Reefs* 29:957–973
- Stella JS, Munday PL, Jones GP (2011) Effects of coral bleaching on the obligate coral-dwelling crab *Trapezia cymodoce*. *Coral Reefs* 30:719–727
- Stewart HL, Holbrook SJ, Schmitt RJ, Brooks AJ (2006) Symbiotic crabs maintain coral health by clearing sediments. *Coral Reefs* 25:609–615
- Stewart HL, Price NN, Holbrook SJ, Schmitt RJ, Brooks AJ (2013) Determinants of the onset and strength of mutualistic interactions between branching corals and associate crabs. *Mar Ecol Prog Ser* 493:155–163
- Stier AC, Geange SW, Hanson KM, Bolker BM (2013) Predator density and timing of arrival affect reef fish community assembly. *Ecology* 94:1057–1068
- Stier AC, Gil MA, McKeon CS, Lemer S, Leray M, Mills SC, Osenberg CW (2012) Housekeeping mutualisms: Do more symbionts facilitate host performance? *PLoS One* 7:2–7
- Stier AC, Leray M (2014) Predators alter community organization of coral reef cryptofauna and reduce abundance of coral mutualists. *Coral Reefs* 33:181–191
- Stier AC, McKeon CS, Osenberg CW, Shima JS (2010) Guard crabs alleviate deleterious effects of vermetid snails on a branching coral. *Coral Reefs* 29:1019–1022
- Veech JA (2013) A probabilistic model for analysing species co-occurrence. *Glob Ecol Biogeogr* 22:252–260
- Vytopil E, Willis BL (2001) Epifaunal community structure in *Acropora* spp. (Scleractinia) on the Great Barrier Reef: implications of coral morphology and habitat complexity. *Coral Reefs* 20:281–288

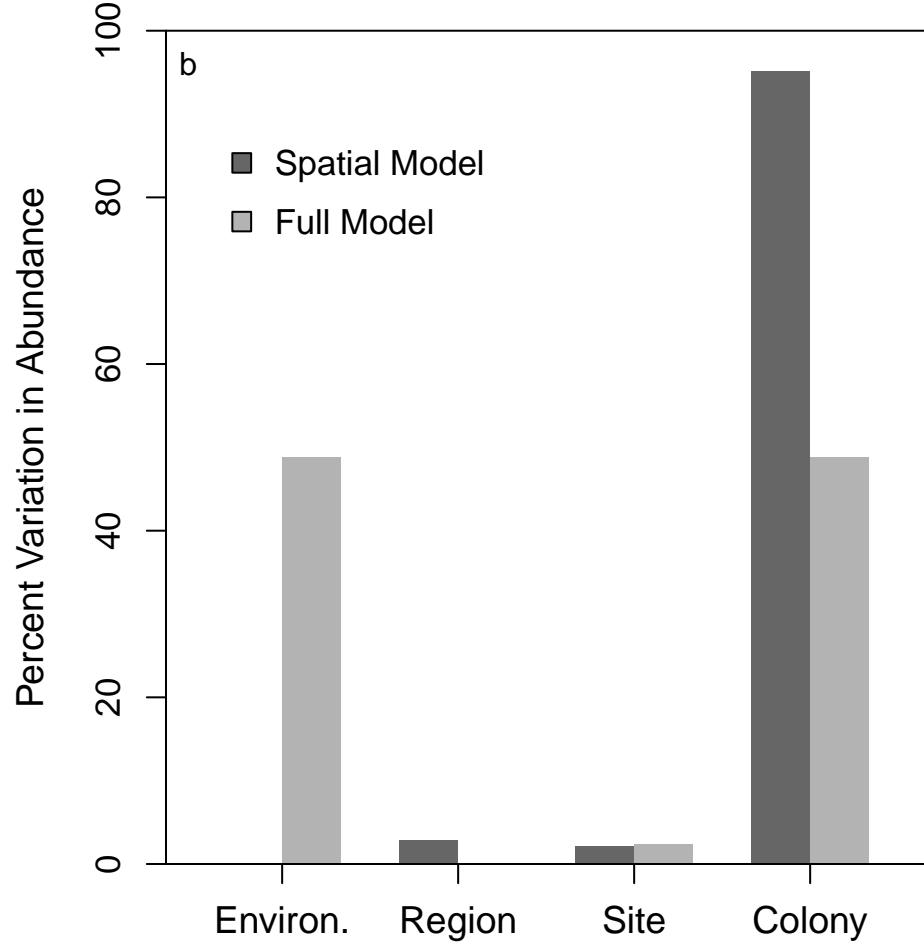
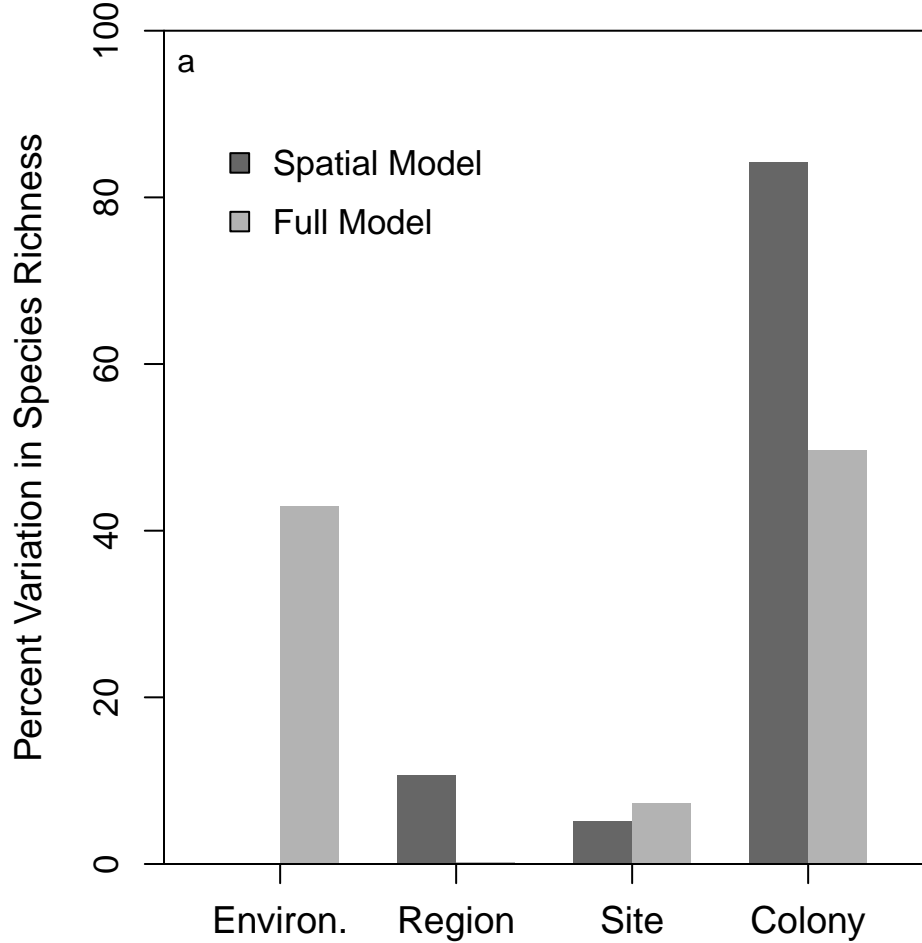
Fig. S1 Paired barplot for percent variation in species richness (a) and community abundance (b) of coral-associated communities (n=708) across spatial scales and environmental factors. Dark gray bars are from a model that included only spatial scales (ESM, Model 1). Light gray bars are from a model that also included environmental factors (ESM, Model 2).

