

**Human disturbance alters Pacific coral reef fish  $\beta$ -diversity at three spatial scales**

by

Logan Douglas Wiwchar  
BSc, University of Alberta, 2011

A Thesis Submitted in Partial Fulfillment  
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biology

© Logan Douglas Wiwchar, 2014  
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by  
photocopy or other means, without the permission of the author.

## **Supervisory Committee**

**Human disturbance alters Pacific coral reef fish  $\beta$ -diversity at three spatial scales**

by

Logan Douglas Wiwchar  
BSc, University of Alberta, 2011

### **Supervisory Committee**

Dr. Julia Baum; Supervisor  
**Department of Biology**

Dr. Verena Tunnicliffe; Departmental Member  
**Department of Biology**

Dr. Jana McPherson; Departmental Member  
**Department of Biology**

Dr. Brian Starzomski; Outside Member  
**School of Environmental Studies**

## Abstract

### Supervisory Committee

Dr. Julia Baum; Supervisor  
Department of Biology

Dr. Verena Tunnicliffe; Departmental Member  
Department of Biology

Dr. Jana McPherson; Departmental Member  
Department of Biology

Dr. Brian Starzomski; Outside Member  
School of Environmental Studies

Coral reefs are the most diverse marine ecosystem, but are increasingly threatened by local and global anthropogenic changes. In this thesis, I examine the impact of local stressors on the spatial variability of coral reef fish community composition by modeling the  $\beta$ -diversity of 35 islands across the Pacific Ocean that are characterized by either low or high human disturbance. By examining  $\beta$ -diversity across three spatial scales (within island, within island group, and across island group), and using null models to control for differences in alpha-diversity or abundance, I reveal previously undocumented effects of human disturbance on coral reef fish assemblages. At all scales, human disturbances alter  $\beta$ -diversity. At the largest-scale, islands with high human disturbance have lower incidence- and abundance-based  $\beta$ -diversity, consistent with biotic homogenization. This pattern was driven by both species with high and low abundances that differed across islands. At the smaller two scales (within islands or island groups), the presence of low abundance species is more variable on islands with high human disturbance (manifest in greater incidence-based  $\beta$ -diversity), but these islands have lower abundance-based  $\beta$ -diversity driven by moderately abundant and widespread species. Multivariate techniques show that islands with high human disturbance have a weaker species-environment relationship, and as such, I suggest that homogenization of coral reef fish assemblages by human disturbances is resulting in greater stochasticity of species composition.

# Table of Contents

<b>Supervisory Committee .....</b>	<b>ii</b>
<b>Abstract .....</b>	<b>iii</b>
<b>Table of Contents .....</b>	<b>iv</b>
<b>List of Tables .....</b>	<b>vi</b>
<b>List of Figures .....</b>	<b>xii</b>
<b>Acknowledgments .....</b>	<b>xvi</b>
<b>1. Introduction .....</b>	<b>1</b>
<b>1.1 Coral Reef Biodiversity .....</b>	<b>2</b>
<b>1.2 Beta Diversity .....</b>	<b>6</b>
<b>1.3 Coral Reef <math>\beta</math>-diversity .....</b>	<b>9</b>
<b>1.4 Objectives and Hypothesis.....</b>	<b>11</b>
<b>2. Methods.....</b>	<b>17</b>
<b>2.1 Study Region .....</b>	<b>17</b>
<b>2.2 Data.....</b>	<b>20</b>
2.2.1 Underwater Visual Censuses.....	20
2.2.2 Explanatory Variables – Environmental.....	20
2.2.3 Explanatory Variables – Human Disturbance.....	23
<b>2.3 Modeling <math>\beta</math>-diversity .....</b>	<b>24</b>
2.3.1 Beta Diversity Study Design .....	25
2.3.2 Beta Diversity Metrics.....	29
<b>2.4 Modeling Drivers of <math>\beta</math>-Diversity.....</b>	<b>32</b>
2.4.1 Modeling Anthropogenic and Biophysical Influences on $\beta$ -Diversity.....	33
2.4.2 Modelling Anthropogenic and Biophysical Influences on Community Composition .	34
<b>3. Results .....</b>	<b>38</b>



<b>3.1 Island <math>\beta</math>-diversity</b> .....	<b>43</b>
3.1.1 Incidence-Based $\beta_{RC}$ Models .....	47
3.1.2 Abundance-Based, $\alpha$ -Constrained $\beta_W$ Models .....	49
<b>3.2 Community Composition: Distance-Based Redundancy Analysis</b> .....	<b>51</b>
3.2.1 Incidence Based Community Composition .....	51
3.2.2 Abundance Based Community Composition .....	52
<b>4. Discussion</b> .....	<b>58</b>
4.1 Effect of Scale on Spatial Variability.....	59
4.2 Effect of Human Disturbance on Spatial Variability.....	61
4.3 Regional Differences in Spatial Variability.....	66
4.4 Effect of Environmental Variables on Spatial Variability.....	68
4.5 Implications and Concluding Remarks .....	71
<b>Appendices</b> .....	<b>74</b>
<b>Appendix A: Null Model Code</b> .....	<b>74</b>
<b>Appendix B: Exploratory Data Analyses</b> .....	<b>78</b>
Regional Overlap of Families .....	78
Correlations Among $\beta$ -diversity Values.....	79
Correlations Among Predictor Variables .....	83
Distance Correction .....	86
<b>Appendix C: Additional <math>\beta</math>-diversity Model Information</b> .....	<b>88</b>
<b>Appendix D: Species Lists</b> .....	<b>90</b>
<b>Appendix E: Testing the Robustness of Human Disturbance Categories</b> .....	<b>99</b>
<b>Appendix F: <math>\beta</math>-diversity Models Without Live Coral Complexity</b> .....	<b>101</b>
<b>Bibliography</b> .....	<b>106</b>

## List of Tables

- Table 1.** Predicted univariate effect of each explanatory variable on  $\beta$ -diversity at different scales. Scale refers to  $\beta$ -diversity calculated within islands (1), within island groups (2), or across island groups (3). ‘All’ denotes that the hypothesis applies to all three scales. H.A., M.A., and A.S. = Hawaiian Archipelago, Mariana Archipelago, and American Samoa respectively. IPC $\Delta$  = Indo-Pacific Coral Triangle; SST<sub>mean</sub> = mean sea surface temperature; SST<sub>var</sub> = interannual variability in sea surface temperature. .... 12
- Table 2.** Island-scale information for all islands surveyed. Surveys not included in the analyses herein indicated by strike through text. .... 19
- Table 3.** Data sources and resolution of the explanatory variables. IPC $\Delta$  = Indo-Pacific Coral Triangle. .... 21
- Table 4.** Mean fish counts and observed species richness ( $\pm$  SEM) at sites on all islands, or low and high human disturbance islands within each island group. Total observed richness of islands or regions within each human disturbance category is also shown. The mean proportion of sites that each species were observed at (Freq. Occur.) is presented for islands with low and high human disturbance. Data is from second survey period, however the same trends were observed in both survey periods. .... 39
- Table 5.** Island richness, number of- and proportion relative to island richness of- species observed only once or twice on each island (singletons and doubletons), as well as the most abundant and most frequently observed species per island. ‘All’ denotes the aggregation of all sites on islands within either human disturbance category (H. Distb). H.A. = Hawaiian Archipelago; M.A. = Mariana Archipelago; A.S.= American Samoa. 41

**Table 6.** Number of species *within* each island group observed on islands of one disturbance category but not the other (# Unique), as well as the median  $L_{\max}$  (maximum length of fish) and general trophic group of those species. H.A. = Hawaiian Archipelago; M.A. = Mariana Islands Archipelago; A.S. = American Samoa; L. Carn = low carnivores (invertivores, small piscivores, and corallivores); Plankt = planktivores; Herb = herbivores; Pisc = piscivores; FG = Functional Group..... 42

**Table 7.** Model Selection and parameter estimation for  $\beta_{RC}$  and  $\beta_W$  at three spatial scales. Positive or negative signs (+ or -) denote whether the model averaged parameter estimate was positive or negative, and the number of symbols indicates the relative size of parameter estimate. Three symbols indicates the largest effect size, two or one symbols notates parameter estimates that are one or two (or greater) orders of magnitude less than the largest effect size. Asterisks are used for categorical variables where positive or negative notation is non-intuitive. All models within  $2 \Delta AIC$  were used for model averaging and the proportion of those models that each parameter was included in is indicated. Values of model weighted parameter estimates are in Appendix C.  $\beta_{IG}$ ,  $\beta_H$ ,  $\beta_{LCC}$ , ( $\beta_{LCC} : \beta_H$ ),  $\beta_{WE}$ ,  $\beta_A$ ,  $\beta_{SST_m}$ , and  $\beta_{SST_v}$  denote effect of island group, human disturbance, live coral cover, the interaction between live coral cover and humans on high disturbance islands, wave energy, reef area, mean sea surface temperature, and interannual variability in sea surface temperature respectively..... 46

**Table 8.** Candidate models of  $\beta_{RC}$  at all three scales: within island  $\beta_{RC,1}$ ; within island group  $\beta_{RC,2}$ ; and across island group  $\beta_{RC,3}$ . AIC weights ( $AIC_w$ ) were calculated to sum to 1 over candidate models with  $\Delta AIC \leq 2$ . Pseudo  $R^2$  indicates the proportion of deviation explained by model relative to deviation in null (intercept only) model.  $SST_v$  and  $SST_v$  denote mean and interannual variability in sea surface temperature..... 47

**Table 9.** Candidate models of  $\alpha$ -constrained  $\beta$ -deviations ( $\beta_W$ ) at all three scales: within island  $\beta_{W,1}$ ; within island group  $\beta_{W,2}$ ; and across island group  $\beta_{W,3}$ . AIC weights ( $AIC_w$ ) were calculated to sum to 1 over candidate models with  $\Delta AIC \leq 2$ . Pseudo  $R^2$  indicates the proportion of deviation explained by model relative to deviation in null (intercept only) model.  $SST_v$  and  $SST_v$  denote mean and interannual variability in sea surface temperature..... 50

**Table 10.** Proportion of variance in community dissimilarity (assessed by Raup-Crick null model dissimilarity) at the within island and within island group scales explained by site-specific environmental variables (Env: mean and interannual variability in sea surface temperature, live coral cover, and coral rugosity) or site-specific + island-specific variables (Total: Env + wave energy, reef area, island composition). Single predictor that explained the highest proportion of variation at each scale (Best) is also included. Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 10 sites from each island (within island scale) or 70 sites from each low or high human disturbance subset from within island group community dissimilarity matrices. H.A. = Hawaiian Archipelago; M.A. = Mariana Archipelago; A.S. = American Samoa; Rug = rugosity; Cover = coral cover. Asterisk indicates significant difference ( $p < 0.05$ ) among low and high human disturbance islands when islands pooled across island groups. 51

**Table 11.** Mean proportion of variability of community composition within islands explained by interannual variability of sea surface temperature ( $SST_v$ ), mean annual sea surface temperature ( $SST_m$ ), reef rugosity (rug.), live coral cover (cover), and a combination of all four (Env.). Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 10 sites from each low or high human disturbance island of Raup-Crick dissimilarity or  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ ) based dissimilarity matrices within islands. .... 53

**Table 12.** Mean proportion of variability of community composition within island groups explained by site-specific variables (Env.), island-specific variables (Island), or both (Max). Site-specific variables include interannual variability and mean sea surface temperature ( $SST_v$  and  $SST_m$  respectively), reef rugosity (rug.), and live coral cover (cover). Island-specific variables include wave energy (Wave), reef area (Area) and island composition (Comp.). Also shown is the variance explained by each variable alone. Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 70 sites from either low or high human disturbance subsets of Raup-Crick dissimilarity or  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ ) based dissimilarity matrices at the within island group scale (which uses comparisons both within and across islands within the same island group). .... 53

**Table 13.** Mean proportion of variability of community composition across island groups explained by site-specific variables (Env.), island-specific variables (Island), or both (Max). Site-specific variables include interannual variability and mean sea surface temperature ( $SST_v$  and  $SST_m$  respectively), reef rugosity (rug.), and live coral cover (cover). Island-specific variables include wave energy (Wave), reef area (Area) and island composition (Comp.). Also shown is the variance explained by each variable alone. Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 353 sites from either low or high human disturbance subsets of Raup-Crick dissimilarity or  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ ) based dissimilarity matrices at the across island group scale (which uses comparisons both within and across islands within the same island group, as well as across island groups). .... 54

**Table 14.** Proportion of variance in community dissimilarity (assessed by  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ )) at the within island and within island group scales explained by site-specific environmental variables (Env: mean and interannual variability in sea surface temperature, live coral cover, and coral rugosity) or site-specific + island-specific variables (Total: Env + wave energy, reef area, island composition). Single predictor that explained the highest proportion of variation at each scale (Best) is also included. Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 10 sites from each island (within island scale) or 70 sites from each low or high human disturbance subset from within island group community dissimilarity matrices. H.A. = Hawaiian Archipelago; M.A. = Mariana Archipelago; A.S. = American Samoa; Rug = rugosity; Cover = coral cover; SST<sub>m</sub> = mean sea surface temperature. .... 56

**Supplemental Table 1.** Model weighted parameter estimation for incidence-based Raup-Crick  $\beta$ -diversity,  $\beta_{RC}$ , weighted using all models within 2  $\Delta$ AIC of best model. Intercept represents an island in the Hawaiian Archipelago. Unconditional variance (Variance) as per Buckland et al. (1997); # Models = the number of models within 2  $\Delta$ AIC that each parameter was included in. .... 88

**Supplemental Table 2.** Model weighted parameter estimation for  $\alpha$ -constraining abundance-based null model  $\beta$ -diversity,  $\beta_w$ , weighted using all models within 2  $\Delta$ AIC of best model. Intercept represents an island in the Hawaiian Archipelago. Unconditional variance (Variance) as per Buckland et al. (1997); # Models = the number of models within 2  $\Delta$ AIC that each parameter was included in..... 89

**Supplemental Table 3.** Species observed on islands with low-, but not high- human disturbance of the Hawaiian Archipelago, their  $L_{max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed only during second survey period (2011 – 2012)..... 91

**Supplemental Table 4.** Species observed on islands with high-, but not low- human disturbance of the Hawaiian Archipelago, their  $L_{max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed only during second survey period (2011 – 2012)..... 92

**Supplemental Table 5.** Species observed on islands with low-, but not high- human disturbance of the Mariana Archipelago, their  $L_{max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed during second survey period (2011 – 2012)..... 93

**Supplemental Table 6.** Species observed on islands with high-, but not low- human disturbance of the Mariana Archipelago, their  $L_{max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed during second survey period (2011 – 2012)..... 94

**Supplemental Table 7.** Species observed on islands with low-, but not high- human disturbance of American Samoa, their  $L_{\max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed only during second survey period (2011 – 2012). ..... 96

**Supplemental Table 8.** Species observed on islands with high-, but not low- human disturbance of American Samoa, their  $L_{\max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed only during second survey period (2011 – 2012). ..... 97

**Supplemental Table 9.** Within island  $\beta_{RC}$  and  $\beta_W$  for heavily populated islands calculated using subsets of sites that had human populations of greater than 50000 individuals living within a 20 km radius..... 99

**Supplemental Table 10.** Mean proportion of variability of community composition within islands explained by interannual variability of sea surface temperature ( $SST_v$ ), mean annual sea surface temperature ( $SST_m$ ), reef rugosity (rug.), live coral cover (cover), and a combination of all four (Env.). Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 10 sites from all sites on each island (All) or only sites with a human populations of greater than 50000 individuals living within a 20 km radius (High Pop. Subset). Raup-Crick dissimilarity and  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_W$ ) based dissimilarity were each used for separate dbRDA..... 100

**Supplemental Table 11.** Candidate models of  $\beta_{RC}$  where live coral complexity was not included as an explanatory variable at all three scales: within island  $\beta_{RC,1}$ ; within island group  $\beta_{RC,2}$ ; and across island group  $\beta_{RC,3}$ . AIC weights ( $AIC_w$ ) were calculated to sum to 1 over candidate models with  $\Delta AIC \leq 2$ . Pseudo  $R^2$  indicates the proportion of deviation explained by model relative to deviation in null (intercept only) model.  $SST_v$  and  $SST_v$  denote mean and interannual variability in sea surface temperature. .... 102

**Supplemental Table 12.** Model weighted parameter estimation for incidence based Raup-Crick  $\beta$ -diversity,  $\beta_{RC}$ , where live coral complexity was not included as a potential explanatory variable. Weighting using all models within 2  $\Delta AIC$  of best model. Intercept represents an island in the Hawaiian Archipelago. Unconditional variance (Variance) as per Buckland et al. (1997); # Models = the number of models within 2  $\Delta AIC$  that each parameter was included in. .... 103

**Supplemental Table 13.** Candidate models of  $\beta_W$  where live coral complexity was not included as an explanatory variable at all three scales: within island  $\beta_{W,1}$ ; within island group  $\beta_{W,2}$ ; and across island group  $\beta_{W,3}$ . AIC weights ( $AIC_w$ ) were calculated to sum to 1 over candidate models with  $\Delta AIC \leq 2$ . Pseudo  $R^2$  indicates the proportion of deviation explained by model relative to deviation in null (intercept only) model.  $SST_v$  and  $SST_v$  denote mean and interannual variability in sea surface temperature. .... 104

**Supplemental Table 14.** Model weighted parameter estimation for  $\alpha$ -constraining abundance-based null model  $\beta$ -diversity,  $\beta_w$ , where live coral complexity was not included as a potential explanatory variable. Weighting using all models within 2  $\Delta$ AIC of best model. Intercept represents an island in the Hawaiian Archipelago. Unconditional variance (Variance) as per Buckland et al. (1997); # Models = the number of models within 2  $\Delta$ AIC that each parameter was included in..... 105

## List of Figures

**Figure 1.** Islands surveyed by NOAA CRED Pacific RAMP between 2009 and 2012 in the Pacific Ocean. Black and red symbols indicate islands with low or high human disturbance respectively in the Hawaiian Archipelago (squares), Mariana Archipelago (circles) and American Samoa (triangles). Scale bar represents distances at equator. .... 18

**Figure 2.** Spatially hierarchical sampling design. A) Overview of  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$ , which represents  $\beta$ -diversity for island  $k$  at three scales: within island, within island group, and across island groups, respectively. Within each island group,  $m - n$ ,  $\beta$ -diversity is calculated for all islands,  $k - l$ , where each island has many sites,  $i - j$ . Only combinations of islands and sites within the same human disturbance regime (low or high; black or red symbols respectively) are used to calculate  $\beta$ -diversity for each island. Dotted line represents distinction between island groups. At the within island scale (B), all sites on island  $k$ ,  $i - j$ , are used to calculate  $\beta_1$  for island  $k$ . At the within island group scale (C), to calculate  $\beta_2$  for island  $k$ , all sites on island  $k$  are compared to all other sites on islands  $k+1 - l$ , within that island group,  $m$ . At the among island groups scale (D), to calculate  $\beta_3$  for island  $k$  in island group  $m$ , all sites on island  $k$  in island group  $m$  are compared to all other sites on islands that are in a different same island group,  $m+1 - n$ . ..... 25

**Figure 3.** Relationship between island area and mean distance between sites on each island (rescaled to mean = 0 and standard deviation = 1). Each point is one island, with either high (red) or low (black) human disturbance in the Hawaiian Archipelago (squares), Mariana Archipelago (circles) and American Samoa (triangles). Only island surveyed in the second survey period are plotted. .... 26



**Figure 4.** Pairwise dissimilarity among sites within islands calculated from an abundance based null model (A – C) and incidence based null model (D – F) as a function of geographical distance (in arbitrary units) between site. Sites on low (black) and high (red) human disturbance islands of the Hawaiian Archipelago (A & D), Mariana Islands Archipelago (B & E) and American Samoa (C & F) are plotted from the most recent survey period only. Lines show linear relationship from all pairwise dissimilarities within a given human disturbance category for each island group. Trends are consistent across larger spatial scales (not shown). . . . . 27

**Figure 5.** Overlap of species observed in the Hawaiian Archipelago (A.S.; blue), Mariana Archipelago (M.A.; pink) and American Samoa (A.S.; yellow) (A), and among low and high human disturbance islands within each island group respectively (B – D). Overlap in A represents species observed during either survey period, whereas B – D represents only the second survey period. . . . . 40

**Figure 6.** Proportion of all species present on islands that have low (left bar) or high (right cross-hatch bar) human disturbance in each island group that are piscivores (red), herbivores (green), planktivores (blue) or low carnivores (invertivores, small piscivores, and corallivores; brown). . . . . 42

**Figure 7.** Island  $\beta$ -diversity values derived from Jaccard dissimilarity (A-C), Raup-Crick dissimilarity (D-F), Bray-Curtis dissimilarity (G-I), Kraft  $\beta$ -deviations (J-L), and Wiwchar  $\beta$ -deviations (M-O). Within island, within island group, and across island group  $\beta$ -diversity are in columns one to three, with islands of low (L; black) and high (H; red) human disturbance plotted for the Hawaiian Archipelago (squares), Mariana Archipelago (circles) and American Samoa (triangles). . . . . 44

**Figure 8.** Distance-based redundancy analysis ordination of Raup-Crick site-site community dissimilarity (A) and  $\alpha$ -constraining  $\beta$ -deviation based community dissimilarity (B) at the Pacific Ocean scale. Sites on low (black) and high (red) human disturbance islands during the second survey period (2011 – 2012) of Hawaiian Archipelago (squares), Mariana Archipelago (circles) and American Samoa (triangles) were used for ordination and are plotted. Length of arrow indicates relative importance of variables. Humans = high disturbance islands; SST<sub>m</sub> and SST<sub>v</sub> = mean and interannual variability in sea surface temperature respectively; cover = live coral cover. . . . . 55

**Supplemental Figure 1.** Overlap of families observed in the Hawaiian Archipelago (H.A.; blue), Mariana Archipelago (M.A.; pink), and American Samoa (A.S.; yellow). Data includes surveys performed during both survey periods (2009 – 2012). . . . . 78

**Supplemental Figure 2.** Correlation among island  $\beta$ -diversity metrics at the within island scale ( $\beta_1$ ). Diagonal panels show distribution of  $\beta_1$ , with y-axis ranging from 0 – 20. Upper panels show spearman rho values that correspond to bottom panels where islands with low (black) and high (red) human disturbance of the Hawaiian Archipelago (squares), Mariana Archipelago (circles), and American Samoa are plotted. Linear correlations are represented with solid lines. Island  $\beta$ -diversity values from both survey periods from are shown..... 80

**Supplemental Figure 3.** Correlations among island  $\beta$ -diversity metrics at the within island group scale ( $\beta_2$ ). Diagonal panels show distribution of  $\beta_2$ , with y-axis ranging from 0 – 20. Upper panels show spearman rho values that correspond to bottom panels where islands with low (black) and high (red) human disturbance of the Hawaiian Archipelago (squares), Mariana Archipelago (circles), and American Samoa are plotted. Linear correlations are represented with solid lines. Island  $\beta$ -diversity values from both survey periods from are shown..... 81

**Supplemental Figure 4.** Correlations among island  $\beta$ -diversity metrics at the across island group scale ( $\beta_3$ ). Diagonal panels show distribution of  $\beta_3$ , with y-axis ranging from 0 – 20. Upper panels show spearman rho values that correspond to bottom panels where islands with low (black) and high (red) human disturbance of the Hawaiian Archipelago (squares), Mariana Archipelago (circles), and American Samoa are plotted. Linear correlations are represented with solid lines. Island  $\beta$ -diversity values from both survey periods from are shown..... 82

**Supplemental Figure 5.** Correlations among sea surface temperature (SST) values for each island.  $SST_{\min}$ ,  $SST_{\text{mean}}$ ,  $SST_{\text{intra}}$ , and  $SST_{\text{inter}}$  represent the mean minimum, mean, intra-annual variance in mean, and interannual variance in mean SST of each site. Island values were derived by taking the mean of all site values for a given island. Diagonal panel shows distribution for each variable (y-axis ranges from 0 to 30), upper panel shows Spearman's rho, and lower panel shows linear correlations. Remote (black) and populated (red) islands from either survey period of the Hawaiian Archipelago (squares), Mariana Island Archipelago (circles), and American Samoa (triangle) are shown..... 84

**Supplemental Figure 6.** Correlations among island predictor variable. Island values were derived by taking the mean of site values for a given island, except for wave energy, distance to IPCT, and reef area where data resolution is at the island level. Diagonal panel shows distribution for each variable (y-axis ranges from 0 to 30), upper panel shows Spearman's rho, and lower panel shows linear correlations. Remote (black) and populated (red) islands from either survey period of the Hawaiian Archipelago (squares), Mariana Island Archipelago (circles), and American Samoa (triangle) are shown..... 85

**Supplemental Figure 7.** Island  $\beta$ -diversity values were calculated from regression of pairwise dissimilarity among sites with geographical distance among sites (in arbitrary units). This was calculated at each of the three spatial scales, within islands (A, D, G, J), within island groups (B, E, G, K), and across island groups (C, F, I, L) and is shown for Agrihan (A – C), Pagan (D – F), Guam (G – I), and Saipan (J – L). Vertical dashed line denotes distance at which island  $\beta$ -diversity was estimated.  $r^2$  and p-value of linear regression shown on plots. .... 87

## Acknowledgments

I thank Julia Baum, my supervisor, for providing me the opportunity to complete this research, of which I have grown more and more fond over the past two years. Thanks to Julia, I developed a deeper understanding of-, and appreciation for- marine ecology. Further, my scientific scepticism was refined and my awareness of the academic world was vastly broadened. By struggling through the conceptual and statistical nuances of null models and  $\beta$ -diversity, Julia helped me garner a stronger appreciation for the scientific discipline. The chance to conduct coral reef research first-hand in Kiritimati, Kiribati, as part of the Baum lab research team, was hugely appreciated.

Thank you Ivor Williams and Rusty Brainard for entrusting the extraordinary NOAA CRED data set to Julia and me. Further, thank you Ivor for helping to interpret the raw data as well as unexpected results. Thanks are in order for all the NOAA CRED scientific divers who collected the data, as well as the technicians and scientists who have compiled the data.

Thanks to my committee, Jana McPherson, Brian Starzomski, and Verena Tunnicliffe: your comment undoubtedly improved this thesis. Additionally, thank you Jana for helping to develop hypotheses relating to the explanatory variables that you know so well.

I thank Mark Vellend for helping me to create a complex null model, being a wealth of biodiversity knowledge, and pushing me towards using multivariate methodologies. Additionally, thank you Mark for hosting me in your lab at the University of Sherbrooke

in Quebec to work with you; I accomplished more in those three weeks than I might have in three months in Victoria. Belaid Moa, thank you for your patience and support implementing my null model on WestGrid servers. And to that end, thank you to Compute Canada for the resources; the expected computation time of my null model code running on my laptop (~70 years) vastly exceeds the standard length of an MSc, and was thankfully completed in a mere 72 hours on Breezy. Thank you Shawn Walbridge for your computational savvy in extracting many predictor variables for me. Eleanore Blaskovich, thank you for dealing with all the graduate students in the department and for providing me with my paycheque every month or so. Beth Rogers and Caroline Fox, thank you for the opportunity to TA at the Bamfield Marine Sciences Centre (BMSC) in the summer of 2013. And Brad Anholt, thank you for continuing to ensure BMSC runs smoothly amid constant budgetary cutbacks, the station is a wonderful place for us scientists.

The collective Baum and Juanes lab and associated members, thank you for many beers shared as well as statistical help and scientific banter. The lab has evolved tremendously over the past two years; I am intrigued to see the direction it will take in the next two and beyond. Thank you to all the intramural teams I played on that kept me from insanity and diabetes: The Dust, we were far better at crushing beers than crushing dingers; Hardwood Cedars, we got far too close to a championship to lose, but it sure was fun; and finally Fortrose Academy, while we ultimately were handed our first lost when it counted most, finishing 11-1 ain't terrible.

To my close friends in Victoria, Sherwood Park, and elsewhere, thank you for your support throughout. It has been a difficult time and there were many moments when I was ready to fly far away from it all... Ultimately I stuck around. I have your support to thank for that, and now I have a graduate degree to show for it.

Mom and Dad, you taught me love and respect of nature. Dad, it was your fondness and appreciation of wilderness that pushed me outdoors, towards being an ecologist. Mom, ultimately it was your love of, and connection to, the ocean that drew me to marine research. Thank you for your everlasting encouragement, support, and love.

**For Mom and Dad**

# 1. Introduction

Coral reefs are the most diverse marine ecosystems (Knowlton 2001). These ecosystems also are increasingly threatened by human mediated changes (Pandolfi et al. 2003). At the local scale, fishing, habitat destruction, and pollution directly alter community composition and biomass (Hughes 1994, Jackson et al. 2001, Friedlander and DeMartini 2002). At the global scale, climate change and the associated impacts of ocean acidification and bleaching impose chronic stresses on communities (McClanahan 2005, Pandolfi et al. 2005, Hoegh-Guldberg et al. 2007). Documented declines in coral cover on the order of 80% (Gardner et al. 2003), reductions in fish biomass exceeding 60% (Friedlander and DeMartini 2002), and massive declines in the ecological status of many coral reef fishes and corals from pristine to depleted, rare, or ecologically extinct (Pandolfi et al. 2003) have called for urgent change and remediation (Bellwood et al. 2004). Coral reef ecology has, to a large extent, become a crisis discipline, documenting humankind's rapid degradation of these ecosystems (Hughes 1994, Hughes et al. 2003, Pandolfi et al. 2005). The most visibly obvious and well-studied result of coral reef degradation is the phase shift from coral- to algal-dominated systems, characterized by huge reductions in coral cover and resulting increases in algal cover (Hughes 1994, Gardner et al. 2003, Rogers and Miller 2006). More recently, these changes have been intensified by coral bleaching: the expulsion of coral symbionts and resulting death of living coral tissue (Hoegh-Guldberg 1999). These human altered reefs tend to have

greater prevalence of coral diseases (Bruno et al. 2003, Sandin et al. 2008, Haapkylä et al. 2011), even further exacerbating reef degradation and ecosystem change (Harvell et al. 1999, Porter et al. 2001, Harvell 2002).

My thesis focuses on the effects of anthropogenic stressors on coral reef fishes, rather than the coral itself. Studies of human impacts on coral reef fishes have consistently demonstrated the loss or vast declines in the mean size and total biomass of large carnivores, including sharks, snappers and groupers (Polunin and Roberts 1993, Friedlander and DeMartini 2002, Graham and Evans 2003, Pandolfi et al. 2003, DeMartini et al. 2008, Nadon et al. 2012). Recent regional studies have also documented vast reductions in the biomass of large herbivorous fish (Friedlander and DeMartini 2002, Williams et al. 2010, Edwards et al. 2013). The direct removal of fish biomass and degradation of habitat has, in some cases, cascaded through reef communities, altering the abundance and species composition of other fish species or taxa (Dulvy et al. 2004a, 2004b, Stevenson et al. 2006). Despite significant changes in coral reef fish communities as a result of fishing and habitat degradation, there have been few cases of marine fish extinctions (Dulvy et al. 2003), and we know surprisingly little about humanity's impact on coral reef fish diversity.