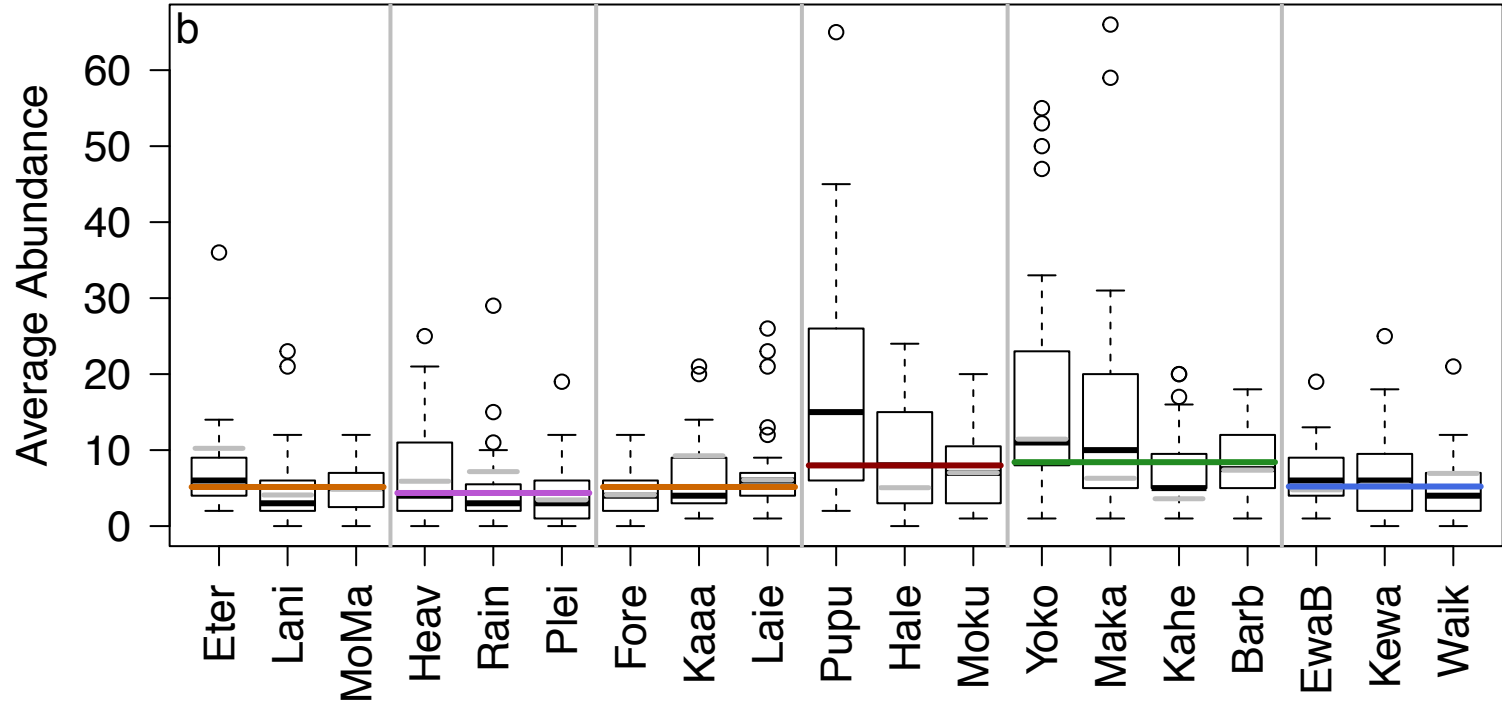
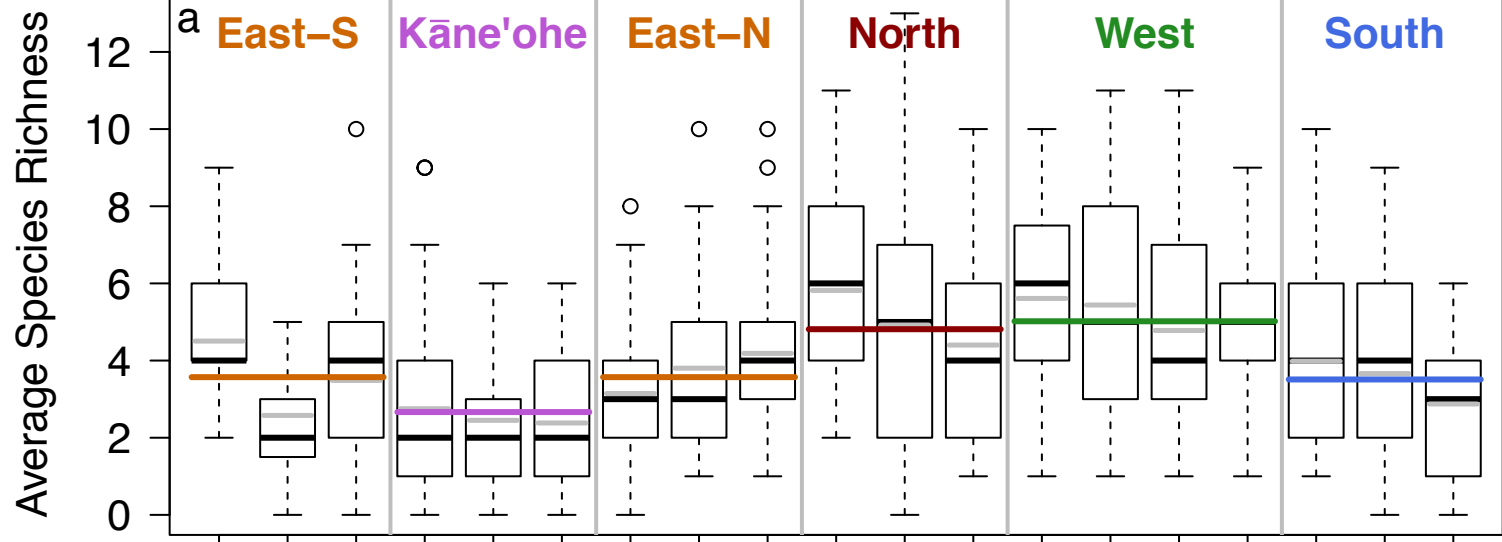
Fig. S2 Mean species richness (a) and community abundance (b) per colony by site, vertical gray lines separating regions. Model estimates of region means are displayed with colored line segments. Within the 25 to 75% quantile box, the light gray horizontal line segments correspond to the model estimates of site means, and the black line segments correspond to the 50% quantile for each site. See Table S4 for site coordinates.