Herein, I first summarize coral reef biodiversity and general beta- ( $\beta$ -) diversity research. I then briefly integrate these two topics, demonstrating the relevance of my primary research, which follows. My thesis research examines the effect of human disturbance on the spatial variability of coral reef fish community composition at three different spatial scales across the Pacific Ocean. In doing so, I am able to examine if humanity's effect is locally contained, only altering community composition and the variability therein at a small scale, or if effects either emerge or pervade across larger scales.

## **1.1 Coral Reef Biodiversity**

Coral reefs have a long history, with evidence of coral species dating back to the Ordovician (~450 Ma) and modern reef fish lineages dating back to at least the late Cretaceous (70 Ma) (Sale 2006). Despite covering less than 0.1% of the oceans' surface (Spalding and Grenfell 1997), coral reefs are inhabited by over 4000 species of fish, 800



species of corals, and 25% of all known marine taxa (McAllister 1995, Burke et al. 2011). The known species are only the tip of the iceberg, with estimates of biodiversity on coral reefs ranging from 1 to 9 million (Reaka-Kudla 1997), making reefs undoubtedly the most diverse marine ecosystem (Knowlton 2001).

Many coral reef fish biodiversity studies have focused on explaining how, or why, there are so many species of coral or reef fishes (Hutchinson 1959, Rocha and Bowen 2008), or how this high diversity is generated and subsequently maintained (Sale 1977). A great deal of the high diversity apparently is due to strong associations of reef fishes with the plethora of microhabitats present on coral reefs (Sale 1977, Syms and Jones 2000, Wilson 2001), however early reports of reef diversity found high numbers of fish on even small, 3 m diameter patches of reef (Smith 1973). Niche partitioning in diet and habitat was an early proximate hypothesis, allowing many species to coexist in small-scale equilibrium communities governed by Lotka-Volterra competition and predation dynamics (Volterra 1926, Sale 1977). However, this hypothesis was more-or-less discredited over forty years ago. Findings indicating that dietary overlap is common among species (Hiatt and Strasburg 1960, Jones and Helfrich 1967), even those with specialized feeding mechanisms (Bellwood et al. 2006a), indicated that food specialists on reefs are in actuality rare. Similarly, niche partitioning of habitat is not consistently strong and there is considerable overlap of habitats among species (Clarke 1977). The “niche-partitioning” hypothesis gave way to a “lottery-dynamic” view of coral reefs, where suitable reef habitat is both limiting and unpredictably available, with no one species having a consistent competitive advantage (Sale 1977). Less than a decade after the advent of the “lottery-dynamic hypothesis”, the “recruitment-limitation hypothesis” was established from findings of exceptionally low larval recruitment and high larval mortality. This hypothesis asserted that coral reef fish communities were, in effect, density-independent because juveniles and adult fish are unable to reach sizes where density-dependent factors become important (Doherty and Fowler 1994). Recent studies propose that post-settlement mortality in juveniles is highly important to fish recruitment and is, in actuality, often density dependent (Forrester 1995, Anderson et al. 2007, White et al. 2010). The hypotheses of lottery-dynamics along with some degree of recruitment-limitation and high post-settlement mortality have become more or less ingrained in coral

reef community ecology over the past thirty years (Hixon 2011). Recently, and in light of the neutral theory of ecology (Hubbell 2001), environmental stochasticity has been emphasized in its role structuring coral reef communities (Connolly et al. 2005, Dornelas et al. 2006).

Much of the of recent coral reef biodiversity research has focused on characterizing large-scale distributional patterns of coral reef species (Hughes et al. 2002, Connolly et al. 2005, Sanciangco et al. 2013), providing us with an ever-increasing awareness of the magnitude of, and factors that influence, diversity on coral reefs (Allen 2008). These studies build off the biogeographical patterns revealed by (Stehli and Wells 1971) that showed the highest coral species richness equatorially and in the Indo-Pacific Coral Triangle (IPCΔ). Recent studies support the IPCΔ as the highest biodiversity of coral reef species, although there is debate over the exact location of the “centre of the centre” of the biodiversity bulls-eye (Carpenter and Springer 2005, Allen 2008). Studies of the IPCΔ have often focused on resolving why the IPCΔ has such high biodiversity. There are four main competing hypothesis with varying levels of support: (1) the “area of overlap hypothesis” relating to the ranges of many species converging in the IPCΔ; (2) the “area of accumulation hypothesis” where ocean currents transport larvae and concentrate species in the IPCΔ; (3) the “area of refuge hypothesis” that posits that the IPCΔ has remained relatively invariant and habitable during periods of great geological and environmental change; and (4) the “centre of origin hypothesis” that suggests the IPCΔ is where coral reef species systems originally evolved and originated, with subsequent dispersal away from this area (see Rosen (1988), and Carpenter and Springer (2005)). From these studies, the importance of shallow-water habitat as a strong predictor of large-scale species richness has emerged (Sale and Douglas 1984, Bellwood and Hughes 2001, Sanciangco et al. 2013). However, different studies tend to favour different hypothesis regarding the IPCΔ. For instance, large range overlap of species in the IPCΔ, also termed the mid-domain effect (MDE), was implicated in large-scale species richness patterns and high richness in the IPCΔ (Hughes et al. 2002, Bellwood et al. 2005). Research from the same group, however, indicated that gyres in the Indian and Pacific Oceans could also be concentrating species in the IPCΔ, resulting in the strong species richness gradient (Connolly and Bellwood 2003). These findings have led to the

possibility that multiple factors implicated in the four main hypotheses work in unison to drive the observed species richness patterns (Wilson and Rosen 1998, Allen and Adrim 2003, Sanciangco et al. 2013).

Over the past decade there also has been increasing interest in studying the diversity of functional traits within coral reef communities (Bellwood et al. 2006a, Fox et al. 2009, Cadotte et al. 2011, Guillemot et al. 2011). Functional diversity studies aim to connect biological diversity with ecosystem functioning to better understand and manage natural processes. By examining the ecological traits of species (e.g. diet, feeding mode, gregariousness, size, etc.), and thereby predicting their ecological role within an ecosystem, groups of species have been identified as particularly important for the natural functioning of ecosystems. For instance, when a function is performed by few species (*low functional redundancy*), but is important for ecosystem functioning (e.g. algal removal), that ecosystem function could be highly susceptible to changes in abundances of few species. In contrast, a function that is performed by many species (*high functional redundancy*) is likely more resilient to changes in abundances of those species. While this framework highlights the immediate importance of subsets of species, even *seemingly* unimportant species are likely important (Nyström 2006). For example, *Platax pinnatus*, a batfish, performed little ecological function in coral dominated reef systems; however, after phase-shifts towards algal dominated reefs, *P. pinnatus* was the most significant herbivore responsible for phase-shift reversal (Bellwood et al. 2006b). Two highly important commonalities emerge from a number of functional diversity studies: many functions have low redundancy (Hoey and Bellwood 2009, Guillemot et al. 2011), and human disturbances can vastly alter and reduce functional diversity (Micheli and Halpern 2005, Pratchett et al. 2011, Martins et al. 2012). Functional diversity studies, however, rely on an assumed correlation between recorded functional group and actual functional role, a potentially fatal drawback. Fox et al. (2009) documented the fallacy of this assumption, as two closely related coral reef herbivores, *Siganus doliatus* and *S. lineatus*, had significant differences in diet, feeding rate, and feeding behaviour. Similarly, Mantyka and Bellwood (2007) found limited functional redundancy among a number of coral reef herbivores. Collectively, there is a call for the validation of functional groups, however such a task would be exceptionally time intensive and likely lead to more

specific functional classifications: a practice that pushes functional diversity studies ever nearer to traditional species diversity studies.

Aside from broad guild-based or functional group-based studies, studies of the effects of human disturbance on coral reef biodiversity are limited. One large-scale study of human disturbances illustrated a depression of the slope of the species-area relationship (Tittensor et al. 2007), however poor taxonomic resolution hampers the species-level interpretation of the results. A number of small-scale studies illustrate the potential effects of human disturbance on coral reef fish diversity in a variety of locations: loss of structural complexity associated with human disturbances in the Seychelles (Indian Ocean) led to decreased species richness and taxonomic distinctness (Graham et al. 2006); fishing pressure in Tuamotu Archipelago (Western Pacific) explained ~60% of the variability in species richness across atolls (Mellin et al. 2008); and finally, experimental habitat disturbance on the Great Barrier Reef resulted in lower fish abundance in addition to decreases in species richness (Syms and Jones 2000). Interestingly, these trends are not always observed. For instance, Jennings and Polunin (1997) found no correlation between removal of large piscivorous fish and species diversity of reef fishes in Fiji. In general, large-scale studies examining the effects of human disturbance on coral reef fish diversity across multiple geographical realms are lacking.

## 1.2 Beta Diversity

Local communities are populated by a subset of the species present at a larger regional scale. Whittaker (1960) proposed a theoretical and mathematical connection between these two scales over fifty years ago, defined as beta ( $\beta$ ) – diversity: “the extent of change of community composition.” Whittaker’s  $\beta$ -diversity formed a direct link between local ( $\alpha$ ) – and regional ( $\gamma$ ) – diversity, most simply defined either multiplicatively ( $\beta = \gamma / \alpha$ ) or additively ( $\beta = \gamma - \alpha$ ) (Whittaker 1960, 1972). Whittaker used this framework, alongside other metrics of community dissimilarity (i.e. Jaccard dissimilarity (Jaccard 1912) or percent (dis)similarity (Gleason 1920)) to calculate changes in floral community composition across environmental gradients (Whittaker 1960). Most simply,

$\beta$ -diversity measures the extent of differentiation among local communities with respect to the species present.

More technically,  $\beta$ -diversity can be divided into two types: turnover and variation (Vellend 2001, Anderson et al. 2011). Turnover measures change in community (i.e. species) composition across predefined spatial or environmental gradients, whereas variation measures variability in community composition among plots independent of gradients. Turnover studies generally ask at what rate does community composition change with respect to a directional gradient, whereas variation studies ask questions regarding the similarity (or dissimilarity) of communities within different sampling areas. Importantly, the two types of  $\beta$ -diversity require different metrics, of which there is a long list (Koleff et al. 2003).

$\beta$ -diversity research has seen a resurgence over the past decade (Anderson et al. 2011), likely owing to the fundamental community ecology underpinnings of  $\beta$ -diversity, together with the recent conservation implications of these studies (Condit et al. 2002, Legendre et al. 2005, Olden 2006).  $\beta$ -diversity provides ecologists insight into the spatial distributions of species and the processes that determine these patterns at various scales (Condit et al. 2002, Chase 2010, Dexter et al. 2012, Myers et al. 2012). These studies provide answers to questions such as: how dissimilar are the species that inhabit different communities within a given area, and to what extent is that spatial variability influenced by specific deterministic factors?  $\beta$ -diversity studies have been used as evidence of processes structuring communities (Hewitt et al. 2005), and as such, these studies have become a hotbed for contrasting views on the neutral theory of ecology (Condit et al. 2002, Dornelas et al. 2006, Dexter et al. 2012). Similarly, many studies have contrasted the relative importance of a suite of deterministic (i.e. environmental) and stochastic (random) processes under different conditions, such as drought (Chase 2007), agriculture (Vellend et al. 2007), or predation (Chase et al. 2009). In general,  $\beta$ -diversity is predicted to increase due to any process that:

- 1) makes areas differentiated with respect to characteristics (i.e. environmental) that drive community composition, or
- 2) increases the randomness with which communities are assembled.

Recently, a wide range of studies have found that human disturbances also frequently alter  $\beta$ -diversity (McKinney and Lockwood 1999).

Biotic homogenization, the decrease in  $\beta$ -diversity over time, is primarily driven by invasions of non-native species and extinctions of rare (i.e. low abundance) species, and has become a well-documented result of human impacts (McKinney and Lockwood 1999, Olden 2006). Olden and Poff (2003) created a conceptual model of fourteen scenarios of species invasions or extinctions that could lead to biotic homogenization or differentiation (increase in  $\beta$ -diversity). Most generally, the degree to which invasions and extinctions are shared among communities is what drives either homogenization or differentiation: shared extinctions or invasions can result in homogenization whereas unshared extinctions or invasions can result in differentiation. Beyond invasions and extinctions, conditions that result in local ecological filters for species (i.e. stressful conditions such as extreme drought) can also result in homogenization (decreased  $\beta$ -diversity) (Chase 2007, Chase and Myers 2011). The consequences of biotic homogenization are extensive, ranging from reductions in species diversity to reductions in ecosystem functioning, stability, and resistance to environmental change on both ecological and evolutionary time scales (Olden et al. 2004). There have been many studies documenting the biotic homogenization of plant and freshwater fish communities (Rahel 2000, Smart et al. 2006, Vellend et al. 2007, Olden et al. 2007), but relatively few documenting biotic homogenization of other fauna (e.g. Olden et al. (2006), Donohue et al. (2009), Burman et al. (2012), and Karp et al. (2012)), and even fewer documenting biotic differentiation (e.g. Taylor (2004), Marchetti et al. (2006), Cassey et al. (2007), Villegger et al. (2011)).

Those studies that did find differentiation were often either relatively small-scale studies (Marchetti et al. 2006), or studies that covered a range of scales and also found homogenization at large scales (Taylor 2004, Cassey et al. 2007, Villegger et al. 2011). The problem of scale is not a new one to ecology (Levin 1992), but is one that  $\beta$ -diversity studies are well suited to address.  $\beta$ -diversity studies typically require definition of regional boundaries that comprise  $\gamma$ -diversity, as well as repeated sampling within those bounds. As such,  $\beta$ -diversity can be calculated at different scales by partitioning  $\beta$ -

diversity into many scales within those regional bounds (Crist et al. 2003, Belmaker et al. 2008, Rodríguez-Zaragoza et al. 2011), or by redefining the spatial extent of regional boundaries during analysis (Cornell et al. 2007, Karp et al. 2012). Subsequent modelling of  $\beta$ -diversity at different scales can help address the scale that different variables influence community composition (Karlson et al. 2007).

Recent methodological developments in  $\beta$ -diversity have focused on multivariate methods to assess  $\beta$ -diversity (e.g. Legendre et al. (2005, Anderson et al. (2006)), as well as addressing and reducing biases inherent in  $\beta$ -diversity metrics using null models (Chase and Myers 2011, Kraft et al. 2011). Null model methodology has been used in ecology for nearly a century, dating as far back as the 1920s (Maillefer 1929) and becoming more prominently used in the 1970s and 1980s for application to species co-occurrence patterns (Connor and Simberloff 1979, Diamond and Gilpin 1982, Gotelli and Graves 1996). Recent advances in computing power have further increased their accessibility and breadth of use. When applying null models to  $\beta$ -diversity studies, one generally asks “how similar are two communities in reference to two randomly assembled communities from a given regional pool of species.” Using this methodology, it is possible to constrain one or more aspects of the random communities, thereby accounting for those effects on  $\beta$  diversity metrics. For instance, recent studies have constrained and removed the effect of  $\alpha$ -diversity (Vellend et al. 2007) or  $\gamma$ -diversity (Kraft et al. 2011) on  $\beta$  diversity. Without the use of null models,  $\beta$ -diversity is negatively related to  $\alpha$ -diversity and positively related to  $\gamma$ -diversity. As a result, a number of patterns have been observed that are simply artefacts of this relationship (Kraft et al. 2011).

### **1.3 Coral Reef $\beta$ -diversity**

The resurgence of  $\beta$ -diversity studies has lagged for coral reef ecosystems, and of the few studies that have explored coral reef  $\beta$ -diversity, there is limited cohesiveness among their themes. Extrapolating from these studies, however, three generalities emerge:

1.  $\beta$ -diversity of coral species is correlated with environmental variables (i.e. live coral cover or depth) (Harborne et al. 2006, Becking et al. 2006, Arias-González et al. 2008),
2.  $\beta$ -diversity of coral species, in addition to variance in other habitat variables, can drive  $\beta$ -diversity of coral reef fishes (Arias-González et al. 2008, Belmaker et al. 2008, Rodríguez-Zaragoza 2008, Macneil et al. 2009, Acosta-González et al. 2013), and
3. scale has a strong effect on the relationship between  $\alpha$ -,  $\beta$ -, and  $\gamma$ -diversity of corals and reef fish (Cornell et al. 2007, Belmaker et al. 2008, Macneil et al. 2009, Francisco-Ramos and Arias-González 2013).

Specifically, Cornell et al. (2007) found that  $\beta$ -diversity of corals decreases as the spatial scale of  $\alpha$ - and  $\gamma$ -diversity increases, indicating that local richness increases at a greater rate than regional richness. Using randomizations, Cornell et al. (2007) attributed a portion of the  $\beta$ -diversity to non-random clustering of species. Macneil et al. (2009) and Francisco-Ramos and Arias-González (2013) found similar results for reef fishes, however Macneil et al. (2009) attributed variability in  $\beta$ -diversity to habitat variability rather than non-random clustering of species. Belmaker et al. (2008) used null models to similarly attribute the  $\beta$ -diversity of coral reef fishes to habitat variability and more specifically, species-specific coral associations.

A number of studies in the Caribbean basin have documented the environmental variables that best correlate with coral reef fish  $\beta$ -diversity. Arias-González et al. (2008) found  $\beta$ -diversity of reef fishes to be less spatially structured than corals and to be driven by a number of environmental variables, including: topographical complexity (i.e. rugosity) depth, live coral cover, soft coral cover, calcareous substrate cover, algal cover, sea grass cover, and rubble cover. Francisco-Ramos and Arias-González (2013) similarly found environmental heterogeneity to be of top importance for maintaining reef fish  $\beta$ -diversity. In addition to attributing across-transect  $\beta$ -diversity to differences in habitat structure, Acosta-González et al. (2013) documented decreases in  $\beta$ -diversity concomitantly with decreases in  $\alpha$ - and  $\gamma$ -diversity over a decade of reef degradation. These decreases were primarily a result of the loss of rare species. While these studies



collectively have highlighted the importance of habitat and environmental variables in maintaining  $\beta$ -diversity, in addition to the importance of considering scale as strongly influencing  $\beta$ -diversity, the spatial scale of human disturbance on coral reef fish  $\beta$ -diversity has not been explored, nor have these results been tested over large geographical ranges.

## 1.4 Objectives and Hypothesis

By modelling  $\beta$ -diversity at three spatial scales across the Pacific Ocean, ranging from island-scale to cross Pacific-scale, I will examine the spatial extent and magnitude of effect of human disturbances, as well as environmental variables, on the spatial variability of coral reef fishes. In this study area, artisanal and recreational fishing is expected to be the greatest aspect of human disturbances influencing the community composition of coral reef fishes. Comparing across three spatial scales will serve two purposes:

- 1) To determine if human disturbances alter  $\beta$ -diversity pervasively across all scales, or only at small, intermediate, or large scales; and
- 2) To determine the scale at which environmental variables are most likely to influence community composition.

Using multivariate techniques I will extend my analysis to examine if human disturbances alter the relationship between community composition and environmental variables. In this thesis I test the effect of a number of environmental variables on  $\beta$ -diversity at all three scales (Table 1). These variables have been shown to influence biodiversity patterns at various scales, as detailed below.

The effect of human disturbances on  $\beta$ -diversity has been frequently documented; however, studies have documented both increases and decreases in  $\beta$ -diversity due to human disturbance. In the biotic homogenization literature, human disturbances most often decrease  $\beta$ -diversity as a result of invasions (Olden and Poff 2003). While invasions of coral reefs are not insignificant (Coles et al. 1999), I expect that the direct removal of fishes as a result of fishing and habitat destruction has a larger impact on  $\beta$ -diversity.

**Table 1.** Predicted univariate effect of each explanatory variable on  $\beta$ -diversity at different scales. Scale refers to  $\beta$ -diversity calculated within islands (1), within island groups (2), or across island groups (3). ‘All’ denotes that the hypothesis applies to all three scales. H.A., M.A., and A.S. = Hawaiian Archipelago, Mariana Archipelago, and American Samoa respectively. IPC $\Delta$  = Indo-Pacific Coral Triangle; SST<sub>mean</sub> = mean sea surface temperature; SST<sub>var</sub> = interannual variability in sea surface temperature.

Group	Variable	Scale:	$\beta$ greatest at:	Hypothesis
Anthropogenic	Human Disturbance	All		Depending on the nature of anthropogenic disturbances, islands with high disturbance could have higher <i>or</i> lower $\beta$ -diversity. If anthropogenic disturbances and fishing results in the removal of rare and spatially variable species, I predict $\beta$ -diversity will be lower on island with high human disturbance <sup>1</sup> . Concomitant increase in invasive species on islands with high human disturbance would exacerbate the decrease in $\beta$ -diversity on these islands <sup>1</sup> . On the other hand, if fishing targets common fishable species, I predict $\beta$ -diversity to be greater on islands with high human disturbance. I predict that islands with high human disturbance will be less well predicted by explanatory variables (i.e. breakdown of species – environment relationship) <sup>2</sup> .
Island Biogeography	Distance between sites	All	Distant sites	As the distance among sites increases, connectivity will decrease and the effect of ecological drift will be increased <sup>3</sup> . As such, $\beta$ -diversity will be greatest among distant sites.
	Distance to IPC $\Delta$	3	H.A.	The smaller realized regional pool of the H.A. will seem more differentiated from the null community than sites with larger proportions of the regional pool present (i.e. M.A. or A.S.).
	Reef Area	1	Greater reef area	On islands with less reef area there is predicted to be lower habitat heterogeneity <sup>4</sup> and smaller species pool <sup>5</sup> . As such, small islands with low reef area are predicted to have lower within island $\beta$ -diversity.
		2 & 3	Lower reef area	Islands with less reef area are more likely to have different communities than islands with large reef area due to ecological drift <sup>2,5</sup> , and would therefore appear less similar to other islands, manifest in greater across island $\beta$ -diversity.
Habitat Heterogeneity	Live Coral Complexity	All	High	At sites with low habitat complexity only a subset of species that would otherwise be present will persist, therefore manifesting in low $\beta$ -diversity. Sites with high complexity have the potential for many different species to live in the different habitats, potentially leading to greater stochasticity in site occupancy of species, and therein greater $\beta$ -diversity.
	Island composition	All	Basalt, Carbonate, Bas/Carb	Habitat heterogeneity is predicted to be greater on basalt, carbonate, or basalt/carbonate islands due to rocky island characteristics not present on atolls (headlands, freshwater outflows, etc.), leading to greater variability in habitat, and therefore in species occurrences across sites.
	Wave Exposure	All	High Wave Exposure	Islands with high variability in wave energy likely have more spatial heterogeneity in wave action (i.e. sheltered sides of headlands, bays, etc.), leading to different species compositions in those areas <sup>6</sup> . As such, $\beta$ -diversity within and across islands is expected to be greater.
Energy	SST <sub>mean</sub>	All	High	Tropical species’ distributions are more likely to be constrained by tolerance of cold rather than warm waters, therefore islands with higher mean SST will have a larger pool of physiologically adapted species and so higher beta diversity <sup>7</sup> .
	SST <sub>var</sub>	All	Low	Islands with low interannual variability in SST are more likely to have a broad range of species that are acclimated to the mean SST and need not adapt for large changes in SST throughout a year that could be physiologically stressful. As such, islands with low SST <sub>var</sub> are expected to have greater species pools leading to greater stochasticity in species distributions and therein greater $\beta$ -diversity.

*Notes.* [1] (Olden and Poff 2003); [2] (Vellend et al. 2007); [3] (Vellend 2010); [4] (Kolasa et al. 2011); [5] (MacArthur and Wilson 1967); [6] (Fulton et al. 2005); [7] (Tittensor et al. 2010)

Human disturbances indeed reduce the abundance of large fishes, however whether disturbances uniformly remove other species throughout large areas will determine whether human disturbances are associated with increases or decreases in  $\beta$ -diversity. If human disturbances result in the introduction of widespread non-native species as well as the reduction of rare species (i.e. species with low abundances) across large areas, similar to that documented by Acosta-González et al. (2013) at smaller scales, I predict that  $\beta$ -diversity will decrease. However if human disturbances (i.e. fishing) target common, easily harvestable species,  $\beta$ -diversity may actually increase as a result of the loss of previously shared species across large geographic areas. Importantly, the effect of human disturbances is likely dependent on the scale of analysis as previously documented (Cassey et al. 2007), and if homogenization is observed, it will likely be most evident at larger scales of study. Regardless of whether human disturbances tend to increase or decrease  $\beta$ -diversity, I hypothesize that our ability to predict species assemblages and therein attribute variability of those assemblages to environmental variables in human disturbed areas will decrease, *sensu* Vellend et al. (2007). This would emerge as a result of similar communities of fishes being present at sites with high human disturbance, regardless of differences in environmental conditions amongst sites (i.e. fish communities not reflecting environmental conditions at sites with high human disturbance).

The geographical distance among areas, distance from the Indo-Pacific Coral Triangle (IPC $\Delta$ ), and reef area could all influence  $\beta$ -diversity, and relate to theories of island biogeography (Table 1). Distance among sites is consistently a strong predictor of  $\beta$ -diversity in studies of plants and animals (Condit et al. 2002, Dexter et al. 2012), as greater distance among sites:

- 1) decreases ecological connectivity resulting in dispersal limitations (MacArthur and Wilson 1967), and
- 2) increases the likelihood that sites differ in environmental characteristics (Karp et al. 2012).

Therefore, as distance among sites increases, I predict that sites will differ more in their species composition, manifest in greater  $\beta$ -diversity among sites.

The Indo-Pacific Coral Triangle (IPC $\Delta$ ) is known for the greatest diversity of corals and coral reef fishes in the world, and species richness of both taxa decreases with increasing distance from the IPC $\Delta$  (Mora et al. 2003, Connolly et al. 2005). Distance to the IPC $\Delta$  would likely influence many metrics of  $\beta$ -diversity, as sites far from the IPC $\Delta$  (i.e. the Hawaiian Archipelago in the Eastern Pacific) are relatively depauperate (Mora et al. 2003), and therefore those sites would appear to have greater  $\beta$ -diversity. Null model methods (described in detail in section 2.3.2) account for differences in  $\alpha$ -diversity and therefore the effect of distance to the IPC $\Delta$  would be a result of differential regional diversity on  $\beta$ -diversity. In that case, depauperate areas further from the IPC $\Delta$  (i.e. the Hawaiian Archipelago) with lower regional diversity might be expected to have lower  $\beta$ -diversity, as  $\beta$ -diversity often correlates with  $\gamma$ -diversity (Cornell et al. 2007).

Islands with greater reef area are expected to have both a larger pool of species present and increased habitat heterogeneity, thereby increasing the potential for sites to be different to one another in their species composition (MacArthur and Wilson 1967), manifest in greater  $\beta$ -diversity. As such, sites on islands with a small reef area might similarly be expected to have a smaller species pool, less habitat heterogeneity, and lower  $\beta$ -diversity. Islands with small reef area and limited connectivity, however, could experience significant ecologically drift relative to one another (MacArthur and Wilson 1967). Therefore, if comparing sites on the same island, I predict  $\beta$ -diversity will be greater on larger islands, however if comparing sites on different islands, I predict  $\beta$  diversity will be greatest between islands with small reef area (Table 1).

In regards to habitat, I test the influence of five different habitat heterogeneity variables: live coral cover, habitat rugosity, wave exposure, reef area, and geological island composition (Table 1). All five variables reflect one common hypothesis: many species have strong preferences for unique habitats, and are more or less likely to be found at sites embodying different environmental characteristics that meet their niche requirements (Hutchinson 1957). Similarly, many coral reef fishes closely associate with their benthic habitat (Syms and Jones 2000, Wilson 2001), and changes to the benthic composition can result in subsequent changes to fish species composition and abundance (Wilson et al. 2006, 2009): live coral cover and habitat complexity (rugosity) are two

such habitat characteristics that the abundance or biomass of many reef fishes often correlate with (Wilson et al. 2008, Cinner et al. 2009). I predict that sites with high living coral cover and habitat complexity will satisfy more species niches, and therefore more species will be able to occupy those sites. Sites with low live coral cover or low habitat complexity will act as an environmental filter and the opposite will be true, whereby only a subset of species would have their ecological niche satisfied. These sites with low live coral cover or habitat complexity are in turn predicted to have lower  $\beta$  diversity (Table 1).

Island composition could influence  $\beta$ -diversity due to its effect on a number of otherwise unmeasured habitat and environmental characteristics (Table 1). In short, I predict that atolls will have less habitat heterogeneity than basalt or carbonate islands that have the potential for large headlands, bays, complex shoreline features, as well as nutrient enrichment and freshwater input. Furthermore, I predict closed atolls have less potential for habitat heterogeneity than open atolls, as open atolls allow for species migration and water flow between the lagoon and outer reef. As such, I predict that  $\beta$ -diversity on islands of different composition will be as follows: basalt/carbonate islands  $\approx$  basalt islands  $\approx$  carbonate islands  $>$  open atolls  $>$  closed atolls.

Wave energy has a strong ecological forcing effect on reefs, influencing both coral and fish communities (Dollar 1982, Storlazzi et al. 2004, Fulton and Bellwood 2005, Fulton et al. 2005). As documented for evidence of the intermediate disturbance hypothesis, high wave energy can destroy a large proportion of coral species and revert communities to early successional stages, whereas low wave energy can result in competitive exclusion of coral species (Connell 1978, Dollar 1982, Storlazzi et al. 2004). Not only would areas with different coral composition (i.e. in different successional stages) differ in the fish species that associate with those corals, wave energy directly influences the type of fishes that live in areas due to constraints on swimming mode. For instance, in areas with high water flow (i.e. high wave energy) pectoral-swimming fishes predominate, whereas pectoral-caudal- and caudal-swimming fishes predominate in areas with low water flow (Fulton and Bellwood 2005). I therefore predict that islands with high variation in wave energy will harbour the greatest variation in species composition, manifest in greater  $\beta$ -diversity (Table 1).