Fig. S3 Plots of the average effect (across regions and sites) of environmental factors from GLMMs (ESM, Model 2) with species richness (a-h) and community abundance (i-p) as the response variables. Gray bars show the 95% confidence interval on model estimates.

Fig. S4 Canonical analysis of principal coordinates constrained by sites (in *R*, function “capscale” in package *vegan*; Oksanen et al. 2017) estimates how much variation in the community composition was explained by variation at the site scale (18%). Sites are shown at their centroids with site codes (see Table S4 for site names).

Fig. S5 Variance in occurrences partitioned across spatial scales based on random effect estimates from a spatial GLMM with occurrence as the response variable, survey structure and species-specific spatial patterns included as random effects (a); and an environmental and spatial GLMM with average and species-specific responses to environmental factors added to the components of the spatial model (b). These models correspond to the “Compositional variation over spatial scales” model (a) and the “Full model” (b, ESM Model 3) in Table S2.





Eter

Lani

MoMa

Heav

Rain

Plei

Fore

Kaaa

Lai

Pupu

Hale

Moku

Yoko

Maka

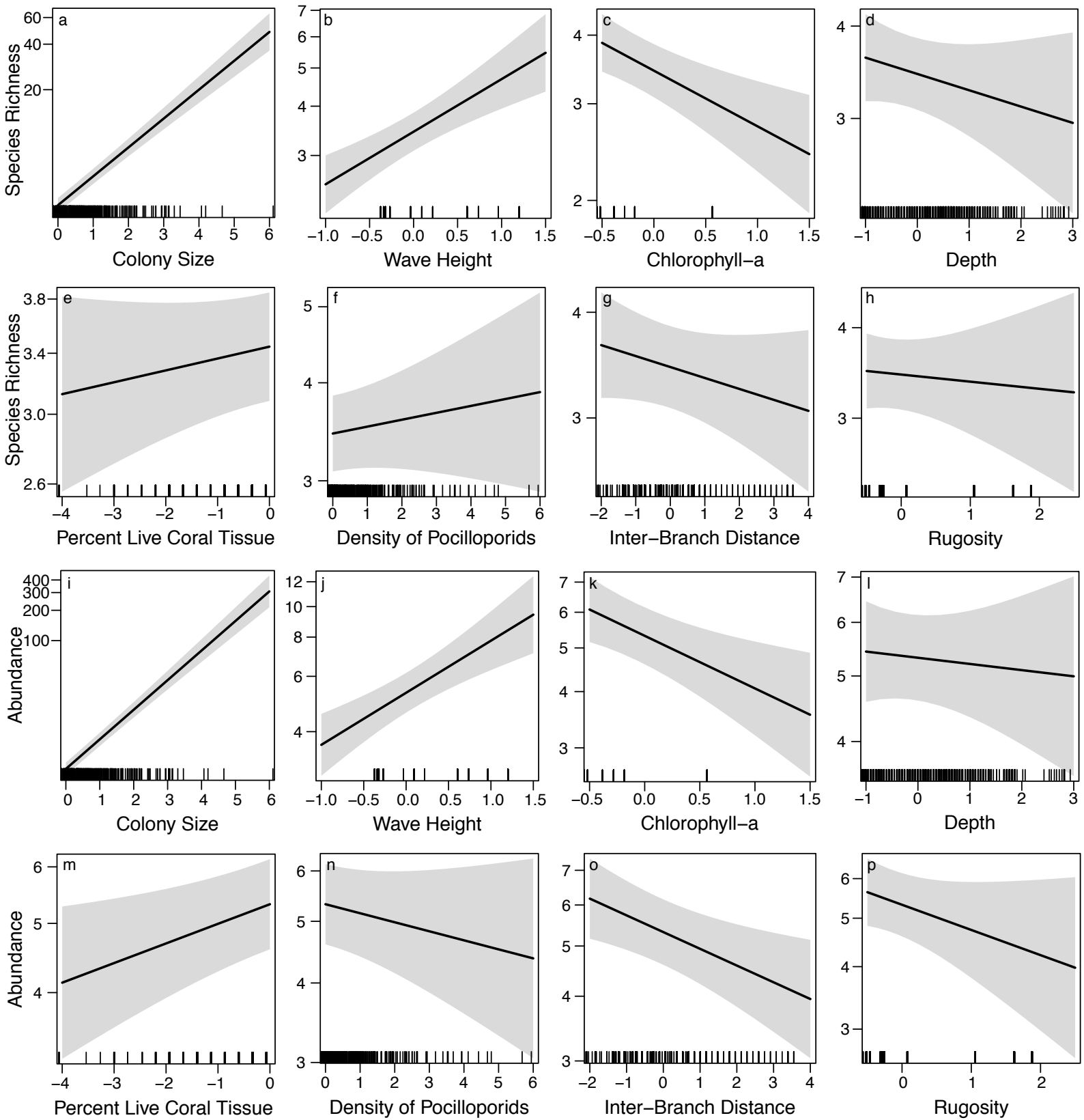
Kahe

Barb

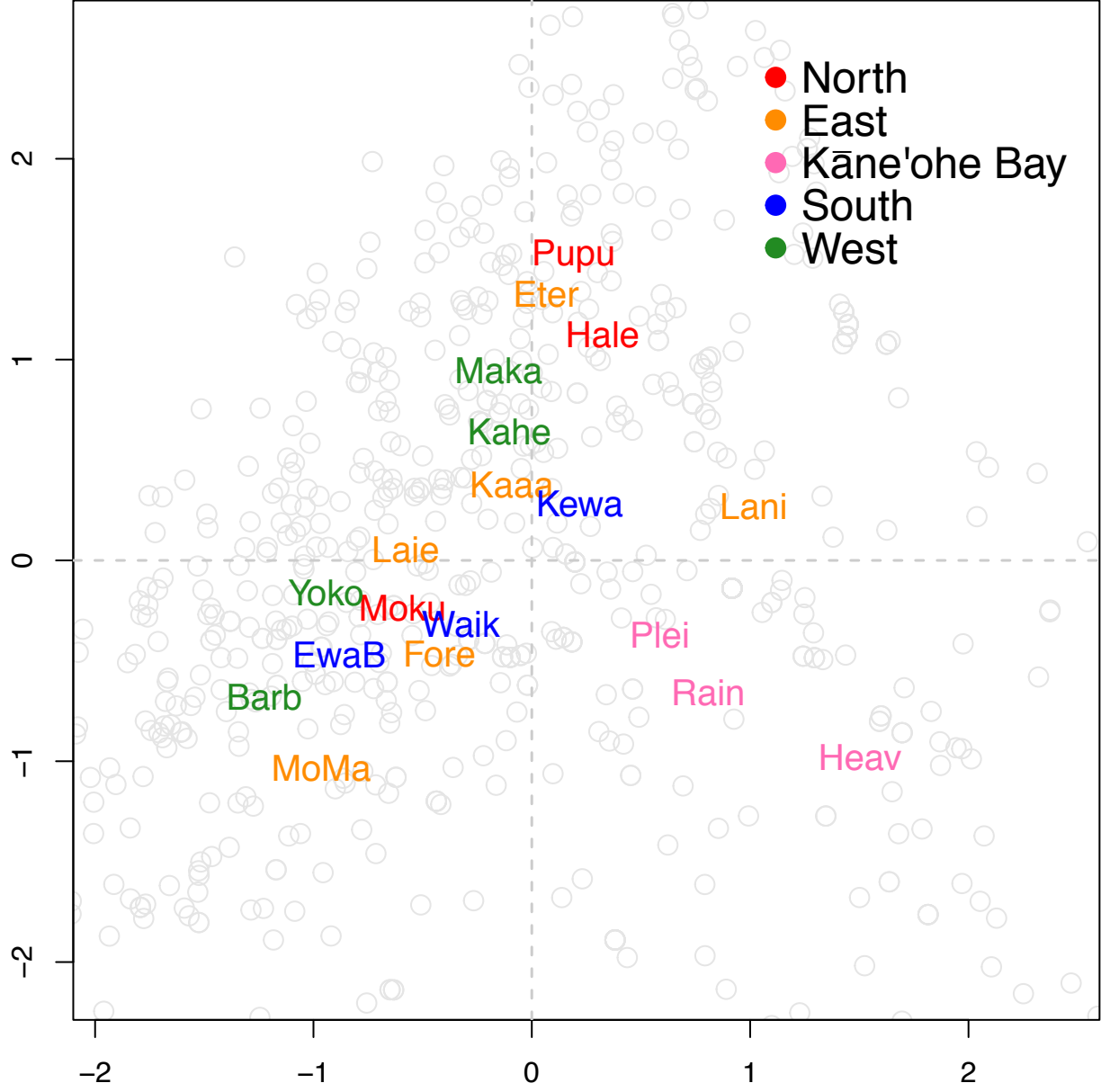
EwaB

Kewa

Waik



Canonical Analysis of Principal Coordinates 2



Canonical Analysis of Principal Coordinates 1

Percent of Variation in Occurrences

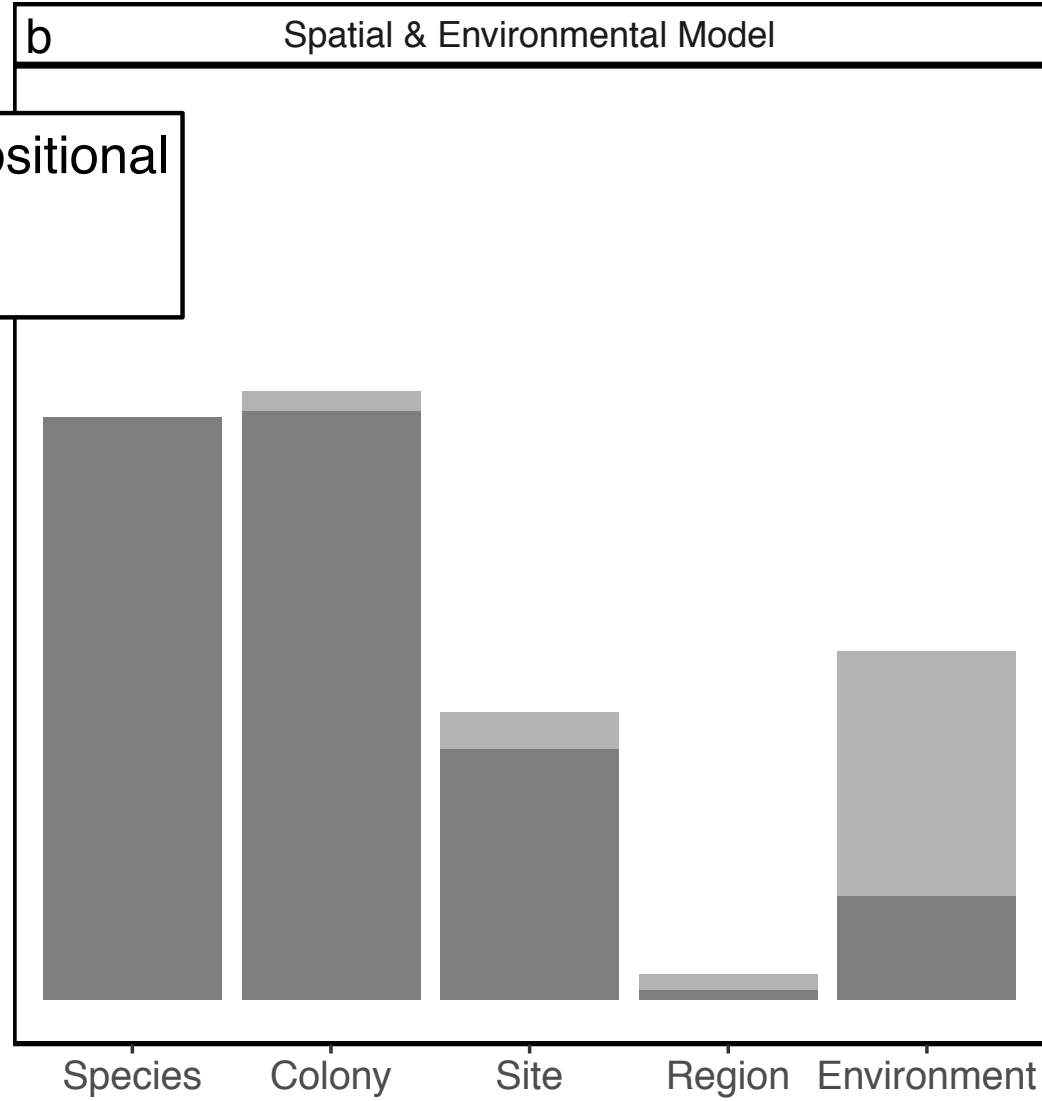
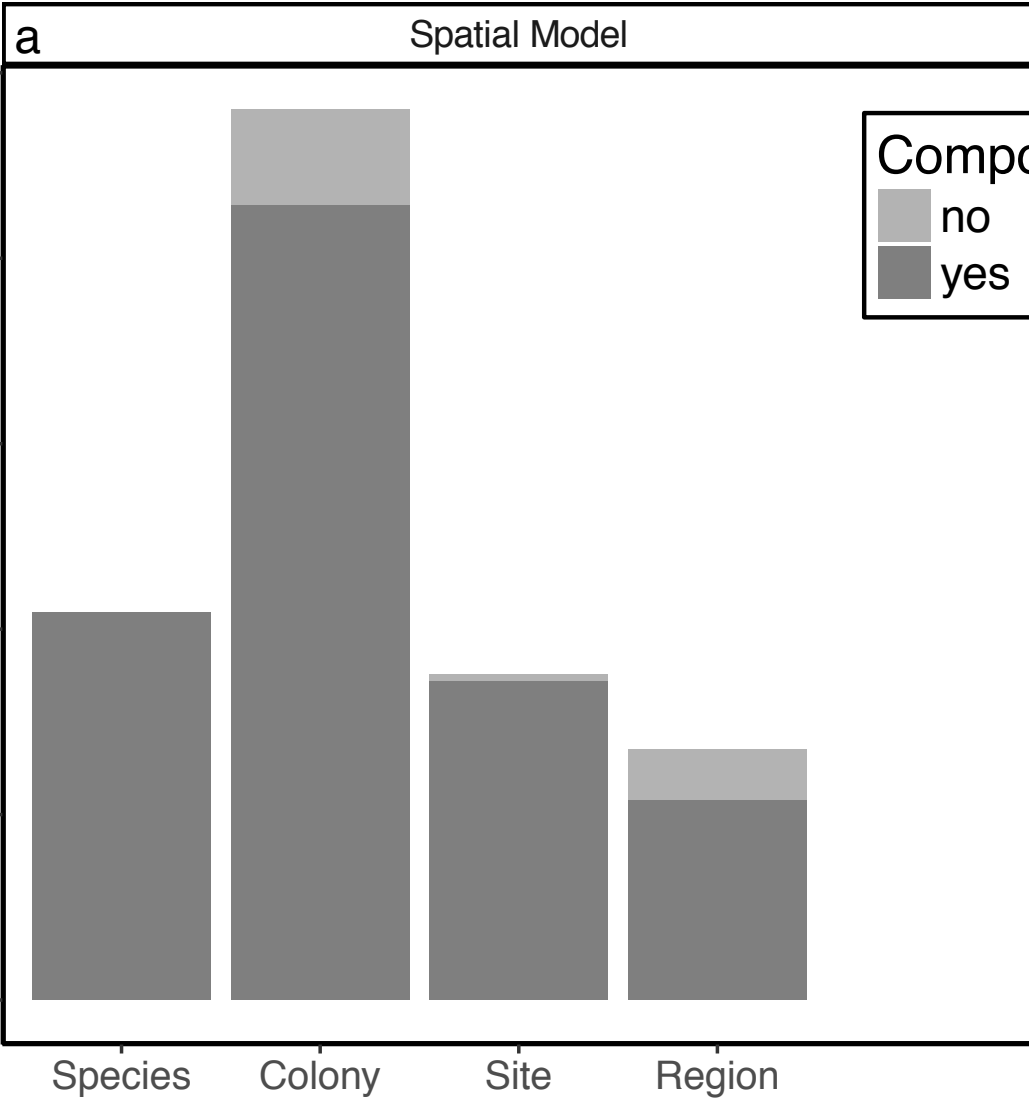