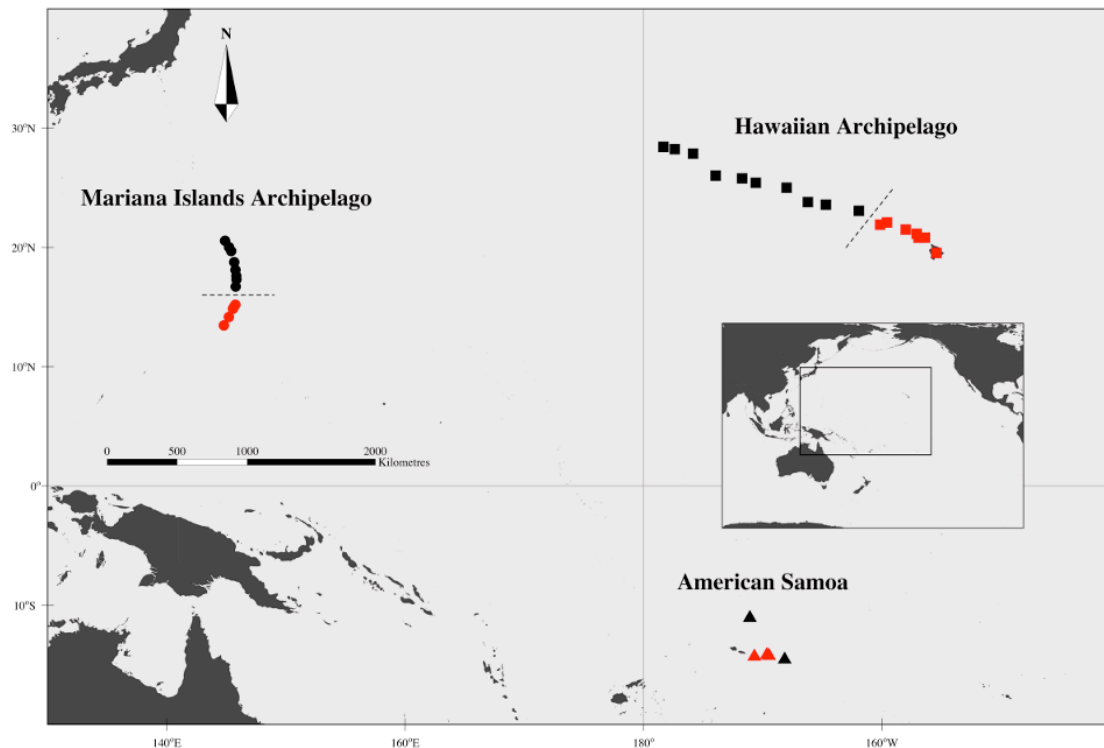
Mean sea surface temperature (SST) and primary productivity correlate strongly with regional species richness of various marine taxa, with warmer or more productive waters tending to have higher species richness (Tittensor et al. 2010). As such, areas with warmer SST (Table 1) or higher primary productivity are predicted to have a greater regional pool of species that could potentially inhabit them and an increased opportunity for stochastic variability in species composition among sites, manifest in greater  $\beta$ -diversity. Similarly, I predict that  $\beta$ -diversity will be greater in areas with higher mean minimum SST because relaxed lower thermal boundaries in areas with high minimum SST would allow a larger regional pool of species to persist (Figueira and Booth 2010), again allowing for greater stochasticity in species composition. Finally, I predict that  $\beta$ -diversity will be greater in areas with less variable SST, both intra- and inter-annually (Table 1), as areas with stable SST have fewer environmentally stressful periods that could constrain community composition. Because both mean SST and primary productivity are hypothesized to similarly effect  $\beta$ -diversity and in this dataset are highly correlated, primary productivity was not modelled to reduce the number of variable modelled.

The results of this analysis will provide insight into the relative impact of human disturbance on  $\beta$ -diversity of coral reef fishes at both large and small spatial scales. Additionally, from this research the spatial scale that environmental variables most strongly influence community composition will be revealed.

## 2. Methods

### 2.1 Study Region

This study focuses on three island groups in the Pacific Ocean, the Mariana Archipelago (MA), the Hawaiian Archipelago (HA), and American Samoa (AS)(Figure 1). Together, these island groups span over 6500 km of longitude and over 4500 km of latitude, from 15°S - 29°N, 154°E - 144°W. Thirteen islands and atolls (henceforth inclusively referred to as islands) were surveyed in the Mariana Archipelago, seventeen in the Hawaiian Archipelago, and five in American Samoa. The islands in these island groups vary greatly in size and composition, from large basalt islands such as Hawaii to fully submerged coral atolls such as Maro Reef (Table 2). The islands also range in their human inhabitancy, from heavily populated islands such as Guam and Oahu to islands devoid of permanent human settlements such as Pagan and French Frigate (Table 2). Between 2009 and 2012, the Coral Reef Ecosystem Division (CRED) of the U.S. National Oceanographic and Atmospheric Administration's (NOAA) Pacific Island Fisheries Science Center (PIFSC; <http://www.pifsc.noaa.gov/>) conducted underwater visual censuses on the forereef of each island as a part of their Pacific Reef Assessment and Monitoring Program (Pacific RAMP). The sampling design is spatially hierarchical: the Pacific Ocean is subset into three main island groups, which are further subset into islands, which are further subset into survey sites. This hierarchical nature allowed me to



**Figure 1.** Islands surveyed by NOAA CRED Pacific RAMP between 2009 and 2012 in the Pacific Ocean. Black and red symbols indicate islands with low or high human disturbance respectively in the Hawaiian Archipelago (squares), Mariana Archipelago (circles) and American Samoa (triangles). Scale bar represents distances at equator.

examine how anthropogenic disturbances influence  $\beta$ -diversity at three spatial scales: across island groups; within island groups; and within islands (Figure 1).

Surveys were completed in two two-year blocks, with between 2 and 133 surveys performed on each island per survey season (Table 2) – roughly proportional to the reef area of each island. All islands of the Mariana Archipelago were surveyed in 2009 and 2011, and all islands of American Samoa were surveyed in 2010 and 2012. Surveys of the Hawaiian Archipelago were done over two years (2009-2010, and 2011-2012), with some islands surveyed twice within a two year period. For those islands, data from the year with the most surveys for a given island were used for analysis (Table 2).



**Table 2.** Island-scale information for all islands surveyed. Surveys not included in the analyses herein indicated by strike through text.

Island Group	Island	Surveys per Year				Lat.	Lon.	Island Composition <sup>1</sup>	Land Area <sup>1</sup> (Km <sup>2</sup> )	Reef Area <sup>2</sup> (Km <sup>2</sup> )	Human	
		09	10	11	12						Population <sup>3</sup>	Disturbance
Hawaiian Archipelago	Hawaii		43			19.53	-155.42	Basalt island	10441.51	193.7	185,079	High
	Kauai		26			22.09	-159.57	Basalt/Carbonate island	1436.7	178.8	66,931	High
	Lanai		16		29	20.82	-156.92	Basalt island	365.37	46.3	3,102	High
	Maui		33		49	20.82	-156.4	Basalt island	1886.32	164.6	144,475	High
	Molokai		10		50	21.14	-157.09	Basalt/Carbonate island	670.22	161.6	7,347	High
	Niihau & Lehua		16			21.9	-160.15	Basalt/Carbonate island	186.82	86.7	160	High
	Oahu		40		35	21.49	-158	Basalt/Carbonate island	1548.99	374.8	953,207	High
	French Frigate		9	<del>2</del>	5	23.79	-166.21	Open atoll	0.2	469.4	0	Low
	Gardner Pinnacles				12	25.00	-168.00	Pinnacle	0.2	0.7	0	Low
	Kure	29	<del>46</del>		14	28.42	-178.33	Closed atoll	0.92	90.2	0	Low
	Laysan	14		23		25.78	-171.73	Carbonate island	3.53	26.4	0	Low
	Lisianski	<del>19</del>	25	<del>9</del>	25	26.01	-173.95	Open atoll	1.5	215.6	0	Low
	Maro	31		21		25.41	-170.58	Open atoll	0	217.5	0	Low
	Midway	31		17		28.23	-177.38	Closed atoll	5.98	85.4	60 <sup>4</sup>	Low
	Necker	13		8		23.58	-164.7	Basalt island	0.12	9.1	0	Low
Nihoa			8		23.06	-161.93	Basalt island	0.72	5.6	0	Low	
Pearl & Hermes		24	<del>9</del>	15	27.86	-175.85	Closed atoll	0.5	374.5	0	Low	
Mariana Archipelago	Aguijan	6		13		14.85	145.55	Basalt/Carbonate island	7.01	2.6	0	High
	Guam	25		133		13.46	144.79	Basalt/Carbonate island	544.34	91.3	159,358	High
	Rota	14		24		14.16	145.21	Basalt/Carbonate island	85.13	12.1	2,527	High
	Saipan	22		30		15.19	145.75	Basalt/Carbonate island	118.98	56.8	48,220	High
	Tinian	14		19		14.99	145.63	Basalt/Carbonate island	101.21	14.7	3,136	High
	Agrihan	14		20		18.76	145.66	Basalt island	44.05	8.6	0	Low
	Alamagan	6		5		17.6	145.83	Basalt island	12.96	3.2	0	Low
	Asuncion	13		20		19.69	145.4	Basalt island	7.86	0.5	0	Low
	Farallon de Pajaros	7		12		20.55	144.89	Basalt island	2.25	0.8	0	Low
	Guguan	6		10		17.31	145.84	Basalt island	4.24	1.1	0	Low
	Maug	21		30		20.02	145.22	Basalt island	2.14	2.1	0	Low
Pagan	20		29		18.11	145.76	Basalt island	47.75	11.1	0	Low	
Sarigan	7		9		16.71	145.78	Basalt island	4.47	1.9	0	Low	
American Samoa	Ofu & Olosega		30		30	-14.17	-169.65	Basalt island	12.61	3.6	353	High
	Tau		24		22	-14.24	-169.47	Basalt island	45.09	3.8	790	High
	Tutuila		105		85	-14.3	-170.7	Basalt island	137.45	35.8	54,359	High
	Rose		24		33	-14.55	-168.16	Closed atoll	0.09	7.9	0	Low
	Swains		24		38	-11.06	-171.08	Carbonate island	2.38	2.4	17	Low

Notes. [1] Island composition and land area from Gove et al. (2013); [2] reef area from shoreline to 10 fathom line (Rohmann et al. 2005); [3] human population data from US Census 2010 (<http://www.census.gov/2010census/>), except [4] Midway where population is ~60 fulltime US Fish and Wildlife Service staff.

## **2.2 Data**

### **2.2.1 Underwater Visual Censuses**

Fish community composition was estimated from stationary point count (SPC) surveys performed by CRED divers between 2009 and 2012. These surveys were conducted by a pair of divers simultaneously monitoring 15 m diameter cylindrical plots, extending from the substrate to the limits of vertical visibility. Each survey consisted of an initial five-minute period, during which divers recorded all species present within their 15m diameter cylinder, followed by a systematic count of the number of each fish on their species list, during which time divers remained mostly stationary and performed a series of visual sweeps of the cylinder for each species. At the end of the tallying period, each diver also swam through their plot area to record estimates of small and cryptic species that tend to be underrepresented in the stationary portion. Together, these survey methods produce estimates of species density of all fish over 5 cm at each site. Any individuals that were not identified to the species level were removed prior to analysis, yielding the removal of 14280 individuals or 2.3% of the 625560 total individuals enumerated. Data from the two cylinders at a site were summed to form an estimate of community composition at that site. Sites where 1, 3, or 4 cylinders were surveyed were removed prior to analysis. At a subset of sites, surveys of cylinders were repeated a second time. In these cases, only the first surveys were used for analysis. Fish abundance data were analyzed both in their raw abundance form as well as in converted incidence (i.e. presence / absence) form. Following fish surveys, divers performed benthic habitat surveys, detailed below (section 2.2.2). Full details of fish and benthic surveys are available in (Williams et al. 2010, Ayotte et al. 2011).

### **2.2.2 Explanatory Variables – Environmental**

I tested the influence of a suite of biophysical variables, obtained from open source datasets, literature values, and benthic habitat surveys performed by CRED scientists during underwater visual censuses, on  $\beta$ -diversity and community composition (Table 3). Each variable relates to a specific hypothesis about its influence on  $\beta$ -diversity as detailed in the Introduction (Table 1).

Distance between sites (in arbitrary pixel units) was calculated as the minimum path distance between sites that does not intersect land. Calculations were done using the R package “gdistance” at a 100m resolution for the within island scale and 1km resolution for the within island group scale. Geodesic distance was calculated between sites across island groups, as well as between sites and the Raja Ampat (the center of the IPC $\Delta$  sensu Allen (2008)), calculated using the R package “fields”. Reef area was measured as the area between island shoreline and the 10 fathom line, derived from charts by Rohmann et al. (2005). Because of correlation with island group, distance to IPC $\Delta$  was not used in final modelling.

**Table 3.** Data sources and resolution of the explanatory variables. IPC $\Delta$  = Indo-Pacific Coral Triangle.

<b>Data Type</b>	<b>Variable</b>	<b>Data resolution</b>	<b>Data source</b>
Anthropogenic	Human Disturbance	Island	SEDAC <sup>1</sup> / US Census 2010 <sup>2</sup>
Biogeographic	Distance among sites	Site	CRED survey coordinates
	Reef Area	Island	Rohmann et al. 2005
	Distance to IPC $\Delta$	Island	CRED survey coordinates
Environmental	Wave Exposure	Island	Gove et al. 2013
	Sea Surface Temperature	Site	Baum and McPherson <sup>3</sup>
	Live Coral Cover	Site	CRED benthic survey
	Habitat Complexity	Site	CRED benthic survey
	Island Type	Island	CRED benthic survey

*Notes.* [1] Socioeconomic Data and Applications Center (<http://sedac.ciesin.columbia.edu/>); [2] <http://www.census.gov/2010census/>; [3] unpublished spatial data product.

Live coral cover and habitat rugosity (vertical relief) were estimated during benthic surveys following fish surveys. Live coral cover was visually estimated by divers as a percentile, and reflects both hard and soft corals. For surveys in 2010 and 2011, habitat rugosity was scored by divers by estimating the percentage of their survey cylinder that falls within six vertical relief categories: low (< 0.25 m), medium-low (0.25 – 0.75 m), medium (0.75 – 1.5 m), medium-high (1.5 – 3 m), high (3 – 5 m), and very high (> 5 m). These percentages are then converted into a single discrete value between one to six, with each integer split into four bins (i.e. 1.0, 1.25, 1.5, 1.75, 2.0, ... , 5.75, 6.0) (Ayotte et al. 2011). In 2012 rugosity measurements were changed to reflect the mean height of substrate within a cylinder (Brandt et al. 2009); these values are not directly comparable

to the 2010-2011 data. To facilitate comparison, I calculated percentiles for each vertical relief categorical bins for American Samoa and the Hawaiian Archipelago in 2010. For example, in 2010 in the Hawaiian Archipelago, sites with a vertical relief value of 1.0 represent the lowest 15.4 percentile. The 2012 vertical relief measurements were ranked for American Samoa and the Hawaiian Archipelago and each site was converted to the 1 – 6 scale based on the percentile rank that it fell into. That is, the site that ranked as the lowest in the Hawaiian Archipelago out of 229 sites would fall into the lowest 15.4 percentile and therefore would receive a complexity value of 1.0. For one aspect of the analysis, live coral cover and habitat rugosity were multiplied into a composite variable, living coral complexity, to reduce the number of variables tested. The two values are slightly correlated (Spearman's  $\rho = 0.30$ ), but in general measure different characteristics of the benthos.

Islands composition was classified by Gove et al. (2013) into five different categories: basalt island, carbonate island, basalt/carbonate island, open atoll, and closed atoll. Gardner Pinnacles were not classified by Gove et al. (2013), but were classified as a closed atoll as this classification most closely reflects the small size of these pinnacles. I predict that islands of different composition will have different levels of  $\beta$ -diversity, generally following the trend: basalt or carbonate islands > open atolls, > closed atolls (Table 1). However, due to covariance of composition with human inhabitancy in Mariana Archipelago and American Samoa, and erratic model behaviour when included as a predictor variable, geological composition was not used in final model selection.

Site-specific wave energy measurements were not available. Instead, I used island level temporal variability in wave energy (measured in  $\text{kW m}^{-1}$ ) as a best-available estimate of site-site variability in wave energy among sites. Standard deviations of wave energy, modelled every 3 hours from wave period and mean significant wave height at a  $1^\circ$  spatial resolution, were calculated from 14 years of data for each island by Gove et al. (2013). Wave energy was not available for Gardner Pinnacles, therefore it was taken to be intermediate of the two most proximate islands, Maro and French Frigate. Because of the linear relationships between offshore wave energy and wave-induced currents on reefs, offshore wave energy has been used as a surrogate for wave forcing on reefs where more precise data are not available (Hearn 1999, Williams et al. 2013). While temporal

and spatial variability in wave energy are not synonymous, areas with high temporal variability are predicted to also have high spatial variability (and vice versa for low temporal variability). Since spatial variability in wave energy is not available, temporal variability is used as a surrogate.

Sea surface temperature (SST), including mean SST ( $SST_{\text{mean}}$  or  $SST_{\text{m}}$ ), mean minimum SST ( $SST_{\text{min}}$ ), mean intra-annual SST variability ( $SST_{\text{intra}}$ ), and mean inter-annual SST variability ( $SST_{\text{inter}}$ ,  $SST_{\text{var}}$ , or  $SST_{\text{v}}$ ) were obtained from Baum & McPherson (unpublished spatial data product). These variables are based upon weekly means between derived from the Coral Reef Temperature Anomaly Database (CoRTAD), a product of AVHRR Pathfinder data (Selig et al. 2010, Casey et al. 2010). Correlations between mean island SST values were plotted and Spearman's rho was calculated to determine if variables were strongly correlated (Supplemental Figure 5). Because  $SST_{\text{mean}}$  and  $SST_{\text{min}}$ , as well as  $SST_{\text{intra}}$  and  $SST_{\text{inter}}$  were strongly correlated (Spearman's rho of 0.83 and 0.82 respectively), I used only one measure of mean SST and one measure of variability in SST. Of the variables available,  $SST_{\text{mean}}$  and  $SST_{\text{inter}}$  were least correlated and therefore chosen for further analysis.

### 2.2.3 Explanatory Variables – Human Disturbance

To assess the impact of human disturbances on species compositions and  $\beta$ -diversity, I modeled human presence. Human population is a commonly used proxy of human disturbance, as detailed information of human disturbances on coral reefs (i.e. intensity of fishing pressure) is often either not available or not in comparable formats across regions (Jennings and Polunin 1997, Nadon et al. 2012).

Total population of each island was obtained from 2010 US Census data, except for Midway Atoll, where the population is composed of approximately 60 US Fish and Wildlife Service employees. Human population density at each site (number of people within 20 km radius) is based on the Socioeconomic Data and Applications Center (SEDAC)'s open source product, Global Rural-Urban Mapping Project (GRUMP) Population Count Grid (v1) for 2000. SEDAC provides human population data (up to the year 2000) at a resolution of 30 arc-seconds (~1 km at the equator).

I classified each island as having either “high” or “low” levels of human disturbance based on its human population, similar to Williams et al. (2010), where any island with a permanent population of less than 100 individuals was considered to have low human disturbance (Table 2), except for Aguijan which is an uninhabited but highly disturbed island 8 km offshore Tinian. This classification of islands reflects the perception that humans’ fishing effort may not be a simple function of distance from population centres and that humans may travel large distances to fish. This classification assumes that populated islands (high human disturbance) have the potential to be disturbed in some manner, however I make no assumptions that fishing is uniform across the whole island. To test this assumption, I also conducted a separate set of analyses in which I restricted the data on heavily populated islands to those sites near population centres, specifically including only those sites with a population of at least 50,000 individuals living within a 20 km radius of the site (Appendix E). This characterization of nearness to population centres reflect the increased exploitation of reefs near population centres (Nadon et al. 2012), however my analysis found no significant or consistent differences in results calculated from sites near population centres and all sites on the same island.

I also examined the utility of the Reefs at Risk Revisited (Burke et al. 2011) data set, to identify other potential sources of human disturbance. Reefs at Risk Revisited aggregates a large number of datasets, including data on coastal development, watershed- and marine-based pollution, and exploitative fishing, to create categorical threat level of reefs. I found, however, that the Reefs at Risk Revisited predictions of human disturbance were inconsistent with the local knowledge of CRED’s experts. This inconsistency, together with the categorical nature of the local threats classification, lead me to pursue human population data as a proxy for disturbance rather than utilizing Reefs at Risk Revisited.

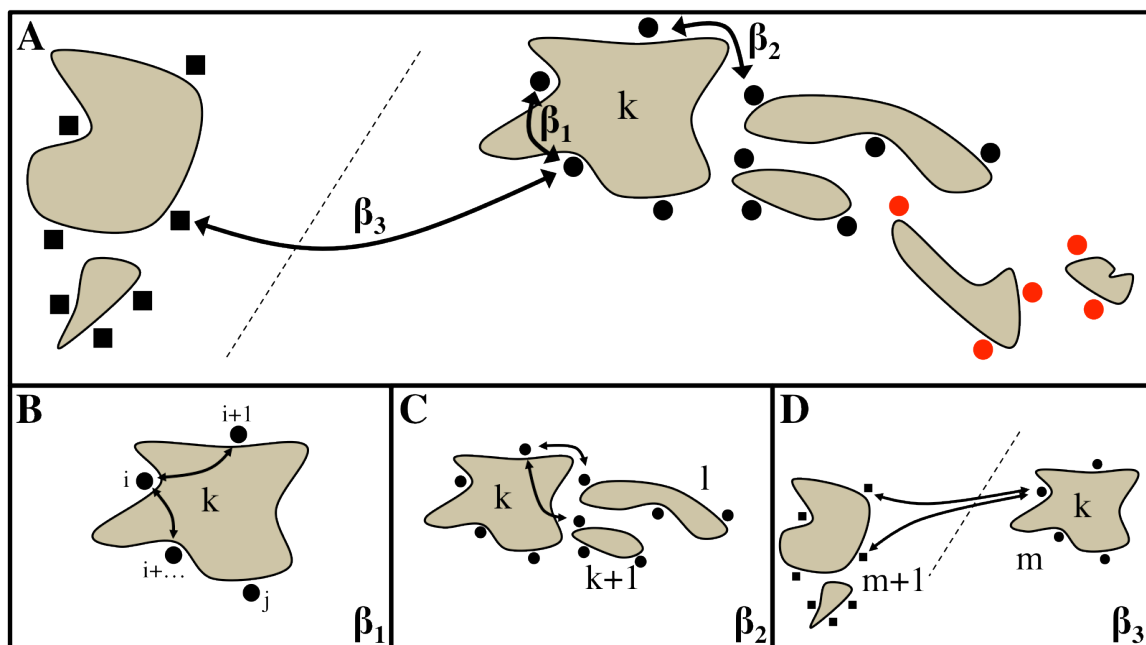
### **2.3 Modeling $\beta$ -diversity**

I examined five different  $\beta$ -diversity metrics, including three distinct null models, at each of three different spatial scales. All five metrics of  $\beta$ -diversity are based on pairwise dissimilarity of communities (i.e. survey sites) that assess the dissimilarity in either

species presences or species abundances between two sites. The five metrics differ, however, in the way that they handle species abundances or account for different biases.

### 2.3.1 Beta Diversity Study Design

The scale at which  $\beta$ -diversity is studied can greatly influence the findings (Cornell et al. 2007). Therefore, in this study I examine  $\beta$ -diversity at three separate scales, within islands; within island groups; and across islands groups (Figure 2). For all three scales of analysis, sites are the unit of measure and  $\beta$ -diversity is calculated for all islands,  $k - l$ , in island groups  $m - n$ , using many pairwise dissimilarities; however which combination of

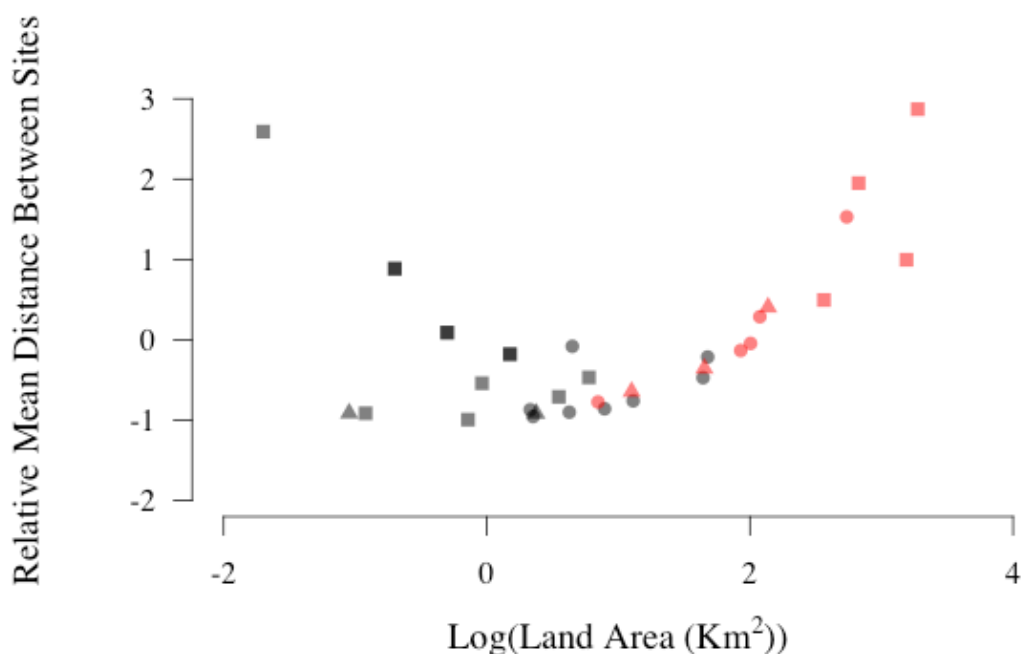


**Figure 2.** Spatially hierarchical sampling design. A) Overview of  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$ , which represents  $\beta$ -diversity for island  $k$  at three scales: within island, within island group, and across island groups, respectively. Within each island group,  $m - n$ ,  $\beta$ -diversity is calculated for all islands,  $k - l$ , where each island has many sites,  $i - j$ . Only combinations of islands and sites within the same human disturbance regime (low or high; black or red symbols respectively) are used to calculate  $\beta$ -diversity for each island. Dotted line represents distinction between island groups. At the within island scale (B), all sites on island  $k$ ,  $i - j$ , are used to calculate  $\beta_1$  for island  $k$ . At the within island group scale (C), to calculate  $\beta_2$  for island  $k$ , all sites on island  $k$  are compared to all other sites on islands  $k+1 - l$ , within that island group,  $m$ . At the among island groups scale (D), to calculate  $\beta_3$  for island  $k$  in island group  $m$ , all sites on island  $k$  in island group  $m$  are compared to all other sites on islands that are in a different island group,  $m+1 - n$ .

sites are used to calculate  $\beta$ -diversity for each island differs depending on the scale of analysis:

- 1) Within island  $\beta$ -diversity  $\beta_1$ , is calculated for each island  $k$  using all pairwise dissimilarities among sites  $i - j$ , on island  $k$  (Figure 2B).
- 2) Within island group  $\beta$ -diversity  $\beta_2$ , is calculated for island  $k$  using all pairwise dissimilarities where site 1 is on island  $k$  and site 2 is on a different island,  $k+1 - l$ , within island group  $m$  (Figure 2C).
- 3) Across island group  $\beta$ -diversity  $\beta_3$ , is calculated for island  $k$  using all pairwise dissimilarities where site 1 is on island  $k$  and site 2 is on an island in a different island group,  $m+1 - n$  (Figure 2D).

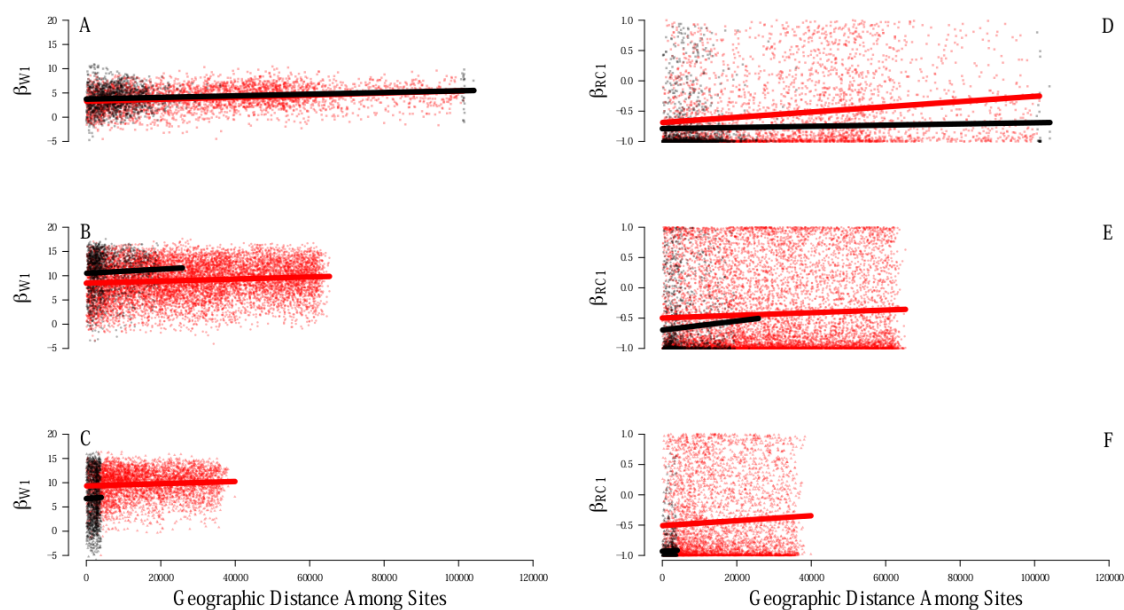
At all three spatial scales, comparisons were only completed among sites that were in the same human disturbance category: either low or high.



**Figure 3.** Relationship between island area and mean distance within sites on each island (rescaled to mean = 0 and standard deviation = 1). Each point is one island, with either high (red) or low (black) human disturbance in the Hawaiian Archipelago (squares), Mariana Archipelago (circles) and American Samoa (triangles). Only island surveyed in the second survey period are plotted.



Because the islands in this study differ substantially in size, the distance between sites at each scale was not consistent among islands: both very large and very small islands increasing distance between sites (Figure 3). As expected, preliminary analyses showed that pairwise site-site dissimilarity tended to increase with increasing distance between sites (Figure 4), although with a high degree of scatter around the relationship and generally with coefficients of determination ( $R^2$ ) values  $< 0.1$  (Supplemental Figure 7). Therefore, simply calculating  $\beta$ -diversity as the mean of many pairwise dissimilarities on islands with different mean distances among sites would result in biases. To reduce this bias, I corrected island  $\beta$ -diversity values to a standard distance. This was done by “unpacking” site-site dissimilarity matrices into vectors of pairwise site-site dissimilarity and regressing these values on geographical distance between sites for each island by year combination independently. Regression slope and y-intercepts were used to estimate  $\beta$ -diversity on each island at a common distance, equal to the largest geographical distance present on all islands. This methodology assumes equal potential for migration



**Figure 4.** Pairwise dissimilarity among sites within islands calculated from an abundance based null model (A – C) and incidence based null model (D – F) as a function of geographical distance (in arbitrary units) between site. Sites on low (black) and high (red) human disturbance islands of the Hawaiian Archipelago (A & D), Mariana Islands Archipelago (B & E) and American Samoa (C & F) are plotted from the most recent survey period only. Lines show linear relationship from all pairwise dissimilarities within a given human disturbance category for each island group. Trends are consistent across larger spatial scales (not shown).

among sites in any direction (i.e. non-directionality of currents), as sophisticated site-specific current models would otherwise be required and were not available. All island  $\beta$ -diversity values presented are distance corrected values.