Table S1 Correlation coefficients between environmental drivers measured at the colony and site scale; scores larger than ± 0.5 listed in bold. Correlation scores for inter-branch distance and *Pocillopora* density are based on a subset of 708 colonies for which measurements of these variables were available. All other scores are based on 751 colonies.

Variable	Colony size	% Live tissue	Inter-branch distance	Depth	Density of <i>Poc.</i>	Wave height	Rugosity	Chl-a
Colony size	1							
% Live tissue	-0.333	1						
Inter-branch distance	0.308	0.018	1					
Depth	-0.341	0.110	-0.056	1				
Density of <i>Poc.</i>	-0.131	0.185	0.011	-0.063	1			
Wave height	-0.411	0.176	-0.101	0.534	0.171	1		
Rugosity	-0.183	0.064	-0.080	0.622	0.017	0.411	1	
Chl-a	0.374	-0.214	0.150	-0.603	-0.256	-0.595	-0.399	1

Table S2 Marginal R^2 (variance explained by fixed effects) and conditional R^2 (variance explained by full model including fixed and random effects) for a set of binomial GLMMs run with probability of occurrence for the species observed on $\geq 5\%$ of colonies as the response variable.

Model set	R^2_{marginal} 1	$R^2_{\text{conditional}}$	Fixed effects: colony size, percent live coral tissue, depth, wave height, & chlorophyll-a	Random effects: species identity & colony nested in site nested in region	Random effects: species by site & species by region	Random effects: species-specific responses to fixed effects
Base model	-----	0.3035		☐		
Compositional variation over spatial scales	-----	0.5714		☐	☐	
Average environmental effects	0.0915	0.3251	☐	☐		
Average environmental effects & compositional variation over environmental factors	0.1297	0.6021	☐	☐		☐
Compositional variation over spatial scales & average environmental effects	0.0891	0.5832	☐	☐	☐	
Full model	0.1316	0.6822	☐	☐	☐	☐

Table S3 Family, genus, species, % of regions, % of sites, % of colonies, average number of individuals observed on a colony, depth range, and % live coral tissue range listed for all fish (n=48) and invertebrate (n=43) species observed on 751 *Pocillopora meandrina*. *These rows are complexes of taxonomically similar species that were not consistently distinguishable in the field. *Sebastapistes* spp. includes observations of *S. fowleri*, *S. galactacma*, and *S. ballieui*.