All three null models require definition of a “*regional species pool*” that encompasses all sites that are used to calculate site-site pairwise dissimilarities, and  $\beta$ -diversity thereafter. To calculate  $\beta$ -diversity at three spatial scales, the regional pool was defined at two different scales accordingly. To calculate  $\beta$ -diversity within islands or within island groups the regional pool was set at the island group scale. Many reef fishes have very large larval dispersal potential (Shanks 2009), and previous  $\beta$ -diversity studies indicate that there is no greater barrier for coral propagule dispersal among islands than among sites within them (Cornell et al. 2007); therefore, I feel justified in assuming island groups represent a suitable regional pool where dispersal among islands is probable. To calculate  $\beta$ -diversity across island groups, the regional pool was set as the whole Pacific Ocean Basin. Exploratory analysis suggested the taxonomic overlap among the island groups was fairly high (65 – 83%), indicating that over geological time the great distance among island groups is not a barrier for the dispersal of many fishes (Kinlan and Gaines 2003, Nunes et al. 2011).

Defining regional pools for the null models allowed for me to account for the unequal sampling effort on each island. To do so, I weighted each site’s contribution to a *corrected regional pool* by the ratio of the number of sites sampled on that island to the reef area of that island (Eq. 1). Weighting ensured that islands that were sampled disproportionately more than others, based on their reef area, do not influence the regional pool more.

$$wf_k = \frac{\text{Num. sites sampled on island } k}{\text{reef area of island } k} \quad \text{Eq. 1}$$

Null models were built for each year by island group combination for the within island and within island group scales ( $n = 6$ ). For the across island group scale, two null models were built, combining surveys from 2009-2010 and 2011-2012 into separate two-year blocks.

### 2.3.2 Beta Diversity Metrics

I first calculated  $\beta$ -diversity from two standard pairwise site-site dissimilarity metrics, Jaccard dissimilarity (Jaccard 1912) and Bray-Curtis dissimilarity (Bray and Curtis 1957). The former (Eq. 2) uses only species incidence information, whereas the latter (Eq. 3) takes into consideration the species abundances:

$$d_{kl} = 1 - \frac{|k \cap l|}{|k \cup l|} \quad \text{Eq. 2}$$

$$d_{kl} = \frac{\sum_i |x_{ik} - x_{il}|}{\sum_i (x_{ik} + x_{il})} \quad \text{Eq. 3}$$

Where  $d_{kl}$  is the Jaccard or Bray-Curtis dissimilarity between sites  $k$  and  $l$ ,  $k \cap l$  and  $k \cup l$  are the intersection (shared) or union (total number) respectively of species among sites  $k$  and  $l$ , and  $x_{ik}$  and  $x_{il}$  are the abundance of species  $i$  at sites  $k$  and  $l$  respectively. Both Jaccard and Bray-Curtis dissimilarity are bound by 0 and 1, where 0 indicates no dissimilarity and 1 indicates complete dissimilarity. While both Bray-Curtis and Jaccard dissimilarity can be biased due to differences among samples (Koleff et al. 2003, Chao et al. 2006), Jaccard dissimilarity is known to be particularly strongly biased by differences in species richness ( $\alpha$ -diversity). The bias is such that sites that have lower  $\alpha$ -diversity appear more dissimilar than otherwise expected by random chance (Chase and Myers 2011).

To overcome potential biases of the Jaccard and Bray-Curtis metrics, I implemented three types of null models. The three null models used generate estimates of  $\beta$ -diversity that remove either the bias of:

- 1)  $\alpha$ -diversity on incidence-based  $\beta$ -diversity,
- 2)  $\alpha$ -diversity on abundance-based  $\beta$ -diversity, or
- 3) the bias of community total abundance on abundance-based  $\beta$  diversity.

These biases were removed individually in separate null models as attempting to remove the effect of both  $\alpha$ -diversity and community total abundance on  $\beta$ -diversity simultaneously might excessively constrain the null models, thereby limiting the interpretability of the results. Both null models generate  $\beta$ -deviation values, sensu Karp et

al. (2012), which are synonymous to  $\beta$ -diversity values and as such will be referred to as  $\beta$ -diversity values hereafter.

First, to account for the bias of  $\alpha$ -diversity on  $\beta$ -diversity while simultaneously down-weighting the importance of highly abundant individuals, I used the Raup-Crick method, an incidence based null model that has seen frequent application (Raup and Crick 1979, Vellend 2004, Vellend et al. 2007, Chase and Myers 2011). This method assesses the probability that two communities share fewer species than expected if the communities were randomly assembled from a regional pool where the probability of a species being present at a site is proportional to the total number of sites at which it was observed. I modified the Raup-Crick model by multiplying the occurrence matrix by the weighting factor of island,  $wf_k$ , thereby ensuring the probability of a species being present in a null community was proportional to the corrected regional pool. In brief, the Raup-Crick null model works for sites  $i$  and  $j$  with  $\alpha = x$  and  $y$  respectively by taking  $x$  and  $y$  samples from the regional pool and assessing the number of species shared among these two sets. This sampling was done 9999 times per site-site combination and is compared to the observed number of species shared among  $i$  and  $j$ . For a given site-site comparison, the Raup-Crick measure of  $\beta$ -diversity (henceforth  $\beta_{RC}$ ) is the proportion of null communities that share the same or greater number of species than observed, scaled between -1 and 1.  $\beta_{RC}$  near -1 indicates highly similar communities, whereas  $\beta_{RC}$  near 1 indicates highly different communities.

Second, to account for differences in alpha-diversity amongst sites, while maintaining the unequal species abundance distributions at sites, I developed an abundance based null model (Wiwchar null model), denoted  $\beta_w$ , which constrains alpha but allows the number of fish at each site to vary. This null model builds upon the Raup-Crick structure. In this null model, null communities are created with  $\alpha$  diversities ( $\alpha_{null}$ ) equal to the observed alpha diversity ( $\alpha_{obs}$ ) at each respective community. A corrected regional pool was created by multiplying species abundances at all sites by each island's weighting factor,  $wf_k$ , and summing species abundances across all sites within the geographical region (i.e. island group or whole Pacific) of study for each survey period. Due to weighting factors, species abundances could be non-integer values, and therefore the abundances were

rounded in the regional pool. If weighting caused species abundances in the regional pool to be less than 1 but greater than 0, abundances were rounded to 1, therein assuring the inclusion of all observed rare species in the corrected regional pool. Rounding also caused the regional pool to have a total abundance that differed from the observed (non-corrected) total species abundances across all sites (difference < 1%); therefore, the difference was made up by randomly adding or removing individuals to or from the regional pool.

Sequentially, individuals from the regional pool were randomly chosen and added to initially empty site X species matrices where site was also chosen randomly. Each individual could be placed in a site provided its placement did not cause that site's richness ( $\alpha_{null}$ ) to exceed the observed richness ( $\alpha_{obs}$ ) at that site. If this condition was not met, the individual would be placed in another site, chosen at random. If the individual could not be placed (i.e. all sites had reached  $\alpha_{null} = \alpha_{obs}$ ), the individual was placed at a new randomly selected site and all the individuals of one of the 10 most abundant species at that site was removed. These displaced individuals were placed back into sites once all species in the regional pool were placed, and were only placed in sites where that species was already present. This simulation was repeated 9999 times. Any iteration with  $\alpha_{null} < \alpha_{obs}$  for any site was discarded and repeated until  $\alpha_{null} = \alpha_{obs}$ , analogous to discarding empty model algorithms (Ulrich and Gotelli 2010). For each iteration, the Bray-Curtis dissimilarity was calculated for all pairwise combinations of sites ( $\beta_{BC,null}$ ). Analogous to the method used by Karp et al. (2012), a  $\beta$ -deviation ( $\beta_w$ ) was calculated for each combination of pairwise site combinations (Eq. 4). For both  $\beta$ -deviation based metrics, large  $\beta$ -deviations indicate highly different communities. The absolute ranges and values of  $\beta$ -deviations are inconsequential to general results and reflect the degree of randomness with which null communities could be built. In general, less constrictive null models will result in larger  $\beta$ -deviations. Full null model R script is available in Appendix A.

$$\beta_{deviation} = \frac{\beta_{BC,obs} - mean(\beta_{BC,null})}{\sigma \beta_{BC,null}} \quad \text{Eq. 4}$$

Finally, to account for the unequal total abundance of fishes among sites, I developed an abundance based null model based on that of Kraft et al. (2011) in which the total abundance of fishes at a given site was constrained to equal the observed number of fishes at the site, but species richness ( $\alpha_{\text{null}}$ ) of the site was allowed to vary. To do so, each randomly selected individual from the corrected regional pool (as composed above in  $\beta_{\text{w}}$ ) was allocated to a site chosen at random, without replacement, until a null site's total abundance reached that site's observed abundance. Again, a  $\beta$ -deviation ( $\beta_{\text{K}}$ ) was calculated from the observed Bray-Curtis similarity ( $\beta_{\text{BC,obs}}$ ), and the mean and standard deviation of  $\beta_{\text{BC,null}}$  (Eq. 4).

## 2.4 Modeling Drivers of $\beta$ -Diversity

I conducted two distinct types of analyses to investigate if  $\beta$ -diversity is greater at islands with low human disturbance, and to elucidate the relative influence of anthropogenic disturbances and environmental covariates on  $\beta$ -diversity. Calculating pairwise dissimilarity generate more data points than sites observed (Eq. 5):

$$\text{Num. pairwise values} = \frac{[(\text{num. sites})^2 - \text{num. sites}]}{2} \quad \text{Eq. 5}$$

Therefore, any analysis that requires independent data points (i.e. regression analysis) cannot be completed on raw data points, as this would be a form of pseudo-replication. For that reason, data were analysed in two different forms:

- 1) regression of *island*  $\beta$ -diversity values generated from many non-independent pairwise site-site dissimilarity values (section 2.3.1). Collapsing pairwise data to single values per island removed issues of non-independence and allowed me to assess if islands with low human disturbance have higher, lower, or equal  $\beta$ -diversity.
- 2) matrices of pairwise site-site dissimilarity were used in distance-based redundancy analysis (dbRDA), a multivariate framework where principle coordinate analysis is the first step and uses information contained within the non-independence of data points to generate independent site axes. This approach partitions variation in community compositional dissimilarity across explanatory variables and addresses whether or not there is a difference in the

variables that drive community composition on island with low or high human disturbance.

Regression of island  $\beta$ -diversity values provides insight into the variables that influence  $\beta$ -diversity in addition to assessing if the  $\beta$ -diversity on islands with high human disturbance differs from those with low human disturbance. dbRDA gives insight into the structuring of communities, addressing which variables are most likely to drive community composition. Full analysis was performed only on Raup-Crick and Wiwchar null-model outputs and  $\beta_{RC}$  and  $\beta_W$ . These approaches are detailed in the following two sections.

#### 2.4.1 Modeling Anthropogenic and Biophysical Influences on $\beta$ -Diversity

Linear regression models were used to estimate the effect of human disturbance and environmental variables on island  $\beta$ -diversity. Models were built using all island-specific explanatory variables as well as the mean of site-specific explanatory variables for each island with the exception of island composition and distance to IPCA (Table 3). All non-categorical variables were standardized prior to modelling. Additionally, an interaction term between human inhabitancy and live coral complexity was included. Thus, at the smallest (within island) scale, I modelled:

$$\beta_{RC,1} \sim \beta_0 + \beta_H + \beta_{IG} + \beta_A + \beta_{WE} + \beta_{LCC} + \beta_{SSTm} + \beta_{SSTv} + (\beta_{LCC} : \beta_H)$$

in which  $\beta_{RC,1}$  is the distance corrected  $\beta_{RC}$  for each island,  $\beta_0$  is the intercept,  $\beta_H$  is the effect of human disturbance,  $\beta_{IG}$  is the effect of island group,  $\beta_A$  is the effect of reef area,  $\beta_{WE}$  is the effect of wave energy,  $\beta_{LCC}$  is the effect of live coral complexity,  $\beta_{SSTm}$  is the effect of  $SST_{mean}$ ,  $\beta_{SSTv}$  is the effect of  $SST_{var}$ , and  $(\beta_{LCC} : \beta_H)$  is the interaction between live coral cover and human inhabitancy. The interaction of  $\beta_{LCC} : \beta_H$  was included because human disturbances could be evident in an altered live coral complexity on islands with high human disturbance.  $\beta_{IG}$  was included to represent unmeasured variability of environmental variables or intrinsic differences in community assembly between the three island groups. I built a similar model at this scale for the distance corrected mean of the abundance-based null model,  $\beta_W$ , as well as for each of these null models,  $\beta_{RC}$  and  $\beta_W$ , at the other two scales. American Samoa could not be modelled at the within island group

scale because there are only two islands with low human disturbance. As such, the geographical distance between sites at this scale generally took one value with very little variance, and therefore linear regression of  $\beta$ -diversity with geographic distance to obtain distance-corrected  $\beta$  values were problematic. Models with subsets of the variables were built using all combinations of the explanatory variables specified above using the R package “glmulti”. Models were also created for both  $\beta_{RC}$  and  $\beta_W$  at all three scales without the inclusion of live coral complexity as an explanatory variable (Appendix F). This was done because as differences in live coral complexity among islands with low and high human disturbance likely reflect the effect of human disturbances, and as such, modelling both live coral complexity and human disturbance could be seen as including the effect of human disturbance twice in one model. The model selection criteria AIC (Burnham and Anderson 1998) was used to determine the model with the best fit, and AIC weights used to create model averaged parameter estimates. Model averaging used all models within 2  $\Delta$ AIC of the best model. Pseudo  $R^2$  was calculated for each model as per Dobson (2003):

$$Pseudo R^2 = \frac{null\ deviance - residual\ deviance}{null\ deviance} \quad Eq. 6$$

#### **2.4.2 Modelling Anthropogenic and Biophysical Influences on Community Composition**

To address the relative importance of habitat or spatial separation in driving local community composition, the basis of  $\beta$ -diversity, I used partial distance-based redundancy analysis (partial dbRDA). dbRDA is a multivariate extension of multiple regression in which community data in the form of dissimilarity matrices, such as pairwise site-site dissimilarity or  $\beta$  deviation values, are regressed on site-specific explanatory variables (Legendre and Legendre 2012). Partial dbRDA differs slightly, as it removes and estimates the effect of set covariates prior to ordination of other explanatory variables (Borcard et al. 1992, Legendre and Legendre 2012). Because of the lack of requirements for Euclidian distance measures and multivariate normality, dbRDA has been utilized frequently in community ecology studies connecting patterns in community



composition or other dissimilarity metrics with predictor variables (e.g. Myers et al. (2012) and Parker et al. (2012)).

In short, dbRDA involves an initial principal coordinate analysis (PCoA) of the pairwise community dissimilarity matrix,  $Y$ . In partial dbRDA, all eigenvectors generated from the PCoA are subsequently regressed on a constrained covariate,  $W$ , generating residuals  $y_{res|W}$ . The residuals of the explanatory variables,  $X_{res|W}$ , are calculated similarly. The third step is the production of canonical eigenvalues, eigenvectors, and matrix  $Z$  containing the canonical axes completed by a principal component analysis (PCA) of  $\hat{Y}$ . Dividing the sum of the constrained eigenvalues by the sum of all eigenvalues yields the variation in community composition accounted for by all explanatory variables (Vellend et al. 2007). Because the denominator of this proportion increases with the number of sites in the analysis (Cottenie 2005), 100 partial dbRDA were conducted on a subset of sites for each scale. 100 repetitions were found to be more than sufficient to converge on stable mean values of proportion variation explained by explanatory variables.

For all three scales, the maximum amount of variation accounted for by each site-specific variable was determined by using that variable as the constraining covariate,  $W$ . This procedure attributes as much variability as possible to that variable, prior to ordination of the community data with other explanatory variables. The amount of variability accounted for by all, or a subset of, environmental variables combined was also determined, and was less than the sum of the variability explained by each variable alone because a proportion of the environmental variability is collinear. The amount of variability accounted for by distance was also initially estimated. To do so, the geographical distance among sites was converted into distance-based Moran's eigenvector maps (dbMEM), and all spatial eigenfunctions generated were used as site-specific explanatory variables. This form of dbMEM is often known as principal coordinates of neighbour matrices (PCNM), and involves the truncation of distances among many objects or sites to a few spatial eigenfunctions to be used as explanatory variables of response data (Borcard and Legendre 2002, Legendre and Legendre 2012). This method is considered more powerful than the alternative method of converting rectangular site-specific variables into distances to compare against geographical

distances in partial mantel tests (Legendre et al. 2008). However, distance tended to over fit the data and account for nearly all variability in community composition. As such, I was unable to accurately compare the amount of variability accounted for by distance, and distance was removed from all dbRDA.

Similar to modeling  $\beta$ -diversity at three spatial scales, partial dbRDA was conducted at three scales: within island, within island group, and among island group. Unlike modeling  $\beta$ -diversity, however, the sites used for comparisons among sites at the larger two scales are not directly comparable to island  $\beta$ -diversity values at these two scales.

At the smallest scale, within island, partial dbRDA was conducted on each island separately, using a random subset of 10 sites per island, thereby excluding analysis of two islands in the Mariana Islands Archipelago (Alamagan and Sarigan), and three islands in the Hawaiian Archipelago (French Frigate, Necker, and Nihoa). 10 sites per island was chosen as a threshold to allow model convergence while reducing the number of islands not analyzed. The mean proportion of variation accounted for by each variable individually, as well as the mean proportion of variation accounted for by all environmental variables was calculated from all islands over 100 iterations. At this scale, explanatory variables included only site-specific environmental data: live coral cover, habitat rugosity,  $SST_{\text{mean}}$  and  $SST_{\text{var}}$ .

For the middle scale, comparing within island groups, it was not possible to conduct partial dbRDA by comparing only sites on different islands. To do so would require matrices with blank elements, which is not amenable to the full matrix requirement of dbRDA. Therefore, I conducted partial dbRDA on subsets of each island group dissimilarity matrix that included both comparisons of sites on different islands as well as comparisons of sites on the same island. While this combines multiple scales into one analysis, reducing my ability to interpret fine scale-specific differences, it maintains a high sample size and statistical power therein. For each island group and human disturbance category, partial dbRDA were conducted on random subsets of  $N_{\text{min}}$  sites, chosen from islands of the human disturbance category under study, where  $N_{\text{min}}$  was the fewest number of sites surveyed in *either* inhabitancy category in all island groups ( $N_{\text{min}} = 70$  sites on Samoan islands with low disturbance and complete predictor data). Sites

were randomly chosen with a probability relative to the proportion of sites on each island to all the sites in the dissimilarity matrix of the island group in question. The proportion of variation accounted for by each variable as well as by all variables combined was calculated from the mean of all 100 iterations. At this scale, I tested both site-specific variables, indicated above, and island-specific variables: wave energy, reef area, and island composition.

Analysis at the largest scale had the same issue as that at the middle scale, where it was not possible to conduct partial dbRDA without including pairwise site-site comparisons pertaining to the smaller two scales, within islands or within island groups. Therefore, I conducted partial dbRDA on subsets of the full community dissimilarity matrix from the largest scale. For each human disturbance category, partial dbRDA was conducted on  $N_{\min}$  randomly selected sites, where  $N_{\min}$  was the fewest number of sites within either human disturbance category ( $N_{\min} = 353$  sites on all islands with low disturbance and complete predictor data). Similar to the middle scale, sites were randomly chosen with a probability relative to the proportion of sites on each island to all the sites in dissimilarity matrix of the Pacific Ocean basin. Once more, the mean proportion of variation accounted for by each variable and by all environmental variables over all 100 iterations was calculated per human inhabitancy category.

The proportion of variability accounted for by environmental variables was compared among islands with low and high disturbance at the smaller two scales using t-tests or ANOVA. At the smallest scale, within-islands, ANOVA was performed comparing the variability accounted for by the combination of site-specific variables on islands with low or high human disturbance in each island group ( $n = 27$ ). At the within island group scale, paired t-tests were performed comparing the variability accounted for by combination of site-specific variables, island-specific variables, or both, among the islands with low or high human disturbance in each of the three island groups ( $n = 3$ ).

All analysis was performed in R 2.14.2, 2.15.1, and 3.0.1 (R Core Team, 2013), except Venn Diagrams produced with Euler APE (<http://www.eulerdiagrams.org/eulerAPE/>).

### 3. Results

In total, 611280 individual fish were counted and identified to species level during 1615 CRED SPC surveys between 2009 and 2012, representing 521 unique species from 182 genera and 55 families. The most frequently observed species, in decreasing frequency of occurrence at all sites, were *Acanthurus nigrofuscus* (a medium sized herbivorous surgeonfish,  $L_{\max} = 38$  cm; observed at 71.7% of sites), *Parupeneus multifasciatus* (a medium sized goatfish,  $L_{\max} = 40$  cm; observed at 65.6% of sites), *Paracirrhites arcatus* (a small carnivorous hawkfish,  $L_{\max} = 14$  cm; observed at 63.7% of sites), *Melichthys vidua* (a medium sized herbivore triggerfish,  $L_{\max} = 34$  cm; observed at 55.8% of sites), and *Stegastes fasciolatus* (a small herbivorous damselfish,  $L_{\max} = 16$  cm; observed at 50.1% of sites). Each of these species was observed in all three island groups. The most numerically abundant species were all planktivorous, and in order of decreasing abundance were *Chromis acares* (a very small damselfish,  $L_{\max} = 6$  cm; 90,247 individuals), *Pseudanthias pascalus* (a small anthias,  $L_{\max} = 17$  cm; 44,391 individuals), *Chromis vanderbilti* (a small damselfish,  $L_{\max} = 15$  cm; 26,332 individuals), *Pomacentrus vaiuli* (a very small damselfish,  $L_{\max} = 10$  cm; 24,459 individuals), and *Pomachromis guamensis* (a very small damselfish,  $L_{\max} = 6$  cm; 22,526 individuals). Of these species, only *C. acares* and *C. vanderbilti* were observed in all three island groups.

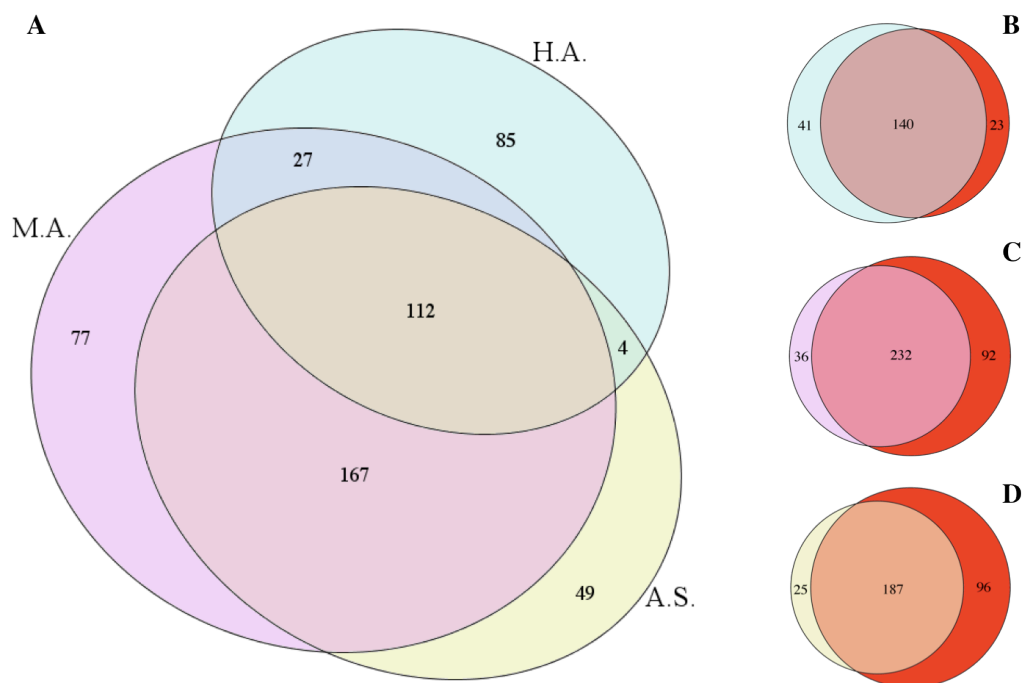
On average, there were considerably fewer fishes observed during surveys in the Hawaiian Archipelago compared to American Samoa or the Mariana Islands Archipelago (Table 4). There also tended to be fewer fishes observed at sites on islands with high

human disturbance in all three island groups (Table 4). The same trend is observed for species richness, albeit weaker, where species richness tended to be lower at sites on islands with high human disturbance (Table 4). The species that were most abundant or most frequently observed on each island were somewhat consistent among islands within each disturbance category for an island group, with the exception of low human disturbance islands in the Hawaiian Archipelago (Table 5).

**Table 4.** Mean fish counts and observed species richness ( $\pm$  SEM) at sites on all islands, or low and high human disturbance islands within each island group. Total observed richness of islands (Island Richness) or regions (Reg. Richness) within each human disturbance category is also shown. The mean proportion of sites that each species were observed at (Freq. Occur.) is presented for islands with low and high human disturbance. Data are from second survey period, however the same trends were observed in both survey periods.

Island Group	Human Disturbance	Site Abundance	Site Richness	Island Richness	Reg. Richness	Freq. Occur.
Hawaiian Archipelago	All	211 $\pm$ 15	27 $\pm$ 0.5	105	204	
	Low	257 $\pm$ 30	30 $\pm$ 0.7	97	181	0.15
	High	169 $\pm$ 15	25 $\pm$ 0.7	125	163	0.12
Mariana Archipelago	All	489 $\pm$ 18	43 $\pm$ 0.6	170	360	
	Low	540 $\pm$ 39	46 $\pm$ 0.9	157	268	0.12
	High	457 $\pm$ 15	41 $\pm$ 0.8	192	324	0.11
American Samoa	All	539 $\pm$ 27	42 $\pm$ 0.6	195	308	
	Low	777 $\pm$ 60	42 $\pm$ 0.7	166	212	0.11
	High	415 $\pm$ 19	42 $\pm$ 0.8	215	283	0.08
All	All	402 $\pm$ 12	37 $\pm$ 0.4	146	491	
	Low	469 $\pm$ 25	38 $\pm$ 0.6	128	392	0.08
	High	356 $\pm$ 10	36 $\pm$ 0.7	175	437	0.07

There was a high degree of overlap between species present in the three island groups and among remote and populated islands of each island group. Of the 521 species observed during surveys, >20% were observed in all three island groups, and of the species present in each island group, only 37% (n = 85), 20% (n = 77), and 15% (n = 49) were unique to the Hawaiian Archipelago, Mariana Archipelago, and American Samoa respectively (Figure 5A). The majority of species present in each island group were observed on both islands with low or high human disturbance (Figure 5B-D); however, several species were observed only on islands with low or high disturbance (Appendix D). The species that were absent from populated islands tended to be larger, and more



**Figure 5.** Overlap of species observed in the Hawaiian Archipelago (H.A.; blue), Mariana Archipelago (M.A.; pink) and American Samoa (A.S.; yellow) (A), and among low and high (red) human disturbance islands within each island group respectively (B – D). Overlap in B represents species observed during either survey period, whereas B – D represents only the second survey period.

often piscivorous or planktivorous (Table 6). In contrast, a large proportion (>40%) of species absent from low human disturbance islands in all three island groups were low carnivores, primarily invertivores (Table 6). This trend was consistent when comparing all species present on either low or high human disturbance islands in each island group, as low human disturbance islands tended to have a greater proportion of piscivorous fish, and lower proportion of low carnivores (Figure 6). Notably absent from high human disturbance islands, but present in islands with low human disturbance, were Galapagos shark (*Carcharhinus galapagensis*), grey reef shark (*Carcharhinus amblyrhynchos*), giant trevally (*Caranx ignobilis*), black jack (*Caranx lugubris*), greater amberjack (*Seriola*

**Table 5.** Island richness, number of- and proportion relative to island richness of- species observed only once or twice on each island (singletons and doubletons), as well as the most abundant and most frequently observed species per island. ‘All’ denotes the aggregation of all sites on islands within either human disturbance category (H. Distb). H.A. = Hawaiian Archipelago; M.A. = Mariana Archipelago; A.S.= American Samoa.

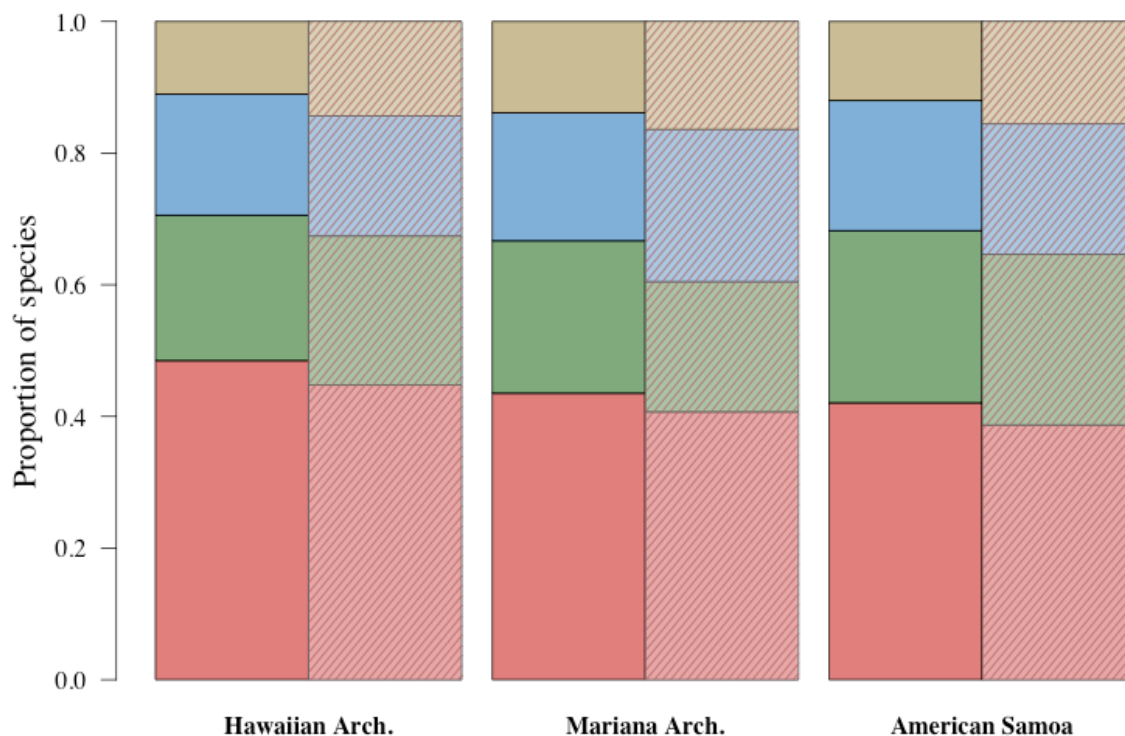
Island Group	Island	H. Distb.	Rich ness	Singletons		Doubletons		Most abundant	Most Frequent
				#	Prop	#	Prop		
H.A.	French Frigate	Low	76	35	0.46	19	0.25	<i>Ctenochaetus strigosus</i>	<i>Parupeneus multifasciatus</i>
	Gardner	Low	94	22	0.23	22	0.23	<i>Chromis vanderbilti</i>	<i>Parupeneus multifasciatus</i>
	Kure	Low	99	24	0.24	20	0.2	<i>Thalassoma duperrey</i>	<i>Thalassoma duperrey</i>
	Laysan	Low	111	26	0.23	18	0.16	<i>Spratelloides delicatulus</i>	<i>Parupeneus multifasciatus</i>
	Lisianski	Low	91	18	0.20	10	0.11	<i>Thalassoma duperrey</i>	<i>Thalassoma duperrey</i>
	Maro	Low	105	24	0.23	20	0.19	<i>Thalassoma duperrey</i>	<i>Parupeneus multifasciatus</i>
	Midway	Low	111	36	0.32	11	0.1	<i>Acanthurus triostegus</i>	<i>Acanthurus nigroris</i>
	Necker	Low	89	36	0.40	15	0.17	<i>Thalassoma duperrey</i>	<i>Acanthurus olivaceus</i>
	Nihoa	Low	94	42	0.45	10	0.11	<i>Chromis vanderbilti</i>	<i>Chromis vanderbilti</i>
	Pearl & Hermes	Low	103	29	0.28	10	0.1	<i>Chromis hanui</i>	<i>Bodianus bilunulatus</i>
	Lanai	High	111	16	0.14	18	0.16	<i>Chromis vanderbilti</i>	<i>Thalassoma duperrey</i>
	Maui	High	137	26	0.19	11	0.08	<i>Chromis vanderbilti</i>	<i>Thalassoma duperrey</i>
	Molokai	High	137	22	0.16	13	0.09	<i>Chromis vanderbilti</i>	<i>Thalassoma duperrey</i>
	Oahu	High	114	33	0.29	14	0.12	<i>Chromis vanderbilti</i>	<i>Canthigaster jactator</i>
All	Low	181	24	0.31	19	0.16	<i>Chromis vanderbilti</i>	<i>Thalassoma duperrey</i>	
	High	163	18	0.2	16	0.12	<i>Chromis vanderbilti</i>	<i>Thalassoma duperrey</i>	
M.A.	Agrihan	Low	175	48	0.27	25	0.14	<i>Pseudanthias pascalus</i>	<i>Naso lituratus</i>
	Alamagan	Low	121	42	0.35	27	0.22	<i>Chromis acares</i>	<i>Acanthurus nigrofuscus</i>
	Asuncion	Low	169	50	0.30	25	0.15	<i>Chromis acares</i>	<i>Halichoeres ornatissimus</i>
	Farallon de Pajaros	Low	143	50	0.35	23	0.16	<i>Chromis acares</i>	<i>Acanthurus nigrofuscus</i>
	Guguan	Low	139	41	0.29	20	0.14	<i>Chromis acares</i>	<i>Acanthurus nigrofuscus</i>
	Maug	Low	178	36	0.20	29	0.16	<i>Chromis acares</i>	<i>Balistapus undulatus</i>
	Pagan	Low	181	41	0.23	21	0.12	<i>Chromis acares</i>	<i>Acanthurus nigrofuscus</i>
	Sarigan	Low	146	62	0.42	22	0.15	<i>Chromis acares</i>	<i>Acanthurus nigrofuscus</i>
	Aguijan	High	137	37	0.27	26	0.19	<i>Pomachromis guamensis</i>	<i>Melichthys vidua</i>
	Guam	High	295	49	0.17	31	0.11	<i>Pomacentrus vaiuli</i>	<i>Acanthurus nigrofuscus</i>
	Rota	High	190	60	0.32	23	0.12	<i>Pomachromis guamensis</i>	<i>Acanthurus nigrofuscus</i>
	Saipan	High	168	41	0.24	27	0.16	<i>Pomachromis guamensis</i>	<i>Acanthurus nigrofuscus</i>
	Tinian	High	168	47	0.28	25	0.15	<i>Pomachromis guamensis</i>	<i>Pomacentrus vaiuli</i>
All	Low	268	50	0.30	24	0.16	<i>Chromis acares</i>	<i>Naso lituratus</i>	
	High	324	54	0.26	30	0.15	<i>Pomachromis guamensis</i>	<i>Acanthurus nigrofuscus</i>	
A.S.	Rose	Low	168	42	0.25	23	0.14	<i>Chromis acares</i>	<i>Cephalopholis argus</i>
	Swains	Low	163	36	0.22	17	0.1	<i>Chromis acares</i>	<i>Balistapus undulatus</i>
	Ofu & Oloseea	High	205	55	0.27	22	0.11	<i>Chromis iomelas</i>	<i>Acanthurus nigericans</i>
	Tau	High	183	46	0.25	23	0.13	<i>Chromis iomelas</i>	<i>Centroloea flavissima</i>
	Tutuila	High	256	40	0.16	28	0.11	<i>Chromis iomelas</i>	<i>Ctenochaetus striatus</i>
	All	Low	212	43	0.24	23	0.12	<i>Chromis acares</i>	<i>Melichthys vidua</i>
High		283	42	0.23	27	0.025	<i>Chromis iomelas</i>	<i>Ctenochaetus striatus</i>	

*dumerili*), Rainbow runner (*Elagatis bipinnulata*), and Kawakawa (*Euthynnus affinis*) in the Hawaiian Archipelago (Supplemental Table 3); whitetip reef shark (*Triaenodon obesus*), tawny nurse shark (*Nebrius ferrugineus*), and rainbow runner in the Mariana Archipelago (Supplemental Table 5); and grey reef shark, blacktip reef shark (*Carcharhinus melanopterus*), bigeye trevally (*Caranx sexfasciatus*), and black jack in American Samoa (Supplemental Table 7). At a higher taxonomic level, the majority (58%) of all families were present in at all three island groups (Supplemental Figure 1), with very few unique to each: Diodontidae and Cheilodactylidae in the Hawaiian

Archipelago; Dasyatidae, Belonidae, Ginglymostomatidae, and Nemipteridae in the Mariana Islands Archipelago; and Ehippidae in American Samoa.

**Table 6.** Number of species *within* each island group observed on islands of one disturbance category but not the other (# Unique), as well as the median  $L_{\max}$  (maximum length of fish in cm) and general trophic group of those species. H.A. = Hawaiian Archipelago; M.A. = Mariana Islands Archipelago; A.S. = American Samoa; L. Carn = low carnivores (invertivores, small piscivores, and corallivores); Plankt = planktivores; Herb = herbivores; Pisc = piscivores; FG = Functional Group.

Human Disturbance	Island Group	# Unique	Proportion of # Unique in Each FG					$L_{\max}$
			L.Carn	Plankt	Herb	Pisc		
Low	H.A.	41	0.29	0.22	0.20	0.29	32	
	M.A.	36	0.17	0.36	0.14	0.33	37.5	
	A.S.	25	0.20	0.36	0.08	0.36	35	
High	H.A.	23	0.43	0.26	0.13	0.17	24	
	M.A.	92	0.41	0.15	0.29	0.14	24	
	A.S.	96	0.44	0.24	0.22	0.10	25	



**Figure 6.** Proportion of all species present on islands that have low (left bar) or high (right cross-hatch bar) human disturbance in each island group that are piscivores (red), herbivores (green), planktivores (blue) or low carnivores (invertivores, small piscivores, and corallivores; brown).

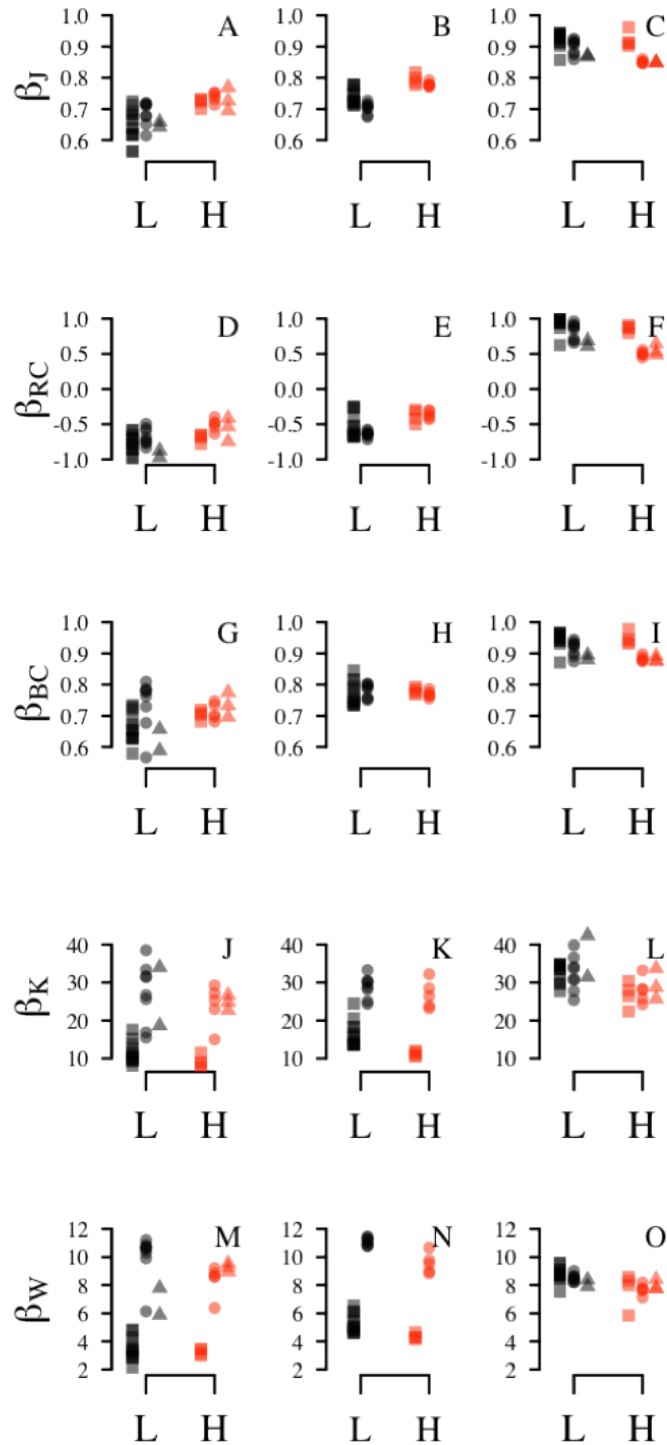


### 3.1 Island $\beta$ -diversity

$\beta$ -diversity varied greatly depending on the type of data used (incidence vs. abundance), scale (within island, within island groups, across island groups;  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  respectively), and the metric used (Jaccard dissimilarity, Bray-Curtis dissimilarity, or null model methods). Most generally, mean  $\beta$ -diversity increased as the scale of analysis increased (Figure 7).

$\beta$ -diversity derived from incidence data ( $\beta_J$  and  $\beta_{RC}$ , derived from Jaccard dissimilarity and Raup-Crick null models respectively) tended to be greater on high disturbance islands in each island group at the within-island and island-group scales. This trend was consistent for each of the three island groups at both of the smaller scales. On average, and relative to the full range of  $\beta_J$  and  $\beta_{RC}$ , islands with high human disturbance had 7% and 6% greater  $\beta_{J,1}$  and  $\beta_{J,2}$ , and 8% and 10% greater  $\beta_{RC,1}$  and  $\beta_{RC,2}$ , respectively (Figure 7 A-B & D-E). Conversely, comparing sites across island groups at the largest scale, mean  $\beta_{J,3}$  and  $\beta_{RC,3}$  tended to be lower on islands with high human disturbance, consistent with biotic homogenization. This trend was also consistent across the three island groups (Figure 7 C & F), albeit only weakly so for  $\beta_{J,3}$  in the Hawaiian Archipelago. Respectively,  $\beta_{J,3}$  and  $\beta_{RC,3}$  were on average 4% and 12% lower on islands with high human disturbance relative to the full range of  $\beta_J$  and  $\beta_{RC}$ . There tended to be little difference between  $\beta_J$  and  $\beta_{RC}$  beyond differences in magnitudes.

With the exception of American Samoa,  $\beta_K$  and  $\beta_W$  were lower on islands with high human disturbance across all three scales, consistent with biotic homogenization. For the Hawaiian Archipelago,  $\beta_{K,1}$  and  $\beta_{W,1}$  were 24% and 7% higher on low human disturbance islands, relative to high disturbance islands. Similarly for the Mariana Archipelago,  $\beta_{K,1}$  and  $\beta_{W,1}$  were 21% and 15% higher on low human disturbance islands. Results for American Samoan islands were variable: both islands with low human disturbance had lower  $\beta_{W,1}$ , whereas  $\beta_{K,1}$  was greater for one island and lower for the other relative to islands with high human disturbance. At the larger two scales, within island group and across island groups,  $\beta_K$  and  $\beta_W$  showed consistency (Figure 7 K-L & N-O).  $\beta_{K,2}$  and  $\beta_{W,2}$  were 45% and 21% lower respectively on islands with high human disturbance in the



**Figure 7.** Island  $\beta$ -diversity values derived from Jaccard dissimilarity (A-C), Raup-Crick dissimilarity (D-F), Bray-Curtis dissimilarity (G-I), Kraft  $\beta$ -deviations (J-L), and Wiwchar  $\beta$ -deviations (M-O). Within island, within island group, and across island group  $\beta$ -diversity are in columns one to three, with islands of low (L; black) and high (H; red) human disturbance plotted for the Hawaiian Archipelago (squares), Mariana Archipelago (circles) and American Samoa (triangles).

Hawaiian Archipelago; and 7% and 16% lower respectively on islands with high human disturbance in the Mariana Archipelago relative to islands with low human disturbance. At the largest scale,  $\beta_{K,3}$  and  $\beta_{W,3}$  on high human disturbance islands were 11% and 22% lower in the Hawaiian Archipelago; 16% and 11% lower respectively in the Mariana Archipelago; and 26% and 2% lower respectively in American Samoa, all relative to low disturbance islands in each island group.

Without accounting for potentially confounding effects of  $\alpha$ -diversity of abundance,  $\beta_{BC}$  (derived from Bray-Curtis dissimilarity) was highly variable across all scales. Relative to low disturbance islands,  $\beta_{BC,1}$  was 6% and 18% greater on high human disturbance islands of the Hawaiian Archipelago and American Samoa respectively, but 3% lower on high disturbance islands of the Mariana Archipelago. Similarly,  $\beta_{BC,2}$  was greater on islands with high human disturbance in the Hawaiian Archipelago but lower in the Mariana Archipelago (Figure 7H). At the largest scale,  $\beta_{BC,3}$  was lower on islands with high human disturbance in the Mariana Archipelago, but not different among disturbance categories in the Hawaiian Archipelago and American Samoa (Figure 7I).

All subsequent analyses were performed on  $\beta_{RC}$  and  $\beta_W$  only. Because of the high degree of correlation within abundance-based null model metrics (Supplementary Figures 1 – 3) I suggest that general trends from  $\beta_W$  are representative of  $\beta_K$ . Any differences that might arise between  $\beta_W$  and  $\beta_K$  are likely reflective of unaccounted-for biases of total community abundance or  $\alpha$ -diversity respectively, of which *neither* are accounted for using traditional  $\beta$ -diversity metrics (i.e.  $\beta_J$  and  $\beta_{BC}$ ). General trends that would be derived from  $\beta_J$  and  $\beta_{BC}$  may differ from results presented, however those trends could arise as an artefact of biases inherent in those metrics (i.e. see Karp et al. (2012)), and therefore interpretation of those results without caution should be avoided.

Model outputs for  $\beta_{RC}$  and  $\beta_W$  showed some consistency across scales, often including the same explanatory variables as important predictors (i.e.  $SST_{var}$  an important predictor of  $\beta_{RC}$  or  $\beta_W$  at both the within and across island group scale), although there are also important differences among scales (i.e. direction of effect of  $SST_{var}$  across aforementioned scales; Table 7). There was a rather high degree of similarity in the

**Table 7.** Model Selection and parameter estimation for  $\beta_{RC}$  and  $\beta_W$  at three spatial scales. Positive or negative signs (+ or –) denote whether the model averaged parameter estimate was positive or negative, and the number of symbols indicates the relative size of parameter estimate. Three symbols indicates the largest effect size, two or one symbols notates parameter estimates that are one or two (or greater) orders of magnitude less than the largest effect size. Asterisks are used for categorical variables where positive or negative notation is non-intuitive. All models within 2  $\Delta AIC$  were used for model averaging and the proportion of those models that each parameter was included in is indicated. Values of model weighted parameter estimates are in Appendix C.  $\beta_{IG}$ ,  $\beta_H$ ,  $\beta_{LCC}$ ,  $(\beta_{LCC} : \beta_H)$ ,  $\beta_{WE}$ ,  $\beta_A$ ,  $\beta_{SSTm}$ , and  $\beta_{SSTv}$  denote effect of island group, human disturbance, live coral cover, the interaction between live coral cover and humans on high disturbance islands, wave energy, reef area, mean sea surface temperature, and interannual variability in sea surface temperature respectively.

Variable	Hypothesis	Presence-Absence (Raup-Crick)						Abundance ( $\alpha$ -constrained Wiwchar)					
		Within Island			Across Island			Within Island			Across Island		
		Within Island	Group	Within Island	Group	Across Island	Group	Within Island	Group	Within Island	Group	Across Island	Group
$\beta_{IG}$	*	***	8/8	*	1/5	***	5/5	***	10/10	***	5/5		
$\beta_H$	-	+++	8/8	+++	5/5	---	5/5	-	2/10	---	5/5	---	5/5
$\beta_{LCC}$	+	---	8/8	+	1/5	+	1/5	-	6/10	-	1/5	++	1/5
$(\beta_{LCC} : \beta_H)$								+++	6/10				
$\beta_{WE}$	+	++	3/8	+++	5/5	---	5/5	-	1/10	+	1/5	--	1/5
$\beta_A$	+	-	3/8	-	1/5	++	3/5	-	2/10	-	1/5	-	1/5
$\beta_{SSTm}$	+	+	1/8	+++	4/5	--	2/5	-	2/10	+	1/5	++	1/5
$\beta_{SSTv}$	-	--	4/8	---	5/5	+++	5/5	-	1/10	--	5/5	+++	5/5

model outputs for  $\beta_{RC}$  and  $\beta_W$  within each scale (i.e. coral complexity only important for both  $\beta_{RC}$  and  $\beta_W$  at the within island scale); however, there are also important and significant deviations from similarities (i.e. wave energy important for  $\beta_{RC}$  but not  $\beta_W$  at the within and across island group scales; Table 7). Model weighted parameter estimates are presented in Appendix C.

### 3.1.1 Incidence-Based $\beta_{RC}$ Models

Incidence-based  $\beta$ -diversity models ( $\beta_{RC}$ ) converged at all three scales, and explained a large proportion of deviance, with pseudo  $R^2$  of best models ( $\Delta AIC \leq 2$ ) of approximately 0.5 for the within island scale, 0.6 for the within island group scale, and 0.8 for the across island group scales (Table 8). AIC weight for the single best model at each scale

**Table 8.** Candidate models of  $\beta_{RC}$  at all three scales: within island  $\beta_{RC,1}$ ; within island group  $\beta_{RC,2}$ ; and across island group  $\beta_{RC,3}$ . AIC weights ( $AIC_W$ ) were calculated to sum to 1 over candidate models with  $\Delta AIC \leq 2$ . Pseudo  $R^2$  indicates the proportion of deviation explained by model relative to deviation in null (intercept only) model.  $SST_v$  and  $SST_m$  denote mean and interannual variability in sea surface temperature.

Model	$AIC_W$	$\Delta AIC$	Pseudo $R^2$
<b>Within island <math>\beta_{RC,1}</math></b>			
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Live Coral Complexity}$	0.227	0	0.511
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{SST}_v + \text{Live Coral Complexity}$	0.181	0.459	0.532
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SST}_v + \text{Live Coral Complexity}$	0.139	0.989	0.551
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Live Coral Complexity} + \text{Reef Area}$	0.106	1.516	0.518
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SST}_v + \text{Live Coral Complexity} + \text{Reef Area}$	0.09	1.846	0.566
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{Live Coral Complexity}$	0.088	1.898	0.512
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{SST}_v + \text{Live Coral Complexity} + \text{Reef Area}$	0.085	1.973	0.538
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Live Coral Complexity} + \text{SST}_m$	0.085	1.977	0.511
<b>Within island group <math>\beta_{RC,2}</math></b>			
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Wave Energy} + \text{SST}_v + \text{SST}_m$	0.39	0	0.639
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SST}_v + \text{SST}_m$	0.17	1.667	0.643
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Wave Energy} + \text{SST}_v + \text{Reef Area} + \text{SST}_m$	0.148	1.946	0.639
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Wave Energy} + \text{SST}_v$	0.147	1.955	0.588
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Wave Energy} + \text{SST}_v + \text{Live Coral Complexity} + \text{SST}_m$	0.146	1.968	0.634
<b>Across island group <math>\beta_{RC,3}</math></b>			
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SST}_v + \text{Reef Area}$	0.265	0	0.781
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SST}_v + \text{Reef Area} + \text{SST}_m$	0.253	0.094	0.792
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SST}_v$	0.21	0.469	0.764
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SST}_v + \text{SST}_m$	0.174	0.848	0.775
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SST}_v + \text{Live Coral Complexity} + \text{Reef Area}$	0.098	1.99	0.781

was 0.23 at the within island scale, 0.39 at the within island group scale, and 0.27 for the among island group scales (Table 8). Model weighted parameter estimates for each scale are available in Supplemental Table 1.

Human disturbance was included in all candidate models within 2  $\Delta$ AIC of the best model at all three scales (Table 8), however the effect direction differed across scales. Within islands and island groups,  $\beta_{RC}$  tended to be greatest on islands populated by humans, whereas across island groups,  $\beta_{RC}$  was greatest on remote islands. Consistent with the trend in  $\beta_{RC}$  at the within island and island group scale, fishes on islands with low human disturbance tended to occupy more sites, on average, than those on high human disturbance islands (Table 4). Similarly, high disturbance islands tended to have greater species richness when sites were pooled on each island (Table 4). When comparing across island groups, the mean frequency of occurrence of species on low and high human disturbance islands tended to be similar (0.078 and 0.074 respectively).

At the within island and island group scales, human disturbance had the largest effect size (Supplemental Table 1). At the across island group scale only island group had a larger effect size than human disturbance (Supplemental Table 1). At the within island group scale,  $\beta_{RC}$  was generally greater in islands of the Mariana Archipelago, whereas across island groups,  $\beta_{RC}$  was greatest in the Hawaiian Archipelago. Island group was only weakly important at the within island group scale, only appearing in one of five models within 2  $\Delta$ AIC of the best model.

Live coral complexity had a relatively strong negative effect on  $\beta_{RC}$  at the within island scale in all models within 2  $\Delta$ AIC of the best model (Table 8). At the larger two scales coral complexity had a positive effect, however was included in only one of five models within 2  $\Delta$ AIC of the best model at each scale. Wave energy had a moderate positive effect on  $\beta_{RC}$  in a minority subset of candidate models within 2  $\Delta$ AIC at the within island scale, a strong positive effect in all candidate models at the within island group scale, and a strong negative effect in all models within 2  $\Delta$ AIC at the among island group scale (Table 8). Reef area had the opposite effect, showing a positive effect at the across island group scale, and negative effect at the within island and within island group

scales (Table 8). Reef area, however, had a relatively small effect size and was not included in all models within  $2 \Delta\text{AIC}$  at any scale.

Mean and interannual variability of sea surface temperature ( $\text{SST}_m$  and  $\text{SST}_v$ ) showed opposing effects at all three scales (Table 8). Within islands and island groups,  $\text{SST}_m$  had a positive effect on  $\beta_{\text{RC}}$ , whereas  $\text{SST}_v$  had a negative effect. The effect of both was generally larger at the within island group scale compared to the within island scale. The effect was opposite at the across island group scale, as  $\text{SST}_m$  had a moderately strong negative effect on  $\beta_{\text{RC}}$ , and  $\text{SST}_v$  had a strong positive effect on  $\beta_{\text{RC}}$ . Both  $\text{SST}_m$  and  $\text{SST}_v$  at the within island group scale and  $\text{SST}_v$  at the across island group scale were included in all models within  $2 \Delta\text{AIC}$  of the best model. Otherwise,  $\text{SST}_m$  and  $\text{SST}_v$  were included in only a subset of candidate models. At all scales  $\text{SST}_v$  had a larger effect size than  $\text{SST}_m$  (Supplemental Table 1)

### 3.1.2 Abundance-Based, $\alpha$ -Constrained $\beta_w$ Models

Models of  $\beta_w$  at all three scales showed convergence, and explained a large proportion of deviance, with pseudo  $R^2$  of top candidate models ( $\Delta\text{AIC} \leq 2$ ) of approximately 0.85 for the within island scale, 0.98 for the within island group scale, and 0.6 among island group scales (Table 9). AIC weight for the single best model at each scale was 0.17 at the within island, 0.36 at the within island group scale, and 0.34 at the across island group scales (Table 9). Model weighted parameter estimates for each scale are available in Supplemental Table 2.

Human disturbance was a strong predictor of  $\beta_w$  at the within- and across- island group scales, and included in all candidate models within  $2 \Delta\text{AIC}$  of the best model (Table 9). At each of these two scales, islands with high human disturbance tended to have lower  $\beta_w$  (Table 9). At the across island group scale human disturbance had the largest effect size (Supplemental Table 2) and at the within island group scale only island group had a larger effect size (Supplemental Table 2). At the within island scale, human disturbance was only included in two of ten candidate models within  $2 \Delta\text{AIC}$  of the best model, however the interaction between humans and coral complexity was included in six of ten (Table 9). The interaction was such that increasing coral complexity correlated with decreasing  $\beta$ -diversity, however on high disturbance islands this relationship was

reversed, such that increasing coral complexity correlated with increased  $\beta$ -diversity (Supplemental Table 2). At this scale, high human disturbance islands tended to have only slightly lower  $\beta_w$ .

At the within island and within island group scales, island group was a strong predictor of  $\beta_w$  and was included in all models within 2  $\Delta$ AIC of the best model (Table 9). At both scales, the Hawaiian Archipelago had considerably lower  $\beta_w$  (Supplemental Table 2).

Wave energy, reef area, and mean sea surface temperature ( $SST_m$ ) were not included in a majority of models of  $\beta_w$  within 2  $\Delta$ AIC of the best model for any of the three scales,

**Table 9.** Candidate models of  $\alpha$ -constrained  $\beta$ -deviations ( $\beta_w$ ) at all three scales: within island  $\beta_{w,1}$ ; within island group  $\beta_{w,2}$ ; and across island group  $\beta_{w,3}$ . AIC weights ( $AIC_w$ ) were calculated to sum to 1 over candidate models with  $\Delta$ AIC  $\leq 2$ . Pseudo  $R^2$  indicates the proportion of deviation explained by model relative to deviation in null (intercept only) model.  $SST_v$  and  $SST_m$  denote mean and interannual variability in sea surface temperature.

Model	$AIC_w$	$\Delta$ AIC	Pseudo $R^2$
<b>Within island <math>\beta_{w,1}</math></b>			
$\beta_{w,1} \sim$ Intercept + Island Group + Live Coral Complexity + (Live Coral Complexity X Humans)	0.173	0	0.869
$\beta_{w,1} \sim$ Intercept + Island Group	0.156	0.205	0.852
$\beta_{w,1} \sim$ Intercept + Island Group + Reef Area	0.105	0.999	0.857
$\beta_{w,1} \sim$ Intercept + Island Group + $SST_v$ + Live Coral Complexity + (Live Coral Complexity X Humans)	0.095	1.194	0.872
$\beta_{w,1} \sim$ Intercept + Humans + Island Group	0.091	1.283	0.856
$\beta_{w,1} \sim$ Intercept + Island Group + Wave Energy + Live Coral Complexity + (Live Coral Complexity X Humans)	0.084	1.448	0.871
$\beta_{w,1} \sim$ Intercept + Island Group + $SST_m$	0.082	1.496	0.855
$\beta_{w,1} \sim$ Intercept + Island Group + $SST_m$ + Live Coral Complexity + (Live Coral Complexity X Humans)	0.079	1.557	0.87
$\beta_{w,1} \sim$ Intercept + Island Group + Live Coral Complexity + Reef Area + (Live Coral Complexity X Humans)	0.07	1.797	0.869
$\beta_{w,1} \sim$ Intercept + Humans + Island Group + Live Coral Complexity + (Live Coral Complexity X Humans)	0.064	1.999	0.869
<b>Within island group <math>\beta_{w,2}</math></b>			
$\beta_{w,2} \sim$ Intercept + Humans + Island Group + $SST_v$	0.363	0	0.981
$\beta_{w,2} \sim$ Intercept + Humans + Island Group + $SST_v$ + Live Coral Complexity	0.185	1.345	0.982
$\beta_{w,2} \sim$ Intercept + Humans + Island Group + $SST_v$ + $SST_m$	0.164	1.592	0.981
$\beta_{w,2} \sim$ Intercept + Humans + Island Group + Wave Energy + $SST_v$	0.153	1.73	0.981
$\beta_{w,2} \sim$ Intercept + Humans + Island Group + $SST_v$ + Reef Area	0.135	1.974	0.981
<b>Across island group <math>\beta_{w,3}</math></b>			
$\beta_{w,3} \sim$ Intercept + Humans + $SST_v$	0.343	0	0.571
$\beta_{w,3} \sim$ Intercept + Humans + $SST_v$ + Live Coral Complexity	0.178	1.312	0.580
$\beta_{w,3} \sim$ Intercept + Humans + Wave Energy + $SST_v$	0.175	1.349	0.579
$\beta_{w,3} \sim$ Intercept + Humans + $SST_v$ + $SST_m$	0.17	1.408	0.578
$\beta_{w,3} \sim$ Intercept + Humans + $SST_v$ + Reef Area	0.133	1.891	0.573



nor was the effect size consistently large for any of the three variables (Table 7).