Family	Genus	Species	% of regions	% of sites	% of Colonies	Avg. per colony	Depth (m)	% Live coral tissue
Acanthuridae	<i>Acanthurus</i>	<i>blochii</i>	20	5	0.1	1.0	16.2	95
Acanthuridae	<i>Acanthurus</i>	<i>triostegus</i>	40	16	0.5	1.8	1.5 – 2.7	45 - 100
Acanthuridae	<i>Ctenochaetus</i>	<i>strigosus</i>	20	5	0.1	1	2.4	30
Acanthuridae	<i>Zebrasoma</i>	<i>flavescens</i>	20	5	0.3	1.5	2.4 - 2.7	55 -95
Antennariidae	<i>Antennarius</i>	<i>commerson</i>	20	5	0.1	1	11.6	90
Apogonidae	<i>Pristiapogon</i>	<i>kallopterus</i>	20	5	0.1	1	2.1	85
Blenniidae	<i>Cirripectes</i>	<i>vanderbilti</i>	20	5	0.1	1	4.6	75
Blenniidae	<i>Exallias</i>	<i>brevis</i>	40	21	0.7	1	4.0 – 14.3	70 - 100
Caracanthidae	<i>Caracanthus</i>	<i>typicus</i>	80	79	9.9	1.5	2.4 – 16.8	50 - 100
Chaetodontidae	<i>Chaetodon</i>	<i>lunula</i>	20	11	0.5	2.3	14.3 – 20.4	40 - 80
Chaetodontidae	<i>Chaetodon</i>	<i>multicinctus</i>	20	5	0.1	1	3.7	75
Cirrhitidae	<i>Amblycirrhitus</i>	<i>bimacula</i>	100	42	4.3	1.2	3.0 - 14.3	0 - 100
Cirrhitidae	<i>Cirrhitops</i>	<i>fasciatus</i>	100	32	1.1	1	2.1 - 16.8	55 - 98
Cirrhitidae	<i>Paracirrhites</i>	<i>arcatus</i>	80	79	23.0	1.3	3.4 – 31.1	0 - 100
Cirrhitidae	<i>Paracirrhites</i>	<i>forsteri</i>	60	16	0.4	1	2.4 – 11.6	90 - 95
Gobiidae	<i>Asterropteryx</i>	<i>semipunctatus</i>	20	11	0.4	1	2.1 - 2.7	30 - 95
Gobiidae	<i>Eviota</i>	<i>susanae</i>	20	11	0.4	1.3	2.4 - 2.7	30 - 70
Labridae	<i>Coris</i>	<i>venusta</i>	20	5	0.5	1	2.1	60 - 95
Labridae	<i>Gomphosus</i>	<i>varius</i>	20	5	0.3	1	2.7	90 - 95
Labridae	<i>Novaculichthys</i>	<i>taeniourus</i>	60	16	0.8	1.2	6.7 - 30.8	0 - 100
Labridae	<i>Pseudocheilinus</i>	<i>octotaenia</i>	20	5	0.1	1	15.5	80
Labridae	<i>Pseudocheilinus</i>	<i>tetrataenia</i>	80	47	2.0	1.2	4.9 - 20.3	45 - 100
Labridae	<i>Stethojulis</i>	<i>balteata</i>	40	16	0.5	1	2.1 – 26.2	20 - 90
Labridae	<i>Thalassoma</i>	<i>ballieui</i>	20	5	0.1	1	2.1	90
Labridae	<i>Thalassoma</i>	<i>duperrey</i>	100	53	6.9	1.4	2.1 – 20.1	5 - 100
Lutjanidae	<i>Lutjanus</i>	<i>kasmira</i>	20	5	0.3	1.5	18.0	0 - 60
Monacanthidae	<i>Cantherhines</i>	<i>verecundus</i>	20	11	0.4	1	11.9 – 18.0	0 - 95
Mullidae	<i>Parupeneus</i>	<i>pleurostigma</i>	20	5	0.3	1	2.4	90 - 95
Muraenidae	<i>Gymnomuraena</i>	<i>zebra</i>	20	5	0.1	1	5.8	95
Muraenidae	<i>Gymnothorax</i>	<i>melatremus</i>	20	11	0.3	1	10.7 – 18.3	30 - 55
Muraenidae	<i>Gymnothorax</i>	<i>undulatus</i>	40	11	0.3	1	2.7 – 9.1	55 - 90
Ostraciidae	<i>Ostracion</i>	<i>meleagris</i>	20	5	0.3	1	2.4	30 - 65
Pomacentridae	<i>Centropyge</i>	<i>potteri</i>	20	5	0.3	1	15.8 – 25.9	80 - 95
Pomacentridae	<i>Chromis</i>	<i>vanderbilti</i>	60	21	0.7	1.6	5.8 – 20.1	95 - 100
Pomacentridae	<i>Dascyllus</i>	<i>albisella</i>	100	58	9.2	5.3	1.5 – 30.2	0 - 100
Pomacentridae	<i>Plectroglyphididon</i>	<i>johnstonianus</i>	100	74	9.2	1.3	2.1 – 26.2	0 - 100

Priacanthidae	<i>Heteropriacanthus</i>	<i>cruentatus</i>	40	11	0.4	1	2.1 – 7.0	60 - 90
Scaridae	<i>Chlorurus</i>	<i>spilurus</i>	20	5	0.3	1	2.4	90 - 95
Scaridae	<i>Scarus</i>	<i>psittacus</i>	20	16	0.8	2.7	2.1 – 3.4	30 – 90
Scorpaenidae	<i>Dendrochirus</i>	<i>barberi</i>	60	21	0.9	1	2.1 – 18.3	30 – 90
Scorpaenidae	<i>Pterois</i>	<i>sphex</i>	20	5	0.1	3	14.3	75
Scorpaenidae	<i>Scorpaenopsis</i>	<i>diabolus</i>	20	5	0.1	1	13.4	100
Scorpaenidae	<i>Sebastapistes</i>	<i>coniorta</i>	100	89	15	1.9	2.1 – 19.5	30 - 100
*Scorpaenidae	<i>Sebastapistes</i>	spp.	100	89	23	1.9	2.4 – 30.5	30 - 100
Tetraodontidae	<i>Canthigaster</i>	<i>amboinensis</i>	20	5	0.1	1	7.3	95
Tetraodontidae	<i>Canthigaster</i>	<i>coronata</i>	40	16	0.4	1	10.7 – 14.3	55 - 100
Tetraodontidae	<i>Canthigaster</i>	<i>jactator</i>	80	37	0.9	1.3	2.1 – 17.1	60 - 100
Zanclidae	<i>Zanclus</i>	<i>cornutus</i>	20	5	0.7	1.6	2.1 – 2.4	30 - 90
Alpheidae	<i>Alpheus</i>	<i>lottini</i>	100	100	33.7	1.4	0.6 – 26.8	30 - 100
Alpheidae	<i>Alpheus</i>	<i>pacificus</i>	40	16	0.4	1	2.1 – 11.6	50 - 90
Amphinomidae	<i>Pherecardia</i>	<i>striata</i>	20	5	0.1	1	7.3	70
Carpiliidae	<i>Carpilius</i>	<i>convexus</i>	20	5	0.1	1	2.4	90
Chromodorididae	<i>Thorunna</i>	<i>kahuna</i>	20	5	0.1	1	2.7	45
Cidaridae	<i>Chondrocidaris</i>	<i>gigantea</i>	40	16	0.7	1.6	12.8 – 26.2	0 - 90
Cidaridae	<i>Eucidaris</i>	<i>metularia</i>	80	26	0.9	1.3	4.3 – 20.1	0 - 95
Cryptochiridae	<i>Utinomiella</i>	<i>dimorpha</i>	80	42	16.5	7.9	3.0 - 18.3	30 - 100
Diadematidae	<i>Echinothrix</i>	<i>calamaris</i>	60	21	0.9	1.1	2.1 – 15.8	30 - 100
Diadematidae	<i>Echinothrix</i>	<i>diadema</i>	40	21	0.8	1	10.4 – 15.5	55 - 100
*Paguroidea	unidentified		60	53	5.3	2.4	2.4 – 19.8	0 - 100
Domeciidae	<i>Domecia</i>	<i>hispida</i>	80	47	2.4	1.4	3.4 – 19.8	40 - 100
Echinometridae	<i>Echinometra</i>	<i>mathaei</i>	80	53	3.6	1.5	4.3 – 29.6	0 - 100
Echinometridae	<i>Heterocentrotus</i>	<i>mamillatus</i>	60	21	0.5	1	4.3 – 14.9	0 - 100
Grapsidae	<i>Percnon</i>	<i>affine</i>	20	5	0.1	1	2.4	90
Grapsidae	<i>Percnon</i>	<i>planissimum</i>	100	37	2.1	1.3	2.1 – 19.2	0 - 98
Hippolytidae	<i>Saron</i>	<i>marmoratus</i>	80	37	2.0	1.6	2.1 – 15.5	50 - 100
Hippolytidae	<i>Saron</i>	<i>neglectus</i>	60	26	1.3	2.1	3.7 – 10.4	0 - 80
Hymenoceridae	<i>Hymenocerca</i>	<i>picta</i>	20	5	0.1	3	11.6	95
*Muricidae	<i>Drupella</i>	spp.	60	16	0.8	2	11.6 - 20.1	0 - 95
Muricidae	<i>Quoyula</i>	<i>monodonta</i>	80	42	2.8	1.6	2.7 – 17.7	50 - 100
Ophidiasteridae	<i>Linckia</i>	<i>multifora</i>	20	5	0.1	1	15.8	85
Ophiocomidae	<i>Ophiocoma</i>	<i>erinaceus</i>	100	79	6.4	1.3	2.7 – 20.4	0 - 100
Ophiocomidae	<i>Ophiocoma</i>	<i>pica</i>	100	100	45	1.6	1.5 – 30.5	0 - 100
Palaemonidae	<i>Harpiliopsis</i>	<i>depressa</i>	100	100	34.1	1.7	0.9 – 24.7	0 - 100
Palaemonidae	<i>Palaemon</i>	<i>pacificus</i>	60	32	1.5	1.1	2.1 – 13.4	20 - 100
Portunidae	<i>Charybdis</i>	<i>hawaiensis</i>	80	47	1.9	1.1	2.1 – 25.3	50 - 98
Portunidae	<i>Thalamita</i>	<i>coerulipes</i>	40	16	0.4	1	1.8 - 2.7	60 – 90
Sabellidae	<i>Sabellastarte</i>	<i>spectabilis</i>	60	32	5.7	2.5	2.1 – 18.3	0 - 100
Sepiolidae	<i>Euprymna</i>	<i>scolopes</i>	20	5	0.1	1	12.8	50
*Serpulidae	<i>Spirobranchus</i>	spp.	60	21	0.7	1.4	12.5 – 16.8	80 - 100
Stenopodidae	<i>Stenopus</i>	<i>hispidus</i>	60	21	0.9	1.4	2.1 – 18.3	0 - 100
Stomatopoda	<i>Gonodactylaceus</i>	<i>falcatus</i>	40	11	0.4	1	2.4 – 15.5	65 - 99
Terebellidae	<i>Loimia</i>	<i>medusa</i>	20	5	0.1	1	2.7	10
Terebridae	<i>Terebra</i>	<i>gouldi</i>	20	5	0.1	1	2.7	45
Trapeziidae	<i>Trapezia</i>	<i>bidentata</i>	100	58	7.5	1.6	2.1 – 18.0	45 - 100