Interannual variability of sea surface temperature ( $SST_v$ ) was included in all models of  $\beta_w$  with  $\Delta AIC \leq 2$  at the within and across island group scales (Table 9). Similar to models of  $\beta_{RC}$ ,  $SST_v$  had a negative effect on  $\beta_w$  at the within island group scale and a positive effect at the across island group scale (Supplemental Table 2).

## 3.2 Community Composition: Distance-Based Redundancy Analysis

### 3.2.1 Incidence Based Community Composition

The amount of variability in within-island site-site community dissimilarity, assessed by Raup-Crick dissimilarity, that was explained by site-specific variables ( $SST_v$ ,  $SST_m$ , live coral cover, and coral rugosity) was greater on high human disturbance islands in all three island groups (Table 10). Combining the three island groups, high human disturbance had significantly greater variability explained by site-specific variables than low human disturbance islands ( $p < 0.05$ ). Modelled separately, either variable describing

**Table 10.** Proportion of variance in community dissimilarity (assessed by Raup-Crick null model dissimilarity) at the within island and within island group scales explained by site-specific environmental variables (Env: mean and interannual variability in sea surface temperature, live coral cover, and coral rugosity) or site-specific + island-specific variables (Total: Env + wave energy, reef area, island composition). Single predictor that explained the highest proportion of variation at each scale (Best) is also included. Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 10 sites from each island (within island scale) or 70 sites from each low or high human disturbance subset from within island group community dissimilarity matrices. H.A. = Hawaiian Archipelago; M.A. = Mariana Archipelago; A.S. = American Samoa; Rug = rugosity; Cover = coral cover. Asterisk indicates significant difference ( $p < 0.05$ ) among low and high human disturbance islands when islands pooled across island groups.

Island Group	Human Disturbance	Within Island		Within Island Group		
		Env*	Best	Env	Total	Best
H.A.	Low	0.472	Rug.	0.270	0.370	Rug.
	High	0.489	Cover	0.295	0.337	Cover
M.A.	Low	0.416	Cover	0.196	0.221	Cover
	High	0.513	Cover	0.175	0.198	Cover
A.S.	Low	0.356	Rug.	0.128	0.144	Cover
	High	0.459	Rug.	0.105	0.136	Rug.

the benthos (live coral cover or reef rugosity), as opposed to  $SST_V$  or  $SST_M$ , explained greater variability in site-site community dissimilarity on both low and high human disturbance islands at the within island group scale (Table 10). This trend held for data at the within island group scale (which uses comparisons both within and across islands within the same island group) (Table 10), and when data from different island groups were combined for both scales (Table 11 Table 12 respectively). The amount of variability explained by site-specific variables, island-specific variables, or a combination of both did not differ between islands with low or high human disturbance at the within island group scale however (Table 12;  $p > 0.05$ ).

Explained variability in site-site community dissimilarity at the largest scale, encompassing comparisons within island, within island groups, and across island groups, was substantially greater on islands with low human disturbance (Table 13). Site-specific variables explained more than twice as much variability on low human disturbance islands, which was driven by the high degree of variability that could be explained solely by  $SST_V$  (Table 13). Similarly, island-specific variables, wave energy, reef area, and island composition, could explain an additional 29% more variability on low human disturbance islands (Table 13). Island-specific and site-specific variables modelled together explained an additional 28% more variability on low human disturbance islands (Table 13). Island groups showed substantial clustering in ordination space (Figure 8A); surprisingly, human disturbance category only explained 5.5% of the variability when both low and high human disturbance islands were modelled together.

### **3.2.2 Abundance Based Community Composition**

The amount of variability in within-island site-site community dissimilarity, assessed using the  $\alpha$ -constraining  $\beta$ -deviations null model ( $\beta_W$ ), that was explained by site-specific variables ( $SST_V$ ,  $SST_M$ , live coral cover, and coral rugosity) was generally higher on populated islands in all three island groups (Table 14), however, not significantly so when all islands were considered irrespective of island group ( $p > 0.05$ ). On both high and low disturbance islands in all island groups, variables describing the benthic habitat (live coral cover or reef rugosity) explained more variability than  $SST_V$  or  $SST_M$  when modelled individually at the within island scale (Table 14). At the within island group

**Table 11.** Mean proportion of variability of community composition within islands explained by interannual variability of sea surface temperature ( $SST_v$ ), mean annual sea surface temperature ( $SST_m$ ), reef rugosity (rug.), live coral cover (cover), and a combination of all four (Env.). Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 10 sites from each low or high human disturbance island of Raup-Crick dissimilarity or  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ ) based dissimilarity matrices within islands.

<b>Model</b>	<b>Human Dist.</b>	<b>Env.</b>	<b><math>SST_v</math></b>	<b><math>SST_m</math></b>	<b>Rug.</b>	<b>Cover</b>
Raup-Crick	Low	0.434	0.113	0.112	0.135	0.142
	High	0.492	0.111	0.120	0.158	0.165
$\alpha$ - Constrained	Low	0.527	0.133	0.137	0.152	0.226
	High	0.580	0.126	0.145	0.213	0.238

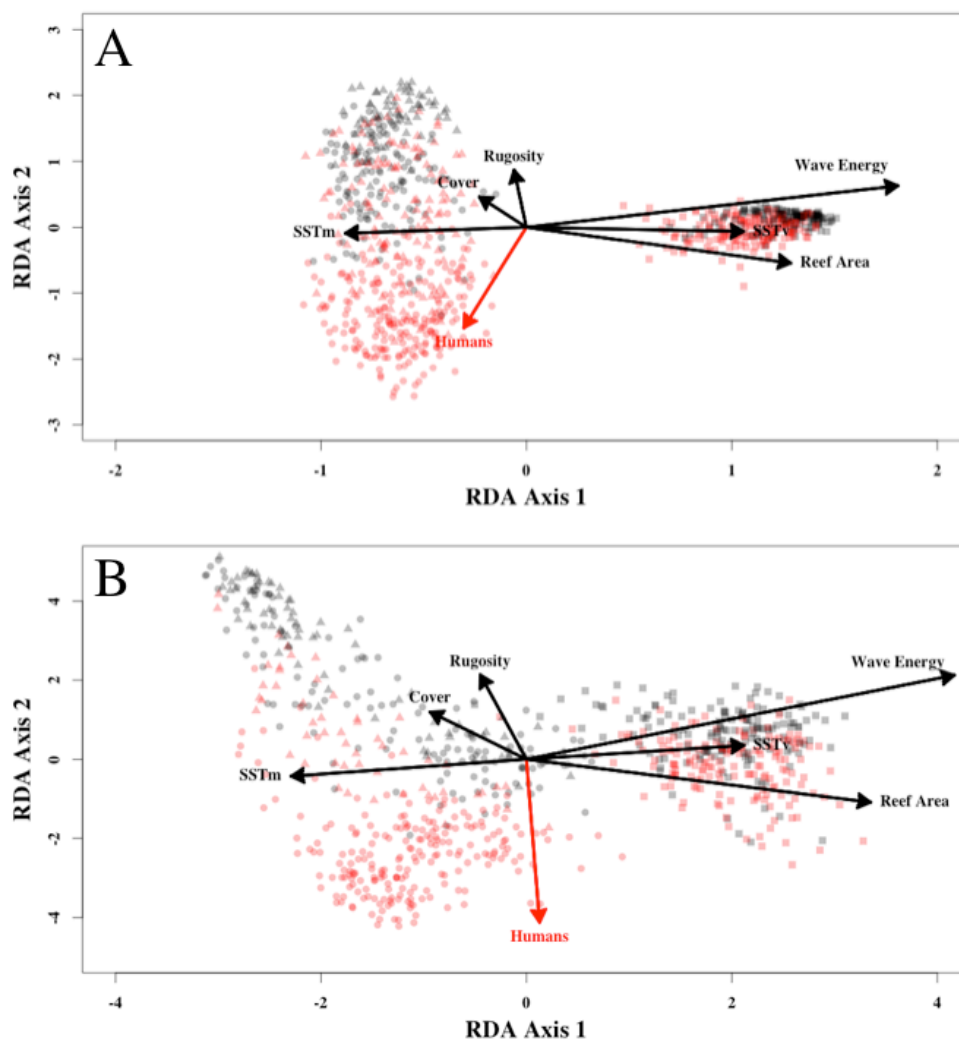
**Table 12.** Mean proportion of variability of community composition within island groups explained by site-specific variables (Env.), island-specific variables (Island), or both (Total). Site-specific variables include interannual variability and mean sea surface temperature ( $SST_v$  and  $SST_m$  respectively), reef rugosity (rug.), and live coral cover (cover). Island-specific variables include wave energy (Wave), reef area (Area) and island composition (Comp.). Also shown is the variance explained by each variable alone. Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 70 sites from either low or high human disturbance subsets of Raup-Crick dissimilarity or  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ ) based dissimilarity matrices at the within island group scale (which uses comparisons both within and across islands within the same island group).

<b>Model</b>	<b>Human Disturbance</b>								<b>Island</b>	<b>Wave</b>	<b>Area</b>	<b>Comp.</b>
	<b>Total</b>	<b>Env.</b>	<b><math>SST_v</math></b>	<b><math>SST_m</math></b>	<b>Rug.</b>	<b>Cover</b>						
Raup-Crick	Low	0.246	0.198	0.034	0.038	0.062	0.079	0.106	0.035	0.045	0.067	
	High	0.224	0.191	0.014	0.014	0.086	0.104	0.055	0.016	0.036	0.011	
$\alpha$ - Constrained	Low	0.256	0.181	0.050	0.051	0.040	0.066	0.128	0.049	0.061	0.092	
	High	0.233	0.199	0.024	0.021	0.068	0.104	0.051	0.023	0.031	0.008	

**Table 13.** Mean proportion of variability of community composition across island groups explained by site-specific variables (Env.), island-specific variables (Island), or both (Total). Site-specific variables include interannual variability and mean sea surface temperature ( $SST_v$  and  $SST_m$  respectively), reef rugosity (rug.), and live coral cover (cover). Island-specific variables include wave energy (Wave), reef area (Area) and island composition (Comp.). Also shown is the variance explained by each variable alone. Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 353 sites from either low or high human disturbance subsets of Raup-Crick dissimilarity or  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ ) based dissimilarity matrices at the across island group scale (which uses comparisons both within and across islands within the same island group, as well as across island groups).

Model	Human							Island	Wave	Area	Comp
	Disturbance	Total	Env.	$SST_v$	$SST_m$	Rug.	Cover				
Raup-Crick	Low	0.786	0.551	0.386	0.108	0.018	0.023	0.758	0.656	0.295	0.387
	High	0.510	0.216	0.003	0.098	0.035	0.022	0.468	0.377	0.207	0.064
$\alpha$ - Constrained	Low	0.477	0.368	0.231	0.083	0.014	0.040	0.433	0.349	0.175	0.184
	High	0.335	0.139	0.002	0.054	0.033	0.025	0.29	0.209	0.106	0.085

scale, which includes comparisons within islands and within island groups, the same trend was seen with the islands in American Samoa with low disturbance, which were best described solely by  $SST_M$  (Table 14). Combining data from the three island groups, live coral cover and reef rugosity tended to explain more variability than  $SST_V$  or  $SST_M$  at either of the within island group scales, and site-specific data always explained more than wave energy, reef area, or island composition (Table 12). Island composition tended to explain more variability than other island-specific variables at the within island group



**Figure 8.** Distance-based redundancy analysis ordination of Raup-Crick site-site community dissimilarity (A) and  $\alpha$ -constraining  $\beta$ -deviation based community dissimilarity (B) at the Pacific Ocean scale. Sites on low (black) and high (red) human disturbance islands during the second survey period (2011 – 2012) of Hawaiian Archipelago (squares), Mariana Archipelago (circles) and American Samoa (triangles) were used for ordination and are plotted. Length of arrow indicates magnitude of correlation. Humans = high disturbance islands;  $SST_m$  and  $SST_v$  = mean and interannual variability in sea surface temperature respectively; cover = live coral cover.

scale, however only for islands with low human disturbance, which tended to have a greater proportion of variability explained by island-specific variables (Table 12). At this scale, the amount of variability explained by site-specific variables, island-specific variables, or a combination of the two, was not significantly different among islands with low or high human disturbance ( $p > 0.05$ ; Table 12).

Explained variability in site-site dissimilarity at the across island group scale (encompassing comparisons within islands, within island groups, and across island groups) was substantially greater on islands with low disturbance (Table 13). Site-specific variables explained more than twice as much variability on remote islands, driven by the high degree of variability explained by  $SST_v$  at this scale (Table 13). Similarly, island-specific variables explained an additional 14% of the variability on islands with low disturbance compared to high disturbance islands (Table 13). Modelled together, site-specific and island specific variables explained an additional 23% of the variance on islands with low disturbance compared to high disturbance islands (Table 13). Island groups tended to show clustering in ordination space and sites on islands with

**Table 14.** Proportion of variance in community dissimilarity (assessed by  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ )) at the within island and within island group scales explained by site-specific environmental variables (Env: mean and interannual variability in sea surface temperature, live coral cover, and coral rugosity) or site-specific + island-specific variables (Total: Env + wave energy, reef area, island composition). Single predictor that explained the highest proportion of variation at each scale (Best) is also included. Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 10 sites from each island (within island scale) or 70 sites from each low or high human disturbance subset from within island group community dissimilarity matrices. H.A. = Hawaiian Archipelago; M.A. = Mariana Archipelago; A.S. = American Samoa; Rug = rugosity; Cover = coral cover;  $SST_m$  = mean sea surface temperature.

Island Group	Human Disturbance	Within Island		Within Island Group		
		Env.	Best	Env.	Total	Best
H.A.	Low	0.572	Rug.	0.233	0.364	Rug.
	High	0.596	Cover	0.303	0.337	Cover
M.A.	Low	0.504	Cover	0.152	0.187	Cover
	High	0.596	Cover	0.199	0.222	Cover
A.S.	Low	0.387	Cover	0.159	0.218	$SST_m$
	High	0.511	Rug.	0.095	0.142	Rug.

high disturbance tended to differentiate from sites on low disturbance islands (Figure 8B). Anthropogenic disturbance regime modelled alone could only explain 8.3% of the site-site community dissimilarity.

## 4. Discussion

In this thesis I provide evidence that islands that are highly disturbed by humans have altered  $\beta$ -diversity and that this effect is manifest at all three spatial scales studied. Unsurprisingly, and similar to other previous findings in other systems (Cassey et al. 2007), the effect of human disturbance on  $\beta$ -diversity is dependent on the scale of analysis. When considering only species incidence, high human disturbance islands are more spatially variable (higher  $\beta$ -diversity) when comparing sites within a single island or across islands within the same island group. In contrast, at the largest spatial scale comparing sites across island groups, highly disturbed islands have less spatial variability (lower  $\beta$ -diversity), consistent with biotic homogenization. Integrating abundances with species incidence, a consistent pattern of biotic homogenization is depicted across scales, as islands with high human disturbance are less spatially variable (lower  $\beta$ -diversity).

In addition to altered  $\beta$ -diversity, the community composition (species assemblage) at sites on islands with high human disturbance is less well explained by environmental variables, and seemingly more stochastically structured. This observation is consistent with or without the integration of species abundances, but given the high variability of reef fish community composition, is only revealed when using the statistical power of a large number of sites at the ocean basin scale of analysis. Contrasting this, when examining the community composition within individual islands only, those islands with



high human disturbance are better explained by site-specific predictor variables: primarily live coral cover or habitat rugosity. The ordination of sites using predictor variables results in considerable clustering of sites on islands with low and high human disturbance, indicating that community composition on the two categories of islands are fundamentally different.

Furthermore, and as previously documented, ignoring the potential biases of  $\alpha$ -diversity or abundance on  $\beta$ -diversity can yield patterns that are reflective of those biases (Kraft et al. 2011, Karp et al. 2012). Had abundance-based  $\beta$ -diversity been modelled based on Bray-Curtis dissimilarity, observed differences among islands with low or high human disturbance may not have been evident.

#### **4.1 Effect of Scale on Spatial Variability**

$\beta$ -diversity generally increased with spatial scale. The magnitude of increase was not, however, consistent across spatial scales. Increasing spatial scale from within island to within island group scale resulted in only a modest increase in  $\beta$ -diversity, whereas the increased spatial scale from within island group to across island group resulted in large increases in  $\beta$ -diversity.

A number of other coral reef fish  $\beta$ -diversity studies have found only slight increases in  $\beta$ -diversity across a variety of scales (Belmaker et al. 2008, Rodríguez-Zaragoza et al. 2011, Francisco-Ramos and Arias-González 2013), and connectivity among those scales is often proposed to be the driver of this result. Specifically, Francisco-Ramos and Arias-González (2013) found the  $\beta$ -diversity of coral reef fishes among connected regions to be lower than expected, and Cornell et al. (2007), finding similar results for corals, proposed that dispersal limitation was no different among two within island group scales. Recent studies of dispersal suggests that a large proportion of larval reef fishes return very close to natal reefs (Jones et al. 1999, Mora and Sale 2002), in many cases within 100 m of their birthing site (Jones et al. 2005). As such, the small proportion of larvae that do not settle near their natal reefs could be transported by oceanographic currents large distances within island or among connected island groups (Shanks 2009). Over evolutionary time

periods, community composition at sites within island groups would therefore become nearly as similar to one another as at sites within islands. As such, the slight increase in  $\beta$ -diversity with increasing scale from within island to within island group likely does not reflect strong dispersal limitations and may instead reflect increased variability of environmental variables across larger geographic areas.

The relatively large increase in  $\beta$ -diversity from within to across island group scales, with the exception of the Mariana Archipelago (described below in section 4.3), more likely reflects dispersal limitation in addition to environmental variability. While the three island groups share a modest portion of species, a larger portion is not shared, unsurprisingly as the island groups span two different biogeographically provinces (Kulbicki et al. 2013). As such, it is also not surprising that  $\beta$ -diversity is greatest at this scale. The limited connectivity and infrequency of dispersal events over such large distances (Cowen et al. 2006) would facilitate ecological drift amongst the island groups, resulting in species and community divergence over long time periods (MacArthur and Wilson 1967). Additionally, these three island groups experience considerably different environmental conditions (Gove et al. 2013) (described below in section 4.4), which could further exacerbate differences among them.

A small portion of the scale-dependent increase in  $\beta$ -diversity could also be an artefact of the null model methodology used. Within island and island group analyses were performed with regional pools set individually for each of the three island groups, whereas the across island group scale required the regional pool to encompass all three island groups. Because of this, the regional pool contained a greater diversity of species and therein a greater variety of species combinations with which to create null communities. These null communities may therefore have larger  $\beta$ -diversities among them, therein leading to larger  $\beta$ -deviations calculated from observed communities. However, because traditional  $\beta$ -diversity metrics ( $\beta_J$  and  $\beta_{BC}$ ) also show similarly large increases in  $\beta$ -diversity between the two scales, the effect of the null models inflating  $\beta$ -diversity at the largest scale is likely minimal.

## 4.2 Effect of Human Disturbance on Spatial Variability

Across all three scales of analysis, the effect of human disturbance on spatial variability of coral reef fish communities ( $\beta$ -diversity) was large and significant, albeit varied. Compared with that examine the effect of human disturbance on  $\beta$ -diversity and find changes in the order of 1-10% (i.e. Vileger et al. (2011)), I document reductions in  $\beta$ -diversity that are up the 24%. When species abundances were ignored and  $\beta$ -diversity was calculated based only on species incidences ( $\beta_{RC}$ ), the effect of human disturbance on spatial variability was scale-dependent. In contrast, when species abundances were integrated with species incidences ( $\beta_W$ ), islands with high human disturbance consistently had lower spatial variability of community composition.

Within islands, the spatial variability of species incidences ( $\beta_{RC,1}$ ) was greatest on islands with high human disturbance. These islands also had lower species richness and total fish abundance, and were missing many species of jacks and sharks: piscivorous fishes that are widely distributed but in low abundance and often targeted by fisheries (Friedlander and DeMartini 2002). Other species on those islands tended to be less cosmopolitan, occurring at a smaller proportion of the sites. In this study, increased  $\beta_{RC,1}$  may be reflective of the locally specific effect of human disturbances on species with low abundances (i.e. rare species), therein decreasing species richness and the proportion of singletons or doubletons, in addition to the well documented removal of moderately common targeted species (Williams et al. 2010). These changes, the local reductions of different rare species across an island and the removal of targeted species at many sites, would result in assemblages that have large differences in the species present, and therein greater  $\beta$ -diversity (Olden et al. 2003). This effect could be exacerbated if the removal of widely distributed targeted species that are in low abundance resulted in locally specific and variable changes to the diversity of untargeted species, as documented by Jennings and Polunin (1997).

When species abundances are considered, the effect of human disturbances on  $\beta_{W,1}$  differed considerably from  $\beta_{RC,1}$ . While the removal of widely dispersed species with low abundances (i.e. jacks, sharks, and other large piscivores) or locally specific rare species likely resulted in differences in species presences among sites and therein increased  $\beta_{RC,1}$ ,

similar abundances of common and shared species among sites likely drove  $\beta_{w,1}$  to be lower on islands with high human disturbance. As  $\beta_w$  uses Bray-Curtis dissimilarity to calculate  $\beta$ -deviations, it is driven more strongly by differences among highly abundant species, rather than those with low abundances. And as such, decreased  $\beta_{w,1}$  on islands with high human disturbance is therefore likely driven by similar abundances of widespread and at least moderately abundant species. Indeed, some groups of species, for instance those that feed on invertebrates, algae, or detritus, tend to thrive in response to human disturbances (Syms and Jones 2000, Wilson et al. 2006, Edwards et al. 2013). I speculate that increased similarity in abundances of those species, particularly low carnivore species that tended to be an increased proportion of the species present on islands with high human disturbance, drove decreased  $\beta_{w,1}$  where human disturbance is high.

Interestingly, at the within-island scale, a larger proportion of spatial variability in community composition was accounted for by site-specific environmental variables, primarily live coral cover and rugosity, on islands with high human disturbance. This is likely reflective of two mechanisms: (i) the relatively low variability in SST at the within-island scale, and (ii) the greater number of sites with low live coral cover and rugosity on islands with high human disturbance. At sites with low live coral cover or rugosity, the species present are likely a predictable subset of species (Bell and Galzin 1984). This subset would not include many species that would otherwise persist if live coral cover or rugosity were high. The species present at those sites would therefore tend to be well predicted by live coral cover and rugosity, and therein inflate the proportion of variability accounted for by environmental variables on those islands. This trend held only when species abundances were not considered. When species abundances were considered there was no difference between islands with low or high human disturbance, possibly owing to high variability of the abundances of species present at sites with low live coral cover or rugosity.

At the within island group scale, the trend reversed, as the spatial variability of community composition accounted for by environmental variables was greater on islands with low human disturbance, regardless of whether or not abundances were integrated. Unfortunately, the low statistical power at this scale (i.e. only three island groups)

prohibits me from concluding if the result was statistically significant. Spatial variability in coral reef fish community composition has previously been associated with environmental variables (Arias-González et al. 2008). A decreased proportion of variability explained by environmental variables on islands with high human disturbance could reflect increased stochasticity of community control on those islands, as observed in disturbed terrestrial systems (Vellend et al. 2007). Indeed, Syms and Jones (2000) attributed a high degree of variability in coral reef fish community composition to stochastic variation, and because human disturbances alter habitat characteristics in addition to direct removal of fishes (Gardner et al. 2003), it is possible that environmental determinism of assemblages is less strong on human disturbed islands. It is also possible that spatially random removal of species on islands with high human disturbance results in species assemblages that differ considerably among one another but have relatively similar environments. Indeed, it appears that the nature of human disturbances on high disturbance islands was spatially patchy, as indicated by the similar proportion of variation accounted for by environmental variables around whole highly disturbed islands island or at areas near large human population on those islands (Appendix E). In this case, rather than reflecting increased stochasticity of community control, reduced explained variability in community composition would be a result of the patchy nature of human disturbance (Kingsford et al. 1991).

In either case, increased stochasticity or spatially random removal of species on islands with high human disturbance would result in greater incidence-based  $\beta$ -diversity at this scale ( $\beta_{RC,2}$ ), which is indeed what is observed. Similar to trends at the within island scale, abundance-based  $\beta$ -diversity at the within island group scale ( $\beta_{W,2}$ ) was unlike incidence-based  $\beta$ -diversity. That is,  $\beta_{W,2}$  was lower on islands with high human disturbance. The magnitude of effect of human disturbances on  $\beta_{RC}$  was similar at the within island and within island group scales.  $\beta_W$ , on the other hand, was impacted to a greater degree by human disturbances at the within island group scale. The increased magnitude of effect may be reflective of removal of similar widespread but low abundance species of fishes (i.e. targeted species) across islands within each island group. As these species are widespread, they would otherwise be shared across islands. While

their removal is still likely somewhat variable within islands, therein increasing  $\beta_{RC}$  at both within island group scales, consistent reductions of abundances of similar types of fishes (e.g. large piscivorous fishes), in addition to moderate abundances of widespread common species across islands with high human disturbance would depress  $\beta_{W,2}$ .

At the largest scale, comparing sites that are in different island groups, separated by vast distances across open oceans, trends consistent with biotic homogenization are most evident. At this scale, islands with high human disturbance have substantially lower incidence-based and abundance-based  $\beta$ -diversity ( $\beta_{RC,3}$  and  $\beta_{W,3}$  respectively). At this scale, sites share many fewer species than at within island group scales and therefore any process that reduces the number of unshared species or increases the number of shared species, even subtly, could significantly decrease  $\beta$ -diversity. The reduction of large targeted species likely contributes heavily to the increased similarity (decreased  $\beta$ -diversity) of islands with high human disturbance. Many large fishes (i.e. jacks and sharks) were present in much greater abundances on islands with low human disturbance, consistent with previous studies (Williams et al. 2010, Nadon et al. 2012). Those groups of fishes were often both spatially variable (infrequently observed, likely due to high mobility) and contained species unique to one or two island groups. The removal of these species on islands with high human disturbance would represent the loss of fishes that were unshared among the three island groups, and therein decrease  $\beta$ -diversity on those islands. In addition, introduced invasive species, such as the intentional introduction of the peacock grouper *Cephalopholis argus* in the Hawaiian Archipelago (Randall 1987), could further increase the similarity of islands with high human disturbance. If invasive species are present in multiple island groups, which is the case for *C. argus*,  $\beta$ -diversity would decrease on islands where invasive species are present. While *C. argus* was present on both islands with low and high human disturbance in the Hawaiian Archipelago, they were observed at 8x more sites and were also 8x more abundant on islands with high human disturbance in the Hawaiian Archipelago (i.e. thus contributing to the lower  $\beta_{W,3}$  of highly disturbed Hawaiian islands relative to those with low human disturbance). Finally, islands with high human disturbance likely had increased, and more similar abundances of a number of common species, such as territorial damselfishes

(Edwards et al. 2013). The cumulative effect of the loss of species that are infrequently observed (i.e. spatially variable within an island group) but present in multiple island groups (i.e. shared) large fishes, introduction of invasive species, and increase in abundance of widespread species results in large-scale homogenization of fish assemblages. While this effect has been documented frequently in terrestrial systems (Olden et al. 2006), this study represents the first evidence of homogenization of coral reef fish communities over large spatial scales.

In addition to having lower  $\beta$ -diversity at the across island group scale, the association between the variability of community composition and environmental variables on islands with high human disturbance was much weaker, largely driven by strong correlations between community composition and sea surface temperature or island-specific variables on islands with low human disturbance. This likely reflects one of two possibilities, similar to the within island group scale: increased stochasticity of species assemblages on islands with high human disturbance, or is a result of widespread but locally variable (i.e. within an island or island group) removal of species resulting in sites with similar environmental variables but different species assemblages. In either case and similar to the within island group scale, coral reef fish communities on islands with high human disturbance are more similar at large scales, but harder to predict spatially based on environmental variables.

Associated with changes to  $\beta$ -diversity and community composition, I documented functional group level changes to community composition at sites with high human disturbance. Similar to frequently observed reductions in biomass of large predators where human disturbance is high (Williams et al. 2010, Nadon et al. 2012), I document decreases in the proportion of species that piscivorous. That is, there is lower numerical diversity (i.e. richness) of different species of piscivorous fishes at sites with high human disturbance. These sites also have a greater proportion of low carnivores (invertivores, small piscivores, and corallivores), and of the species that are present on islands with high human disturbance but not low human disturbance, the largest proportion are low carnivores. Additionally, these species are substantially smaller, likely reflecting the loss of large fishes in these areas.

### 4.3 Regional Differences in Spatial Variability

Across most scales of analysis, for incidence-based and abundance-based  $\beta$ -diversity ( $\beta_{RC}$  and  $\beta_W$  respectively), there were regional differences in the magnitude of  $\beta$ -diversity of islands. Generally, the Hawaiian Archipelago had lower within island  $\beta$ -diversity and greater across island group  $\beta$ -diversity. The Mariana Archipelago had greater abundance-based, but lower incidence based  $\beta$ -diversity at the within island group scale.

Consistent with trends of decreased species richness with increased distance from the Indo-Pacific Coral Triangle (IPCT) (Stehli and Wells 1971, Connolly et al. 2005), within island  $\beta$ -diversity ( $\beta_1$ ) was lower in the Hawaiian Archipelago than the Mariana Archipelago or American Samoa. Interestingly, both  $\beta_{RC,1}$  and  $\beta_{W,1}$  account for biases in local species richness that would otherwise tend to make depauperate sites (i.e. those in the Hawaiian Archipelago) appear to have *higher*  $\beta$ -diversity using traditional metrics. The models do not account for differential *regional* species richness, however. As  $\beta$ -diversity generally correlates strongly with regional richness ( $\gamma$ -diversity) for corals and coral reef fish (Belmaker et al. 2008), it is likely that low regional richness of the Hawaiian Archipelago is in part driving lower  $\beta$ -diversity in that region.  $\beta$ -diversity in the Hawaiian Archipelago tended to be greater than the other two island groups at the across island group scale. As the Hawaiian Archipelago is in a different biogeographical province than the other two island groups (Kulbicki et al. 2013), it shared the fewest number of species with the Mariana Archipelago or American Samoa, which shared a large proportion of their species. As such, it is not surprising that the Hawaiian Archipelago was more differentiated at this scale and therein had higher  $\beta$ -diversity. Interestingly, within island groups, the Hawaiian Archipelago had lower abundance-based  $\beta$ -diversity ( $\beta_{W,2}$ ) than the Mariana Archipelago but relatively similar incidence-based  $\beta$ -diversity ( $\beta_{RC,2}$ ). This indicates that relatively abundant species were more consistently similar across islands of the Hawaiian Archipelago compared to the Mariana Archipelago. In contrast, the proportion of rare species across islands in the two island groups was relatively similar. This is supported by the similar proportion of singletons and doubletons on islands in each of the two island groups.



Interestingly, the Mariana Archipelago shows somewhat different trends in  $\beta_w$  across scales compared to the other two island groups. While  $\beta_w$  of both the Hawaiian Archipelago and American Samoa increase with increasing scale,  $\beta_{w,1}$  and  $\beta_{w,2}$  were similar magnitudes for the Mariana Archipelago, and  $\beta_{w,3}$  was lower. This indicates that, in the Mariana Archipelago the spatial variability within islands tends to be just as high as that among islands, likely reflecting high connectivity of the islands. The decreased magnitude of  $\beta_{w,3}$  may reflect the high number of species the Mariana Archipelago shares with other island groups relative its the high regional richness. As such, there is a large number of species that can vary within the island group, therein increasing  $\beta_{w,1}$  and  $\beta_{w,2}$ , but also a large number of species that are shared with either the Hawaiian Archipelago or American Samoa, therein decreasing  $\beta_{w,3}$ . Additionally, as the Mariana Archipelago spans a large latitudinal gradient, there is likely a significant degree of turnover among species present in the southern compared to northern islands, also increasing  $\beta_{w,1}$  and  $\beta_{w,2}$  but not  $\beta_{w,3}$ .

American Samoa also shows trends that are not consistent with trends observed in the Hawaiian Archipelago and Mariana Archipelago.  $\beta_{w,1}$ , which was greater on islands with low human disturbance in both the Hawaiian Archipelago and Mariana Archipelago, was greater on islands with high human disturbance in American Samoa. Williams et al. (2010) similarly found different trends for American Samoa compared to the Hawaiian Archipelago and Mariana Archipelago, and attributed the inconsistencies to limited differences in human impacts among islands categorized as either low or high human disturbance. Indeed, unlike other islands categorized as low human disturbance, Swains has a permanent population and Rose is nearer than most other low human disturbance islands. Swains and Rose also have little reef area compared to other islands, and therefore human disturbances on those islands may have a larger impact on fish communities.

The proportion of variability in community composition explained by environmental variables at the within island and within island group did not differ substantially among the three island groups. To a small extent, a greater proportion of variability was explained in the Hawaiian Archipelago, however only marginally so. The smaller

regional pool and lower local richness in the Hawaiian Archipelago could result in more predictable species assemblages across the island group. Similarly, the mean proportion of sites each species was observed at was higher in the Hawaiian Archipelago, and there were numerically fewer rare species (singletons or doubletons), a trend observed in less diverse reef fish assemblages (Guillemot et al. 2011). The relative scarcity of low abundance species and lower local and regional richness could all contribute to a more predictable species assemblages in the Hawaiian Archipelago, and therefore increased variability in community composition explained by environmental variables.

#### **4.4 Effect of Environmental Variables on Spatial Variability**

Island group and human disturbance were the only explanatory variables consistently associated with  $\beta_{RC}$  or  $\beta_W$  across the three spatial scales. Live coral complexity was only a strong predictor of  $\beta$ -diversity at the smallest scale, and at this scale no other environmental predictor consistently influenced  $\beta$ -diversity. At the within and across island group scales, wave energy and sea surface temperature were strong predictors of  $\beta$ -diversity; reef area was rarely a strong predictor of  $\beta$ -diversity at any scale.

As reef fish community composition often correlates strongly with coral identity and cover (Sale 1977, Arias-González et al. 2008, Messmer et al. 2011), it is not surprising that within island  $\beta$ -diversity was influenced strongly by live coral complexity. To the same end, habitat rugosity and live coral cover, the two variables underpinning live coral complexity, accounted for the greatest variability in spatial variability of community composition within islands. Surprisingly, however, increased live coral complexity was associated with decreased  $\beta_{RC,i}$  on all islands, and increased  $\beta_{W,i}$  only on islands with high human disturbance. Loss of live coral complexity was predicted to act as an environmental filter, whereby only a subset of species present regionally would persist at sites with low live coral complexity. While loss of live coral complexity likely caused reductions in some species that associate closely with coral (Jones et al. 2004), other fishes have species-specific preferences for corals (Messmer et al. 2011) and similar magnitudes of reductions in live coral complexity might not reflect reductions in specific species. As such, low live coral complexity might be related to greater  $\beta_{RC,i}$  because sites with low live coral complexity may harbour distinct assemblages of rare fishes depending

on the species of coral present. On islands with high live coral complexity and human disturbance, increased  $\beta_{w,1}$  is likely driven by spatial variability of fishes with moderate or high abundances. Islands with high human disturbance as well as relatively high live coral complexity likely have high degree of variability in live coral complexity among sites. Therefore, it is probable that the fishes that inhabit sites with low or high live coral complexity likely differ. For instance, large mobile fishes that tend to have larger home ranges (Kramer and Chapman 1999) are likely more abundant at sites where coral complexity is higher and prey are more available. Sites with low live coral complexity, on the other hand, are likely dominated by different fishes, such as those feeding on invertebrates, detritus, or plankton (Wilson et al. 2006). On islands with low human disturbances, live coral complexity may not drive  $\beta_{w,1}$  for two reasons. First, habitat rugosity, a component of live coral complexity, is greater (Appendix F) and therefore even islands with relatively low live coral complexity could support rather complex fish assemblages. Secondly, if prey are abundant, large fishes may thrive in both sites with high and low live coral complexity (Friedlander and Parrish 1998, Connell and Kingsford 1998). While high live coral complexity may harbour greater prey abundances (Jennings et al. 1996), prey also have greater opportunity to hide from predators, and predators can be more efficient where prey refuge is lower. In general, live coral complexity was associated with within island  $\beta$ -diversity in a non-intuitive way, but was not associated with  $\beta$ -diversity at larger scales. Historically, high live coral complexity was likely not limiting at any moderately large scale (i.e. island-scale or larger) and therefore has not limited large-scale distributional patterns of species.

Wave energy was expected to influence  $\beta$ -diversity at all scales, but was only strongly associated with incidence-based  $\beta$ -diversity at the within and across island groups scales ( $\beta_{RC,2}$  and  $\beta_{RC,3}$  respectively). At the within island group scale, increased wave energy was associated with increased  $\beta_{RC}$ , as hypothesized. Likely driving increased  $\beta_{RC,2}$  on islands with high wave energy were distinct assemblages of fishes that associate with high wave energy. Fish fin morphology generally relates to the wave energy of their habitat (Fulton et al. 2005), and distributional differences of species with certain swimming modes have previously been associated with habitat wave energy (Fulton and Bellwood 2005).

Interestingly, wave energy only accounted for a large portion of the variability in community composition explained by environmental variables at the across island group scale. Decreased  $\beta_{RC,3}$  associated with high wave energy was not expected, but could be related to a relative absence of rare fishes where wave energy is high. If high wave energy represents an ecological filter where only some species can persist, then there might be fewer rare species at islands with high wave energy (i.e. Maug in the Mariana Archipelago or Kure and Pearl & Hermes in the Hawaiian Archipelago), which would have the effect of increasing similarity across the ocean basin scale. Alternately, driving this association could have also been the high negative correlation among wave energy and mean sea surface temperature ( $SST_{mean}$ ): islands with low  $SST_{mean}$ , such as northern islands of the Hawaiian Archipelago and Mariana Archipelago, also had high wave energy. Had site-specific wave energy measurements been available rather than using temporal variability in wave energy as a proxy, it may have been possible to disentangle whether or not the association was driven by correlations with  $SST_{mean}$ . Furthermore, site-specific wave energy may have accounted for a larger proportion of variability in community composition at the within island and island group scale.

Sea surface temperature was strongly associated with both incidence and abundance based  $\beta$ -diversity ( $\beta_{RC}$  and  $\beta_W$  respectively), but only to a great extent at the within and across island group scales. The interannual variability of sea surface temperature ( $SST_{var}$ ) was generally more strongly associated with  $\beta$ -diversity than mean sea surface temperature. At the within island group scale, both  $\beta_{RC}$  and  $\beta_W$  were greater on island with lower  $SST_{var}$ , as hypothesized. Higher  $SST_{var}$  could result in more frequent or more extreme temperature stresses that greatly reduce the diversity of species present (Riegl 2002), and therefore those areas would be likely to be populated by similar subsets of species that are resilient to temperature extremes or fluctuations. At this scale, increased  $SST_{mean}$  was related to greater  $\beta_{RC}$ .  $SST_{mean}$  correlates strongly with species richness (Gaston 2000, Mellin et al. 2010, Tittensor et al. 2010), and within island groups the species that populate islands with low  $SST_{mean}$  would be a general subset of the species present in the whole island group. As such, there would be fewer species with which to populate those islands, and likely less variability and stochasticity of species composition on those islands. Across island groups, the trend reverses, and  $\beta_{RC}$  and  $\beta_W$  were greater at

islands with high  $SST_{var}$ . Driving this trend could be different subsets of species that are resilient to temperature fluctuations in the three island groups. If the species resilient to large temperature fluctuations differ in the three island groups,  $\beta_{RC,3}$  and  $\beta_{W,3}$  would be greater on islands with high  $SST_{var}$ . Likewise, if species that are associated with more stable SST are generalist species, they may be more likely to be shared across island groups and therein decrease  $\beta_{RC,3}$  and  $\beta_{W,3}$  at islands with low  $SST_{var}$ .  $SST_{var}$  only accounted for a large portion of the variability in community composition at the across island group scale. This could be an artefact of the increased statistical power at this scale, or the greater variability in  $SST_{var}$  that is present across the Pacific basin rather than within individual island groups. At no scale did  $SST_{mean}$  explain a very large portion of the variability accounted for by environmental variables.

At no scale was reef area consistently and strongly associated with  $\beta$ -diversity nor did it explain a large portion of the variability in community composition. This measure was an imprecise measure of reef area, reflecting the area around islands within the 10 fathom depth line (Rohmann et al. 2006). Had a more precise measure of reef area been available, based on observed habitat, reef area may have been more strongly associated with  $\beta$ -diversity, as it is often a strong predictor of coral reef local and regional richness (Bellwood et al. 2005). Reef area was predicted to influence  $\beta$ -diversity because of its influence on habitat heterogeneity and regional richness. Since habitat heterogeneity is also integrated into live coral complexity and other variables influence regional richness (i.e. SST or wave energy due to their potential influence as ecological filters), these influence of reef area alone on  $\beta$ -diversity may have been weak compared to other variables.

#### **4.5 Implications and Concluding Remarks**

The homogenization of biota can have wide ranging consequences, from reduced ecosystem stability and functioning to altered patterns of evolution (Olden et al. 2004). Only recently have human induced changes in  $\beta$ -diversity been documented in systems beyond plant and freshwater fish communities; this study is one of the first to relate human disturbances to altered  $\beta$ -diversity of marine taxa. It also provides the first

evidence of large-scale homogenization of coral reef fish assemblages. To prevent further homogenization, incorporating management and conservation of  $\beta$ -diversity should be considered.

Early in the resurgence of  $\beta$ -diversity research, both Condit et al. (2002) and Legendre et al. (2005) noted the importance of  $\beta$ -diversity for management and conservation efforts. Because  $\beta$ -diversity reflects processes that act over larger spatial scales, management plans that consider  $\beta$ -diversity can help protect those processes and also conserve large portions of biodiversity. Further, management of areas with high  $\beta$ -diversity increases the likelihood that patterns generating high  $\beta$ -diversity are maintained, and importantly, regional resilience is protected. In areas where  $\beta$ -diversity is low, environmental changes (e.g. increased frequency of stressful hot water events) or epidemics (e.g. coral disease outbreaks) that cause the reduction of certain species could lead to similar changes to species composition at sites throughout geographical large areas. Since different species are not present in the regional pool to colonize sites that experience further disturbance (because  $\beta$ -diversity is low), changes in community composition could be exacerbated and phase-shifts or further site degradation may be more likely imminent.

Unfortunately, preserving areas of high  $\beta$ -diversity often requires areas far from human disturbances (Miller and Hobbs 2002), or large areas near human settlements, a luxury not typically afforded. When vast stretches of sea or land are not available, assessing sites' "local contribution to  $\beta$ -diversity" (LCBD) (Legendre and De Cáceres 2013) could be of aid. The LCBD of sites reflects how similar or dissimilar particular sites are from all other sites, and therein shows the contribution of sites to the overall  $\beta$ -diversity of an area. Prioritizing the protection of sites with high LCBD could help preserving a large portion of  $\beta$ -diversity while reducing area requirements.

Unfortunately, as this study indicated, over large scales coral reef fish communities are more stochastically driven where human disturbances pervade. In communities where species closely associate with environmental characteristics, protection of areas based on species assemblages and environmental characteristics could yield communities that are relatively stable over time. However, where environmental determinism is low,

community composition might be more likely to vary over time. Therefore, areas that were selected for protection at one time may stochastically change and no longer be of great importance for protection at a later time. The relative lack of environmental determinism also makes selection of areas for protection more difficult: since specific habitat characteristics are less associated with fish communities, using environmental data to select pertinent habitat to conserve is less likely to harbour high biodiversity or specific species assemblages.

With biotic homogenization added to the list of deleterious anthropogenic driven changes to coral reef ecosystems, protection of the biodiversity that remains is essential. The implementation of vast marine protected areas (MPAs), such as the Papahānaumokuākea Marine National Monument, the Marianas Trench National Monument, and the National Marine Sanctuary of American Samoa that covers low human disturbance islands of the Hawaiian Archipelago, Mariana Archipelago, and American Samoa could indeed help to protect  $\beta$ -diversity and biodiversity in general. However, human impacts are greatest where people live, and to maintain the biodiversity that humanity both desires and requires, large protected areas near human population are necessary. Sufficiently large protected areas, in addition to prioritizing management of areas with high  $\beta$ -diversity, provide our greatest opportunity to protect both the processes that drive biodiversity as well as the multitude of species that enhance it.

## Appendices

### Appendix A: Null Model Code

Null model is also available as sourceable R script online at:

[github.com/baumlab/betadiversity/](https://github.com/baumlab/betadiversity/)

```
### <><><><><><><><><><><><><><><><> ###
### Abundance based null model ###
### <><><><><><><><><><><><><><> ###

##### The idea for this abundance based null model is as follows:
# 1 # Create a regional pool that is a vector of all the species, each repeated X number
##### of times equal to the number of times it is present in all sites summed over the
##### region. Each element in the vector is an individual of a given species.
# 2 # Create an empty species X site matrix.
# 3 # Choose an individual from the regional pool at random, and stick it in a site.
##### Repeat.
# 4 # When you choose a species, Z, that is to be placed in site y, where sites y has
##### already reached its observed alpha richness and Z is a new species for site y, find
##### a different site for Z to occupy. If all sites have reached their alpha richness
##### and there is still a (rare) species that has no home (i.e. Z), choose a site at
##### random and kick out an abundant species from that site, placing Z in its place. The
##### displaced individuals will be placed into sites where that species is already #####
##### present.
# 5 # Calculate the null beta diversity among all sites, compare to observed. The
##### function expects a site X species matrix, with site name in the first column, and a
##### vector of equal length as # sites that is for the weighting of sites (if
##### unequal_weight_sites=T)

abundance_null_1<-function(spxsite, weight_vector=FALSE,
                          unequal_weight_sites=FALSE, site_names_in_col1=TRUE,
                          reps=99, metric="bray", n_abund=5, n_rare=10){

# Necessary for calculation of Bray-Curtis or other abundance based metric:
require("vegan")

# This section sets the row names of the spXsite matrix, and removes that column:
if(site_names_in_col1){
  row.names(spxsite)<-spxsite[,1]
  spXsite<-spxsite[,-1]
}
spxsite<-as.matrix(spxsite)

# This section removes and columns (species) that have a total regional abundance of
# zero:
spxsite<-spxsite[,colSums(spxsite)>0]

# Save this original siteXsite matrix
original_spXsite<-spxsite

# This section calculates the site richness and total abundance, which will be used
# later:
site_richness<-rowSums(spxsite>0)
site_abundance<-rowSums(spxsite)

# This section renames the columns to numbers to make the computation slightly quicker:
colnames(spxsite)<-c(1:ncol(spxsite))

# This section multiplies through the weighting vector if present:
if(unequal_weight_sites){
  spXsite<-t(t(spxsite)*rep(weight_vector,each=ncol(spxsite)))
}

# This section ensures any species that was seen on a survey retains its presence, even
# if after weighting its abundance is <1. Any species with an abundance <1 after the
# weighting gets a value of 1, and is removed from the species totals (regional pool)
# that is to be "un-weighted":
spxsite[which(spxsite<1 && spXsite>0)]<-1 # any species who had an ab <1 get ab = 1
rare<-ceiling(which(spxsite==1)/nrow(spxsite)) # extract those rare species (ab <=1)
spxsite[which(spxsite==1)]<-0 # remove those extreme rare species from spXsite
```



```

# This line calculates the weighted regional pool species abundances:
speciestotals<-apply(spXsite, MARGIN=2, FUN=sum)

# This section creates the regional species pool, rounding non-integer weighted species
# abundances to integers. The first lines is to make the regional pool of approximate
# same number of individuals as the original data. The "if" function removes or adds some
# species if n_original != n_weightedregion, however the lines above the "if" ensure
# that only relatively non-rare species (set by n_rare=) are the ones whose added or
# removed:
n_original<-round(sum(colSums(original_spXsite))) # original matrix total abundance

# re-weights regional pool so total ab ~ = original ab:
roundedtotals<-speciestotals/(sum(speciestotals)/n_original)

# if after down-weighting region, total ab <1, gets 1:
roundedtotals[which(roundedtotals<1)]<-1

# removes somewhat rare species from chance of being added/removed after potential
# rounding "errors":
rare2<-roundedtotals[which(roundedtotals<n_rare)]
rare2<-round(rare2,digits=0)

# create vector of "less-rare" (ab<n_rare) species regional pool:
rare2<-as.numeric(rep(names(rare2),rare2))
roundedtotals<-roundedtotals[which(roundedtotals>=n_rare)]
roundedtotals<-round(roundedtotals, digits=0)

# create vector of of rest of species regional pool:
regionalvector<-rep(names(roundedtotals),roundedtotals)

# combine regional pool of very rare (ab<1) and "less-rare" (ab<n_rare):
rare<-c(rare,rare2)

# calculate how many species are in total regional pool:
n_weightedregion<-(length(regionalvector)+length(rare))

# This "if" control ensures that the weighted regional pool has the same number of
# individuals as the original matrix
if(n_weightedregion>n_original){
  regionalvector<-sample(x=regionalvector,n_original-length(rare))
}

if(n_weightedregion<n_original){
  regionalvector<-c(regionalvector,sample(x=regionalvector,(n_original-
n_weightedregion)))}

regionalvector<-c(rare,regionalvector) # now n_original == length(regionalvector)
regionalvector<-sample(regionalvector) # shuffle order of regional vector

# This section will begin to fill an empty matrix, based on a few rules:
# 1. An individual, selected at random, will be placed in a site, selected at random,
# so long as it being added to that site does not make alpha(null) > alpha(obs)
# 2. If all sites have reached their observed alpha and species need to be placed, a
# site will be selected at random, and all individuals of one of the more abundant
# species (1 in n_abund most abundant species at that site) will be removed from the
# site to allow the original species to be placed. The displaced individuals will
# randomly be placed into sites after all individuals in the regional pool have been
# placed, where they are only to be placed in sites where they are already present.

site_vector<-1:nrow(spXsite)

# the array to fill with null beta deviation values:
all_null_sim1<-array(0,dim=c(nrow(spXsite),nrow(spXsite),reps))
all_null_sim2<-array(0,dim=c(nrow(spXsite),nrow(spXsite),reps))

# calculate how many individuals are at each site, each rep:
cumulative_site_abundance<-rep(0,nrow(spXsite))

# ensure there are no deviations in alpha(null)==alpha(obs):
null_site_richness_diff<-rep(0,nrow(spXsite))

# used to see if / how many repetitions are discarded (if alpha(null)<alpha(obs)):
iteration_count<-0

pb<-txtProgressBar(min=0,max=reps,style=3) # For progress bar
for(h in 1:reps){

# repeat is to allow repetitions where alpha(null)<alpha(obs) are discarded:
  repeat {
# create null matrix:

```

```

null_matrix<-matrix(data=0,ncol=ncol(spxsite),nrow=(nrow(spxsite)))
unplaced<-vector() # where displaced individuals are kept

for(i in 1:length(regionalvector)){ # for each individual in the regional pool...
# randomize the selection order of which site to try placing individual "species" in:
site_vector<-sample(site_vector)
species<-as.numeric(regionalvector[i]) # what species is the individual?

for(j in 1:length(site_vector)){ # try to place the individual in site[j]
site<-site_vector[j]

# If rule one is satisfied, individual will be placed
if(sum(null_matrix[site,]>0)<site_richness[site] ||
null_matrix[site,species]>0){
null_matrix[site,species]<-null_matrix[site,species]+1

break # move to next individual if it was placed
}

# If rule one is not satisfied, and there are still sites that have yet to be tried, move
# onto next site
if(sum(null_matrix[site,]>0)==site_richness[site] &&
null_matrix[site,species]==0){

# try a new site to place that individual (if there are still sites to try):
if(j<length(site_vector)) {next
}

}

# If all sites have been tried and the individual cannot be placed (i.e. it would cause
# alpha(null)>alpha(obs) at a site), use this loop to displace individuals of an abundant
# species at a random site.

# this for loop is only used if no sites can handle that individual:
if(j==length(site_vector)){
site2<-sample(nrow(null_matrix),1) # randomly select a site

# choose one of the n_abund most abundant species to remove:
species_to_remove<-which(rank(null_matrix[site2,],ties.method="random") ==
sample(ncol(null_matrix):(ncol(null_matrix)-n_abund),1))
n_remove<-null_matrix[site2,species_to_remove]
speciesXn<-rep(species_to_remove,n_remove)
null_matrix[site2,species_to_remove]<-0 # remove that species

# add the original species:
null_matrix[site2,species]<-null_matrix[site2,species]+1

# add the displaced individuals to the vector:
unplaced<-c(unplaced,speciesXn)
}
}

# This section places the displaced ("unplaced") after all individuals in regional pool
# have been placed:
if(length(unplaced)>0){

# For each species that has been displaced...
for(k in 1:length(unique(unplaced))){

# What species is that individual?
species2<-unique(unplaced)[k]

# How many individuals are there of that species?
n_species2<-sum(unplaced==species2)

# In what sites is that species present?
sites_where_k_present<-which(null_matrix[,species2]>0)

# If it is present in more than one site, choose how many will be placed in each site:
if(length(sites_where_k_present)>1){

# Could include prob of a displaced individual being placed in a site relative to that
# species abundance at that site, or abundance of all species at that site, currently not
# included:
sites_to_place<-sample(sites_where_k_present,n_species2,replace=T)

# If there is only one site where that species is present, place all individuals (of that
# species) there.
} else {

```

```

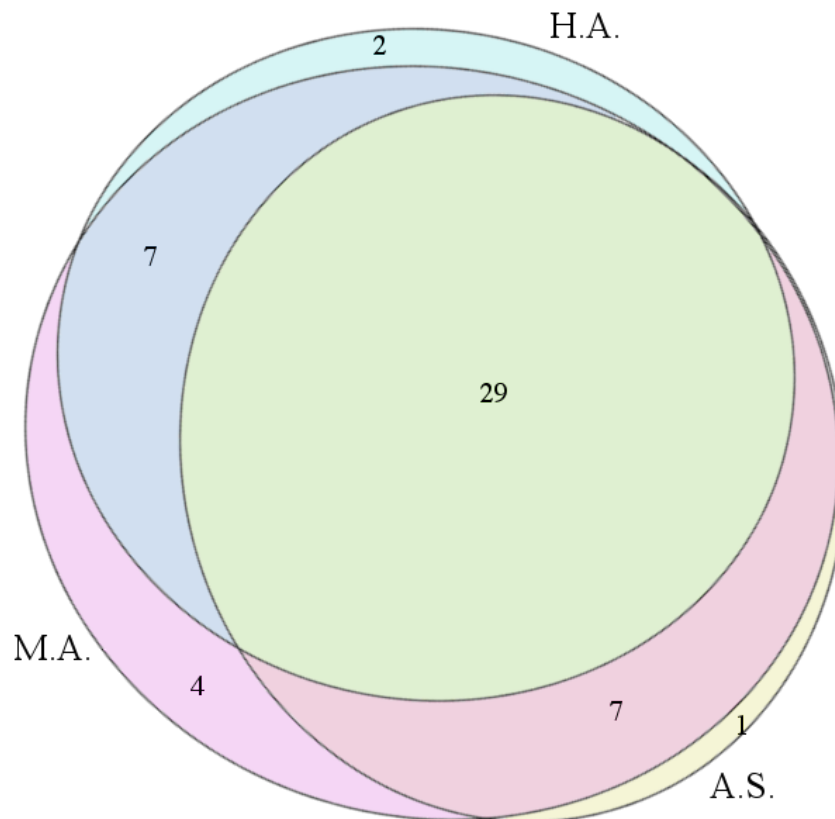
        sites_to_place<-rep(sites_where_k_present,n_species2)
    }
# Place the species:
  for(l in 1:length(unique(sites_to_place))){
# the site chosen to add x number of individuals to:
  site3<-unique(sites_to_place)[l]
# X number of individuals:
  n_species_to_add_site3<-sum(site3==sites_to_place)
  null_matrix[site3,species2]<-null_matrix[site3,species2] +
    n_species_to_add_site3
  }
}
# If site_alpha(null)==site_alpha(obs), calculate null beta and break out of repeat loop,
# continuing onto next repetition. Otherwise, repeat and do not calculate beta.
# determine if repetitions were discarded:
  iteration_count<-iteration_count+1
  if(sum(rowSums(null_matrix>0)!=site_richness)==0){ # If control
# to calculate how many individuals were present at each site per repetition, on average:
  cumulative_site_abundance<-(cumulative_site_abundance+rowSums(null_matrix))
# To show if, somehow, there were sites where alpha(null)!=alpha(obs):
  null_site_richness_diff<-(null_site_richness_diff+(site_richness-
    rowSums(null_matrix>0)))
# Calculate beta of random/null community, keeping both upper and lower triangle:
  null_sim1<-t(as.matrix(vegdist(null_matrix, method=metric, diag=T)))
  null_sim2<-t(as.matrix(vegdist(log(null_matrix+1), method=metric, diag=T)))
# append that beta dissimilarity matrix to the array of all repetitions:
  all_null_sim1[,,h]<-null_sim1
  all_null_sim2[,,h]<-null_sim2
# will break out to next iteration of "h in repetition" if alpha(null)==alpha(obs):
  break
  }
}
# update progress bar
  Sys.sleep(0.1)
  setTxtProgressBar(pb, h)
}
close(pb) # close progress bar
# Calculate the mean beta among pairwise sites for all random repetitions (n=reps):
mean_sim1<-apply(all_null_sim1,MARGIN=c(1,2),FUN=mean)
mean_sim2<-apply(all_null_sim2,MARGIN=c(1,2),FUN=mean)
# Calculate the variance in mean beta among pairwise null sites:
var_sim1<-apply(all_null_sim1,MARGIN=c(1,2),FUN=var)
var_sim2<-apply(all_null_sim2,MARGIN=c(1,2),FUN=var)
# Calculate the OBSERVED pairwise beta in the original matrix:
obs_sim1<-t(as.matrix(vegdist(original_spXsite, method=metric, diag=T)))
obs_sim2<-t(as.matrix(vegdist(log(original_spXsite+1), method=metric, diag=T)))
# Get a standardized score of the beta deviation from null for each pairwise comparison:
standard_score1<-(obs_sim1-mean_sim1)/(sqrt(var_sim1))
standard_score2<-(obs_sim2-mean_sim2)/(sqrt(var_sim2))
# calculate difference in #individual per run of the iteration vs observed:
mean_abundance_deviation_per_rep<-(cumulative_site_abundance -
(site_abundance*reps))/reps
# Show if any iterations had alpha(null)!=alpha(obs):
  null_site_richness_diff<-null_site_richness_diff/reps
  abundance_info<-data.frame(site_richness, null_site_richness_diff, site_abundance,
mean_abundance_deviation_per_rep)
# how many iterations were discarded (because alpha(null)<alpha(obs))?
  unused_iterations<-iteration_count-reps
# return the output:
  return(list(standard_score1, standard_score2, abundance_info, unused_iterations))
}

```

## Appendix B: Exploratory Data Analyses

### Regional Overlap of Families

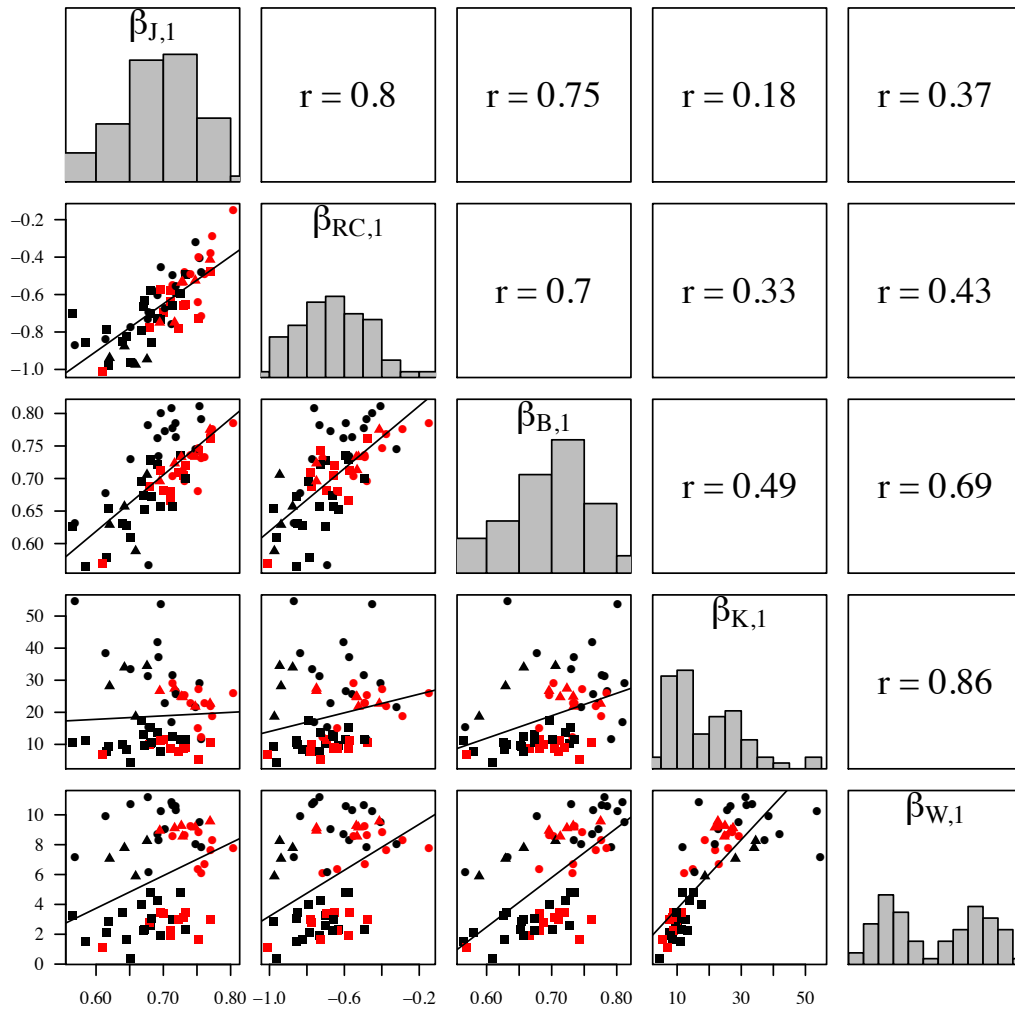
Overlap of families among island group was high, with the majority of families observed in all three island group (29/50), and few families unique to only one island group (Supplemental Figure 1). The families observed only in one island group were Diodontidae and Cheilodactylidae in the Hawaiian Archipelago; Dasyatidae, Belonidae, Ginglymostomatidae, and Nemipteridae in the Mariana Islands Archipelago; and Ehippidae in American Samoa.