Trapeziidae	<i>Trapezia</i>	<i>digitalis</i>	100	79	16.5	1.5	2.1 – 19.5	45 - 100
Trapeziidae	<i>Trapezia</i>	<i>flavopunctata</i>	100	68	6.1	1.7	2.1 – 16.8	50 - 100
Trapeziidae	<i>Trapezia</i>	<i>intermedia</i>	100	100	59.0	1.6	0.6 - 31.1	20 - 100
Trapeziidae	<i>Trapezia</i>	<i>tigrina</i>	100	100	27.7	1.8	1.5 – 26.2	30 - 100
*Vermetidae	unidentified		20	5	0.1	1	2.7	50
Xanthidae	<i>Liomera</i>	<i>rubra</i>	40	11	0.3	1.5	9.1 – 18.0	0 - 100
Xanthidae	<i>Pseudoliomera</i>	<i>speciosa</i>	100	68	2.3	1.4	2.1 – 14.3	75 - 100

Table S4 List of site names, abbreviations, coordinates, regions, and depth ranges for 19 survey sites around O‘ahu.

Site name	Code	Latitude	Longitude	Region	Depth (m)
Heaven	Heav	21.4516	-157.7904	Kāne‘ohe	2.1 – 2.4
Pleiades	Plei	21.4564	-157.7945	Kāne‘ohe	2.4 – 3.7
Rainbow	Rain	21.4549	-157.7947	Kāne‘ohe	2.4 – 3.0
La‘ie	Laie	21.6636	-157.9155	East	11.9 – 14.3
Ka‘a‘awa	Kaaa	21.5664	-157.8436	East	11.0 – 13.4
Kāne‘ohe Forereef	Fore	21.5087	-157.8051	East	9.4 – 30.5
Moku Manu	MoMa	21.4710	-157.7209	East	15.2 – 20.7
Lanikai	Lani	21.3906	-157.7086	East	0.6 – 2.7
Eternity Beach	Eter	21.2812	-157.6766	East	7.3 – 10.4
Waikīkī	Waik	21.2687	-157.8378	South	7.9 - 15.8
Kewalo	Kewa	21.2904	-157.8655	South	6.1 – 17.7
‘Ewa Beach	EwaB	21.2930	-158.0102	South	12.2 - 16.2
Barber's Point	Barb	21.3112	-158.1276	West	9.8 – 25.0
Kahe Point	Kahe	21.3528	-158.1318	West	3.4 – 7.0
Mākaha	Maka	21.4748	-158.2267	West	3.7 – 14.0
Yokohama	Yoko	21.5339	-158.2348	West	13.7 – 18.3
Hale‘iwa	Hale	21.5955	-158.1105	North	2.4 – 8.5
Pupukea	Pupu	21.6521	-158.0634	North	5.8 – 10.4
Mokulē‘ia	Moku	21.5910	-158.2153	North	10.7 – 20.1

Fig. S1 Paired barplot for percent variation in species richness (a) and community abundance (b) of coral-associated communities (n=708) across spatial scales and environmental factors. Dark gray bars are from a model that included only spatial scales (ESM, Model 1). Light gray bars are from a model that also included environmental factors (ESM, Model 2).

Fig. S2 Mean species richness (a) and community abundance (b) per colony by site, vertical gray lines separating regions. Model estimates of region means are displayed with colored line segments. Within the 25 to 75% quantile box, the light gray horizontal line segments correspond to the model estimates of site means, and the black line segments correspond to the 50% quantile for each site. See Table S4 for site coordinates.

Fig. S3 Plots of the average effect (across regions and sites) of environmental factors from GLMMs (ESM, Model 2) with species richness (a-h) and community abundance (i-p) as the response variables. Gray bars show the 95% confidence interval on model estimates.

Fig. S4 Canonical analysis of principal coordinates constrained by sites (in *R*, function “capscale” in package *vegan*; Oksanen et al. 2017) estimates how much variation in the community composition was explained by variation at the site scale (18%). Sites are shown at their centroids with site codes (see Table S4 for site names).

Fig. S5 Variance in occurrences partitioned across spatial scales based on random effect estimates from a spatial GLMM with occurrence as the response variable, survey structure and species-specific spatial patterns included as random effects (a); and an environmental and spatial GLMM with average and species-specific responses to environmental factors added to the components of the spatial model (b). These models correspond to the “Compositional variation over spatial scales” model (a) and the “Full model” (b, ESM Model 3) in Table S2.

Model 1: For the community metrics the model with only random effects was:

$$\begin{aligned}
 Y_{ijk} &\sim \text{Poisson}(\lambda_{ijk}) \\
 \log(\lambda_{ijk}) &= \beta_0 + \alpha_k + \eta_{jk} + \varepsilon_{ijk} \\
 \alpha_k &\sim \text{Normal}(0, \sigma_{\text{region}}) \\
 \eta_{jk} &\sim \text{Normal}(0, \sigma_{\text{site}}) \\
 \varepsilon_{ijk} &\sim \text{Normal}(0, \sigma_{\text{colony}})
 \end{aligned}$$

where Y_{ijk} is community abundance or species richness of colony i at site j in region k , λ_{ijk} is the community metric at colony i in site j in region k , α_k is the random effect for region k , normally distributed with mean zero and standard deviation σ_{region} , η_{jk} is the random effect for site j in region k , normally distributed with mean zero and standard deviation σ_{site} , ε_{ijk} is the random effect for colony i in site j in region k , normally distributed with mean zero and standard deviation σ_{colony} , and β_0 is the overall mean community metric across samples.

R code for this model: `glmer(community_metric ~ (1|Colony)) + (1|Site) + (1|Region), data=oahu_commonspecies, family=poisson)`

Model 2: For the community metrics, the model with fixed and random effects was:

$$\begin{aligned}
 Y_{ijk} &\sim \text{Poisson}(\lambda_{ijk}) \\
 \log(\lambda_{ijk}) &= \beta_0 + \alpha_k + \eta_{jk} + \varepsilon_{ijk} + \beta_1 \times \text{depth}_i + \beta_2 \times \text{percent_live_coral_tissue}_i \\
 &\quad + \beta_3 \times \text{colony_size}_i + \beta_4 \times \text{density_of_Pocillopora}_i + \beta_5 \times \text{branch_distance}_i \\
 &\quad + \beta_6 \times \text{wave_height}_j + \beta_7 \times \text{rugosity}_j + \beta_8 \times \text{chlorophyll_a}_j \\
 \alpha_k &\sim \text{Normal}(0, \sigma_{\text{region}}) \\
 \eta_{jk} &\sim \text{Normal}(0, \sigma_{\text{site}}) \\
 \varepsilon_{ijk} &\sim \text{Normal}(0, \sigma_{\text{colony}})
 \end{aligned}$$

where Y_{ijk} is community abundance or species richness of colony i at site j in region k , λ_{ijk} is the mean

community metric at colony i in site j in region k , α_k is the random effect for region k , normally distributed with mean zero and standard deviation σ_{region} , η_{jk} is the random effect for site j in region k , normally distributed with mean zero and standard deviation σ_{site} , ε_{ijk} is the random effect for colony i in site j in region k , normally distributed with mean zero and standard deviation σ_{colony} , and β_0 is the overall mean community metric across samples.

R code for this model: `glmer(community_metric ~ depth + percent_live_coral_tissue + colony_size + density_of_Pocillopora + branch_distance + wave_height + rugosity + chlorophyll_a + (1|Colony) + (1|Site) + (1|Region), data=common_species, family=poisson)`

Model 3: For community composition, the full model (Table 2: Model 4) was:

$$Y_{ijk}^n \sim \text{Bernoulli}(p_{ijk}^n)$$

$$\text{logit}(p_{ijk}^n) = \beta_{0,ijk}^n + \beta_1^n \times \text{depth}_i + \beta_2^n \times \text{percent_live_coral_tissue}_i + \beta_3^n \times \text{colony_size}_i + \beta_4^n \times \text{wave_height}_j + \beta_5^n \times \text{chlorophyll_a}_j$$

$$\beta_{0,ijk}^n = \bar{\beta}_0 + \varepsilon_i + \eta_j + \alpha_k + \delta^n + \varphi_j^n + \kappa_k^n$$

$$\varepsilon_i \sim \text{Normal}(0, \sigma_{\text{coral}}); \eta_j \sim \text{Normal}(0, \sigma_{\text{site}}); \alpha_k \sim \text{Normal}(0, \sigma_{\text{region}}); \delta^n \sim \text{Normal}(0, \sigma_{\text{species}});$$

$$\varphi_j^n \sim \text{Normal}(0, \sigma_{\text{site:species}}); \kappa_k^n \sim \text{Normal}(0, \sigma_{\text{region:species}})$$

$$\beta_1^n = \bar{\beta}_1 + b_1^n; b_1^n \sim \text{Normal}(0, \sigma_{\text{depth}})$$

$$\beta_2^n = \bar{\beta}_2 + b_2^n; b_2^n \sim \text{Normal}(0, \sigma_{\text{percent_live_coral_tissue}})$$

$$\beta_3^n = \bar{\beta}_3 + b_3^n; b_3^n \sim \text{Normal}(0, \sigma_{\text{colony_size}})$$

$$\beta_4^n = \bar{\beta}_4 + b_4^n; b_4^n \sim \text{Normal}(0, \sigma_{\text{wave_height}})$$

$$\beta_5^n = \bar{\beta}_5 + b_5^n; b_5^n \sim \text{Normal}(0, \sigma_{\text{chlorophyll_a}})$$

where p_{ijk}^n is the probability of observing species n on colony i at site j in region k , Y_{ijk}^n is the estimated probability of observing species n on colony i at site j in region k , $\bar{\beta}_0$ is the overall mean probability of occurrence, ε_i is the random effect for colony i , η_j is the random effect for site j , α_k is the random effect

for region k , δ^n is the random effect for species n , φ_j^n is the random effect for species n at site j , κ_k^n is the random effect for species n at region k , each random effect is normally distributed with mean zero and a unique (for each random effect) standard deviation σ , residual variance in this model is at the scale of colony:species. This model is analogous to a constrained ordination, but uses a full statistical model and allows for hierarchical random variation.

R code for this model: `glmer(presence_absence ~ depth + percent_live_coral_tissue + colony_size + wave_height + chlorophyll_a + (1|Colony) + (1|Site) + (1|Region) + (1 + depth + percent_live_coral_tissue + colony_size + wave_height + chlorophyll_a |species) + (1|Site:species) + (1|Region:species), data=really_common_species, family=binomial)`