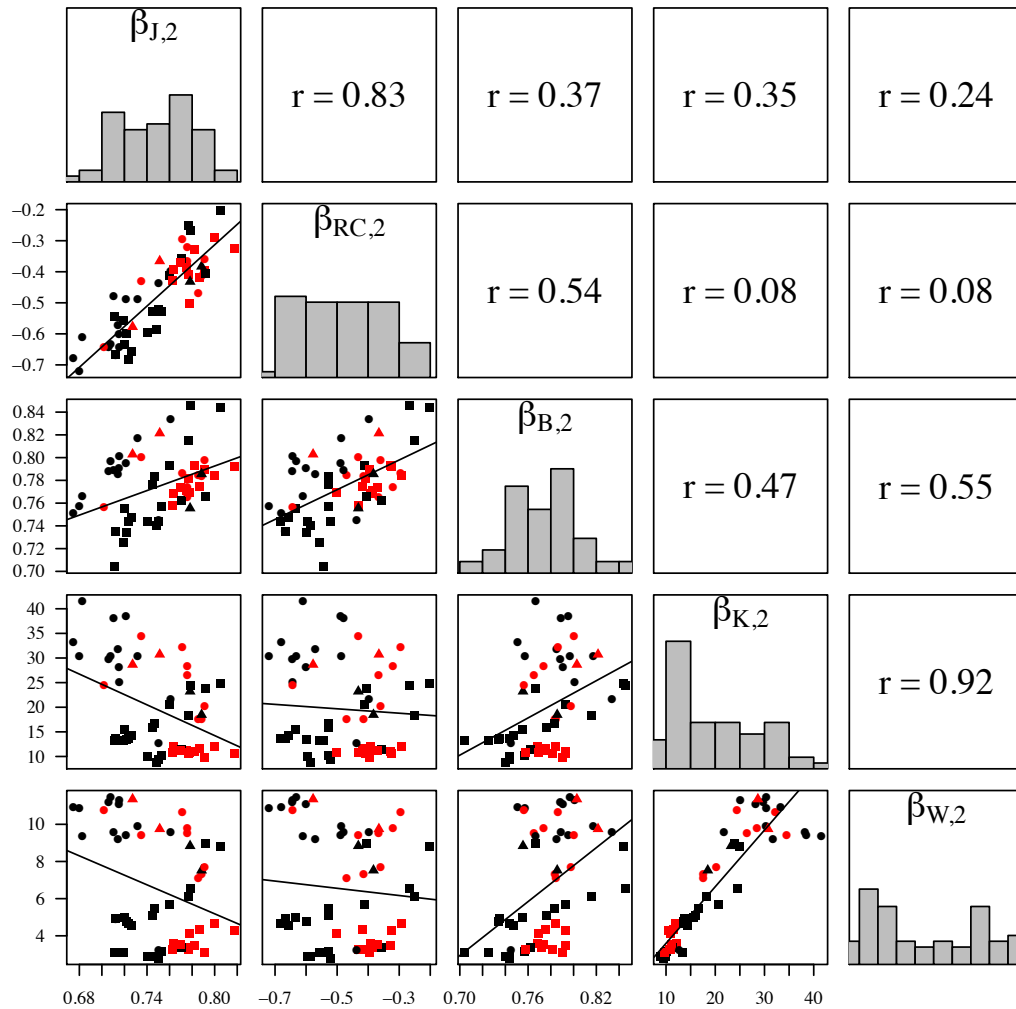
**Supplemental Figure 1.** Overlap of families observed in the Hawaiian Archipelago (H.A.; blue), Mariana Archipelago (M.A.; pink), and American Samoa (A.S.; yellow). Data includes surveys performed during both survey periods (2009 – 2012).

## Correlations Among $\beta$ -diversity Values

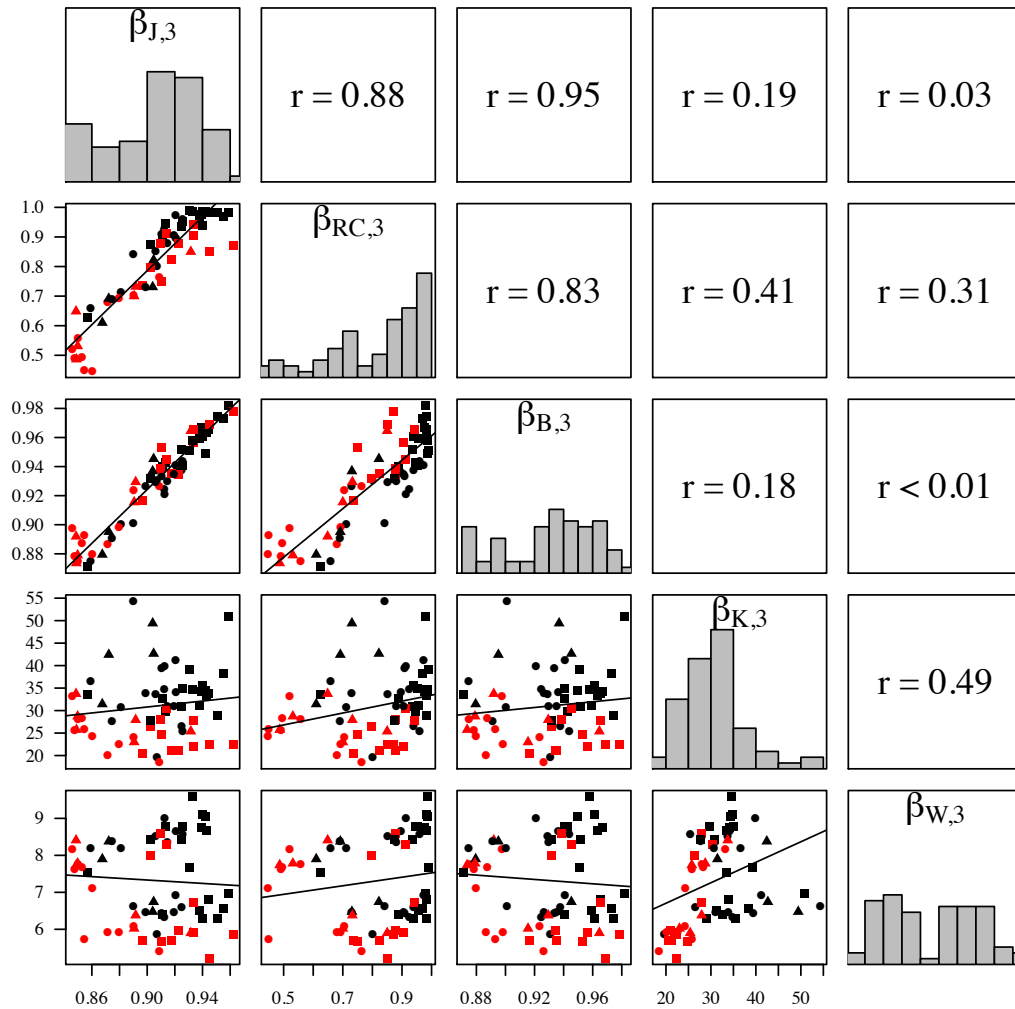
Correlations among island  $\beta$ -diversity values derived from different null models or  $\beta$ -diversity metrics were examined to discern how similar or dissimilar different metrics were at all three scales (Supplemental Figure 2, Supplemental Figure 3, & Supplemental Figure 4 respectively). Two incidence-based and three abundance-based metrics of  $\beta$ -diversity were compared: Jaccard  $\beta$ -diversity ( $\beta_J$ ), Raup-Crick null model  $\beta$ -diversity ( $\beta_{RC}$ ), Bray-Curtis  $\beta$ -diversity ( $\beta_B$ ), abundance-constrained null model  $\beta$ -deviations ( $\beta_K$ ), and  $\alpha$ -constrained null model  $\beta$ -deviations ( $\beta_W$ ).  $\beta_{RC}$  and  $\beta_J$  are incidence based metrics, and  $\beta_W$ ,  $\beta_K$ ,  $\beta_B$  are abundance based metrics. Spearman's rho was calculated for correlations among all island  $\beta$ -diversity values for both survey periods at each spatial scale. Both abundance based null models metrics ( $\beta_W$  and  $\beta_K$ ) were highly correlated at the within island and within island group scale ( $r = 0.86$  and  $0.92$  respectively), but to a lesser degree at the across island group scale ( $r = 0.49$ ). Neither  $\beta_W$  nor  $\beta_K$  correlated strongly ( $r > 0.5$ ) with either incidence based  $\beta$ -diversity metric at any scale.  $\beta_J$  and  $\beta_{RC}$  were highly correlated at all three scales ( $r \geq 0.8$ ).



**Supplemental Figure 2.** Correlation among island  $\beta$ -diversity metrics at the within island scale ( $\beta_1$ ). Diagonal panels show distribution of  $\beta_1$ , with y-axis ranging from 0 – 20. Upper panels show spearman rho values that correspond to bottom panels where islands with low (black) and high (red) human disturbance of the Hawaiian Archipelago (squares), Mariana Archipelago (circles), and American Samoa are plotted. Linear correlations are represented with solid lines. Island  $\beta$ -diversity values from both survey periods from are shown.



**Supplemental Figure 3.** Correlations among island  $\beta$ -diversity metrics at the within island group scale ( $\beta_2$ ). Diagonal panels show distribution of  $\beta_2$ , with y-axis ranging from 0 – 20. Upper panels show Spearman rho values that correspond to bottom panels where islands with low (black) and high (red) human disturbance of the Hawaiian Archipelago (squares), Mariana Archipelago (circles), and American Samoa are plotted. Linear correlations are represented with solid lines. Island  $\beta$ -diversity values from both survey periods from are shown.

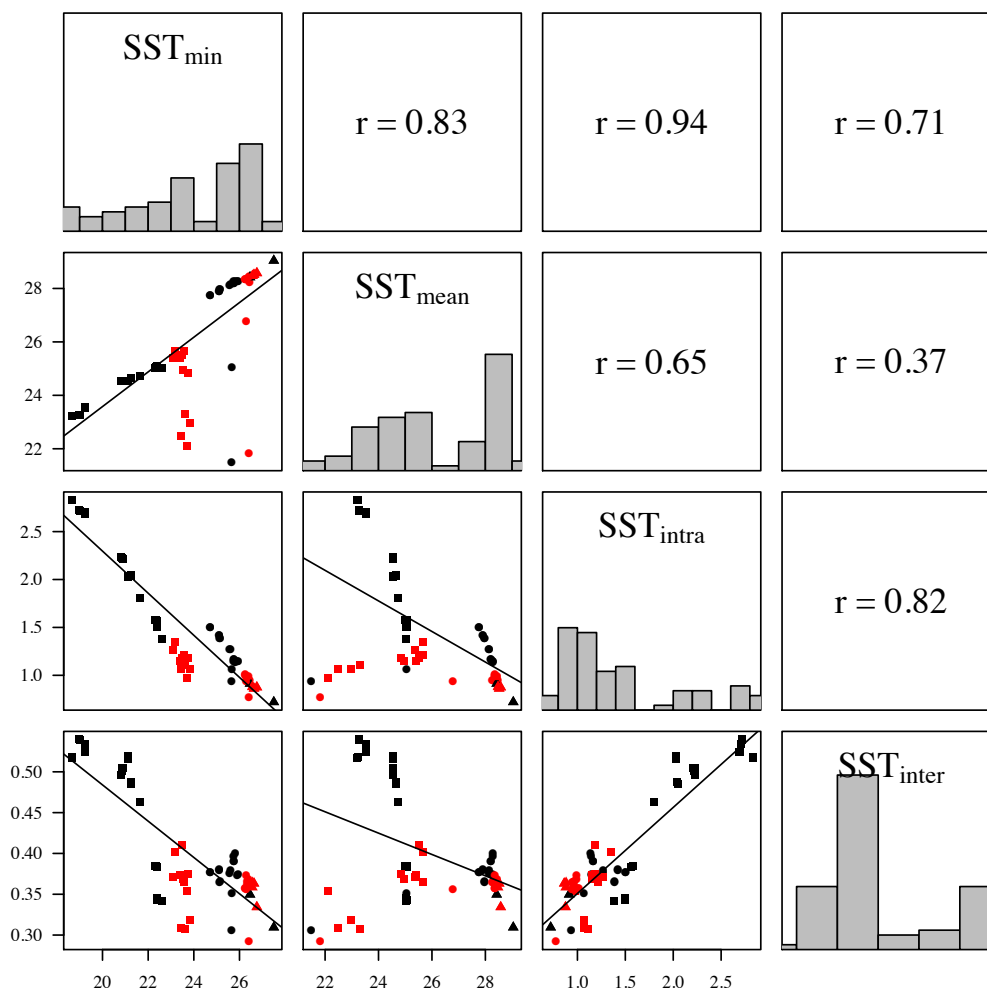


**Supplemental Figure 4.** Correlations among island  $\beta$ -diversity metrics at the across island group scale ( $\beta_3$ ). Diagonal panels show distribution of  $\beta_3$ , with y-axis ranging from 0 – 20. Upper panels show Spearman rho values that correspond to bottom panels where islands with low (black) and high (red) human disturbance of the Hawaiian Archipelago (squares), Mariana Archipelago (circles), and American Samoa are plotted. Linear correlations are represented with solid lines. Island  $\beta$ -diversity values from both survey periods from are shown.

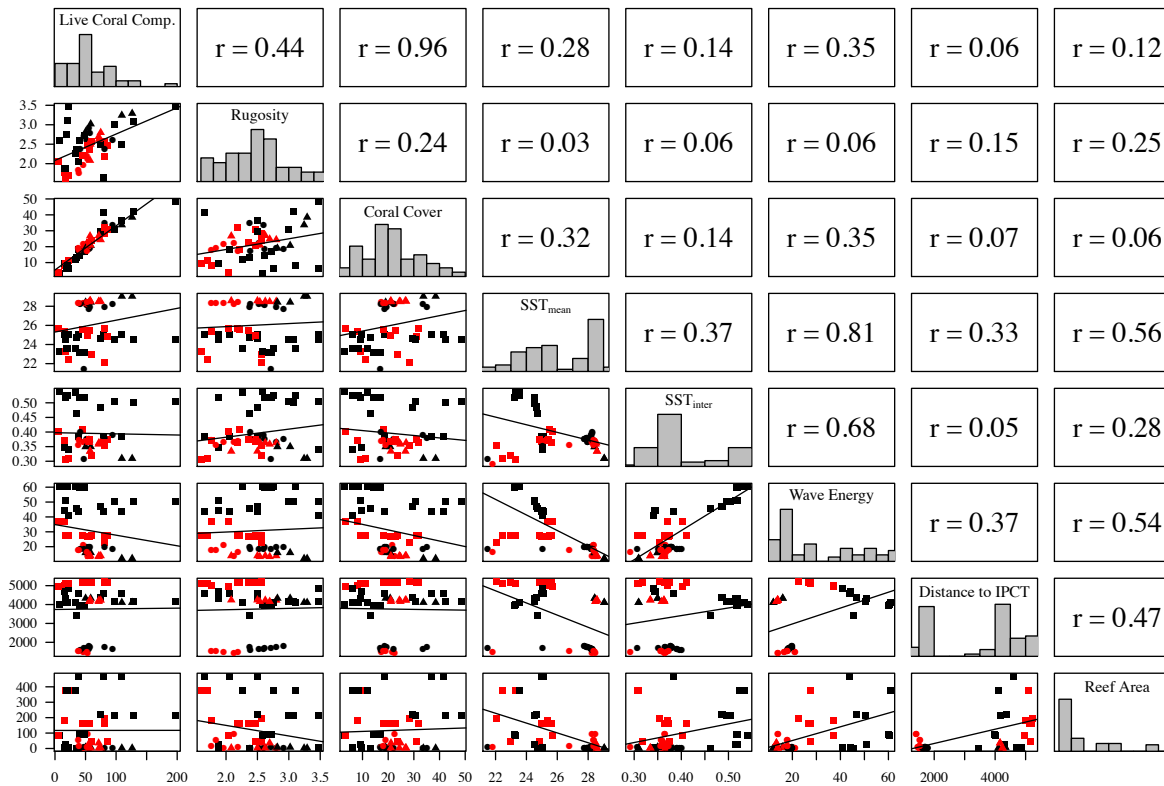


## Correlations Among Predictor Variables

Correlations among predictor variables were visualized and Spearman's rho calculated for island values using data from both survey periods. For data that has site-level resolution (sea surface temperature (SST), live coral complexity, rugosity, live coral cover), island values were calculated as the mean of all sites values per island. As expected, the four SST variables were highly correlated (Supplemental Figure 5), therefore full analysis was performed only on mean SST ( $SST_{\text{mean}}$  or  $SST_{\text{m}}$ ) and interannual variability in SST ( $SST_{\text{inter}}$  or  $SST_{\text{v}}$ ) as they were least correlated and represent different aspects of energy forcing; mean and variability therein. Correlations of other predictor variables were generally much lower (Supplemental Figure 6). Live coral complexity and live coral cover were highly correlated ( $r = 0.96$ ), however both were not used together in any analysis. Wave energy was highly correlated to both  $SST_{\text{mean}}$  and  $SST_{\text{inter}}$  ( $r = 0.81$  and  $0.68$  respectively), driven in part from the north-western islands of the Hawaiian Archipelago that have low human disturbance, low  $SST_{\text{mean}}$  but high  $SST_{\text{var}}$ , and high wave energy. Wave energy and  $SST_{\text{mean}}$  were also correlated to reef area ( $r = 0.54$  and  $0.56$  respectively), as smaller islands (lower reef area) tended to have lower wave energy but higher  $SST_{\text{mean}}$ .



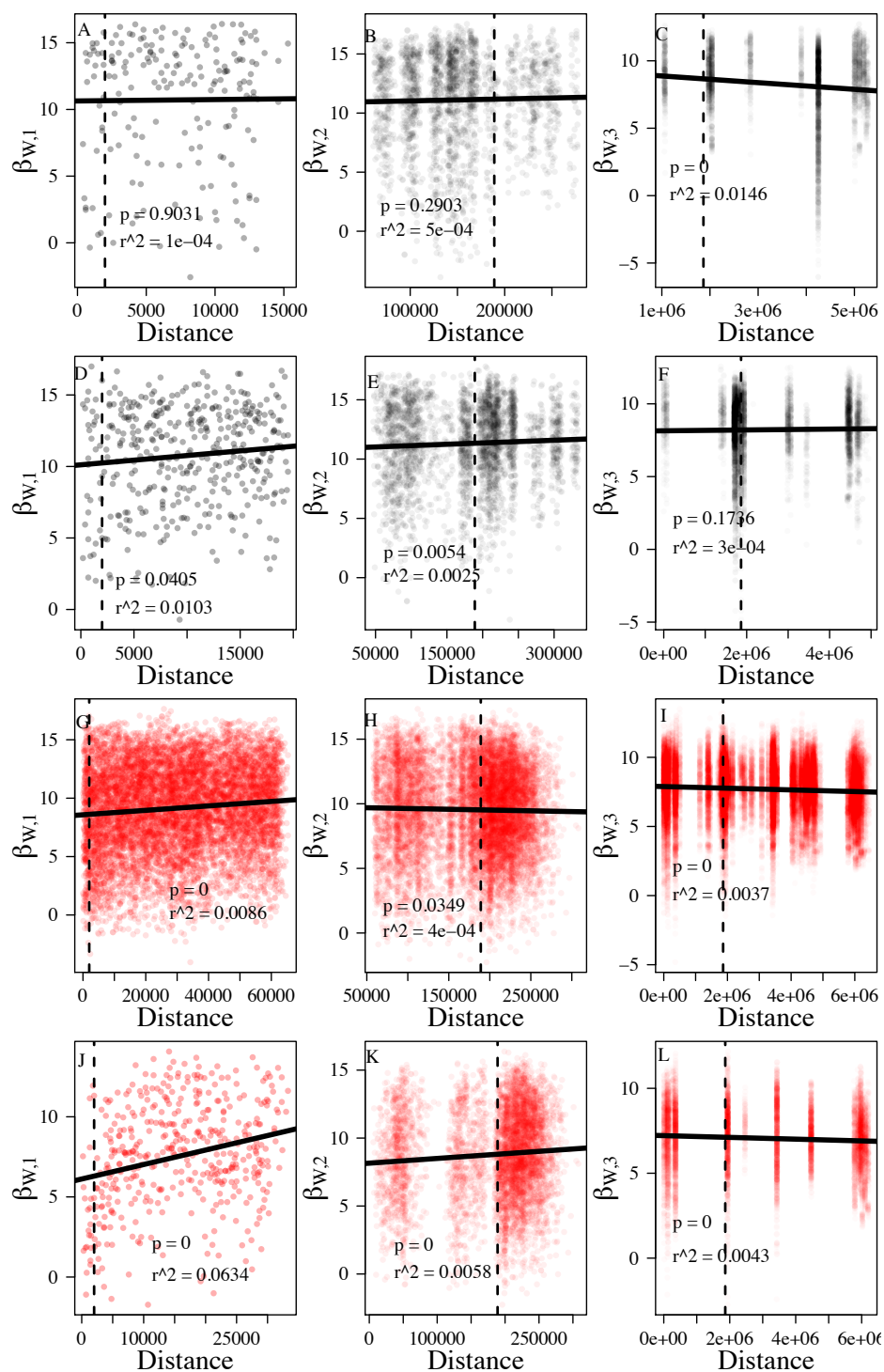
**Supplemental Figure 5.** Correlations among sea surface temperature (SST) values for each island.  $SST_{\min}$ ,  $SST_{\text{mean}}$ ,  $SST_{\text{intra}}$ , and  $SST_{\text{inter}}$  represent the mean minimum, mean, intra-annual variance in mean, and interannual variance in mean SST of each site. Island values were derived by taking the mean of all site values for a given island. Diagonal panel shows distribution for each variable (y-axis ranges from 0 to 30), upper panel shows Spearman's rho, and lower panel shows linear correlations. Remote (black) and populated (red) islands from either survey period of the Hawaiian Archipelago (squares), Mariana Island Archipelago (circles), and American Samoa (triangle) are shown.



**Supplemental Figure 6.** Correlations among island predictor variable. Island values were derived by taking the mean of site values for a given island, except for wave energy, distance to IPCT, and reef area where data resolution is at the island level. Diagonal panel shows distribution for each variable (y-axis ranges from 0 to 30), upper panel shows Spearman's rho, and lower panel shows linear correlations. Remote (black) and populated (red) islands from either survey period of the Hawaiian Archipelago (squares), Mariana Island Archipelago (circles), and American Samoa (triangle) are shown.

## Distance Correction

Island  $\beta$ -diversity values were calculated from regression between geographical distance between sites and pairwise site-site dissimilarity for each metric of  $\beta$ -diversity. Regressions were done independently on each island, for either study period, and the y-intercept and slope for each regression was used to estimate  $\beta$ -diversity at a common distance among all islands (see section 3.2.1). This was done to reduce biases associated with different sized islands (Figure 3) and the known positive correlation of  $\beta$ -diversity with geographical distance among sites (e.g. Condit et al. (2002)). Included are examples of regressions of islands in the Mariana Archipelago, performed at each of the three spatial scales. p-values may be low due to artificially high sample size as calculating pairwise dissimilarity generates more data points than sites observed (Eq. 5). As such, interpreting p-values is cautioned against.



**Supplemental Figure 7.** Island  $\beta$ -diversity values were calculated from regression of pairwise dissimilarity among sites with geographical distance among sites (in arbitrary units). This was calculated at each of the three spatial scales, within islands (A, D, G, J), within island groups (B, E, G, K), and across island groups (C, F, I, L) and is shown for Agrihan (A – C), Pagan (D – F), Guam (G – I), and Saipan (J – L). Vertical dashed line denotes distance at which island  $\beta$ -diversity was estimated.  $r^2$  and p-value of linear regression shown on plots.

## Appendix C: Additional $\beta$ -diversity Model Information

Here, I present the precise value of each model weighted parameter estimate, together with the unconditional variance, and the number of models each parameter was included in, for each model selection process of  $\beta_{RC}$  and  $\beta_W$  at all three scales.

**Supplemental Table 1.** Model weighted parameter estimation for incidence-based Raup-Crick  $\beta$ -diversity,  $\beta_{RC}$ , weighted using all models within 2  $\Delta AIC$  of best model. Intercept represents an island in the Hawaiian Archipelago. Unconditional variance (Variance) as per Buckland et al. (1997); # Models = the number of models within 2  $\Delta AIC$  that each parameter was included in.

	Estimate	Variance	# Models
<i>Within Island <math>\beta</math>-diversity Models</i>			
Intercept	-0.791970235	0.002397176	8
Humans	0.140263849	0.002524915	8
Mariana Islands Archipelago	0.122304327	0.005813051	8
American Samoa	0.012700813	0.008428307	8
Live Coral Complexity	-0.045827706	0.000456857	8
SSTv	-0.023825815	0.001127596	4
Reef Area	-0.005510298	0.000156456	3
Wave Energy	0.021432909	0.002140345	3
SSTm	0.000432239	1.25E-05	1
<i>Within Island Group <math>\beta</math>-diversity Models</i>			
Intercept	-0.568047566	0.001623593	5
Humans	0.203357039	0.002483034	5
Wave Energy	0.17293568	0.003479215	5
SSTv	-0.15553665	0.001532516	5
SSTm	0.050721216	0.001265021	4
Live Coral Complexity	0.000466718	1.12E-05	1
Reef Area	0.000738599	1.79E-05	1
Mariana Islands Archipelago	-0.008753273	0.000659541	1
<i>Across Island Group <math>\beta</math>-diversity Models</i>			
Intercept	1.047667617	0.002174136	5
Humans	-0.225370344	0.002084652	5
Mariana Islands Archipelago	-0.332632157	0.008281251	5
American Samoa	-0.395257254	0.010822807	5
Wave Energy	-0.181787875	0.004977427	5
SSTv	0.104656765	0.001365439	5
Reef Area	0.018906766	0.000499323	3
SSTm	-0.015922925	0.000638164	2
Live Coral Complexity	0.000166712	4.15E-06	1

**Supplemental Table 2.** Model weighted parameter estimation for  $\alpha$ -constraining abundance-based null model  $\beta$ -diversity,  $\beta_w$ , weighted using all models within  $2 \Delta AIC$  of best model. Intercept represents an island in the Hawaiian Archipelago. Unconditional variance (Variance) as per Buckland et al. (1997); # Models = the number of models within  $2 \Delta AIC$  that each parameter was included in.

	Estimate	Variance	# Models
<i>Within Island <math>\beta</math>-diversity Models</i>			
Intercept	3.546576771	0.152615929	10
Mariana Islands Archipelago	5.953872575	0.361500098	10
American Samoa	4.688449923	0.69818864	10
Live Coral Complexity	-0.085331176	0.027934946	6
Live Coral Complexity X Humans	1.034870123	1.34416635	6
SSTm	-0.039310692	0.014821811	2
Reef Area	-0.04343885	0.011724167	2
Humans	-0.038407445	0.00843976	2
SSTv	-0.026105334	0.00478737	1
Wave Energy	-0.020968506	0.002576883	1
<i>Within Island Group <math>\beta</math>-diversity Models</i>			
Intercept	5.627414139	0.029591475	5
Humans	-1.626452824	0.038137566	5
Mariana Islands Archipelago	5.263999054	0.056644878	5
SSTv	-0.480477079	0.011887652	5
Reef Area	-0.00192732	0.000222942	1
Wave Energy	0.021289757	0.004399299	1
SSTm	0.012262904	0.001128807	1
Live Coral Complexity	-0.010903485	0.000678156	1
<i>Across Island Group <math>\beta</math>-diversity Models</i>			
Intercept	8.519340628	0.021741624	5
Humans	-0.567065171	0.035569064	5
SSTv	0.371351277	0.010786626	5
Reef Area	-0.0036105	0.000229829	1
SSTm	0.011188452	0.00074303	1
Wave Energy	-0.019492025	0.002140246	1
Live Coral Complexity	0.011493486	0.00072239	1

## **Appendix D: Species Lists**

A large number of species observed in each island group were recorded only on islands with low or high human disturbance (Figure 5 B–D). Of those species, a greater proportion on low human disturbance islands tended to be piscivorous, whereas a greater proportion on high disturbance islands tended to be low carnivores (Table 6). Tables to follow show those species, and represent the species used to produce Figure 5 B–D and Table 6.



**Supplemental Table 3.** Species observed on islands with low-, but not high- human disturbance of the Hawaiian Archipelago, their  $L_{\max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed only during second survey period (2011 – 2012).

Family	Common name	Scientific name	$L_{\max}$	# Sites	# Observed
<b>Piscivores</b>					
Carangidae	Giant trevally	<i>Caranx ignobilis</i>	165	37	178
	Greater amberjack	<i>Seriola dumerili</i>	190	4	6
	Rainbow runner	<i>Elagatis bipinnulata</i>	180	2	2
	Black jack	<i>Caranx lugubris</i>	100	1	1
Carcharhinidae	Galapagos shark	<i>Carcharhinus galapagensis</i>	370	27	58
	Grey reef shark	<i>Carcharhinus amblyrhynchos</i>	240	4	6
Scombridae	Kawakawa	<i>Euthynnus affinis</i>	100	1	1
Scorpaenidae	Hawaiian turkeyfish	<i>Pterois sphex</i>	22	6	10
	False stonefish	<i>Scorpaenopsis diabolus</i>	30	1	1
	Titan scorpionfish	<i>Scorpaenopsis cacopsis</i>	51	1	1
Synodontidae	Two-spot lizard fish	<i>Synodus binotatus</i>	17	2	2
	Sand lizardfish	<i>Synodus dermatogenys</i>	24	2	2
<b>Low Carnivores</b>					
Chaetodontidae	Chevron butterflyfish	<i>Chaetodon trifascialis</i>	18	3	14
Cheilodactylidae	Hawaiian morwong	<i>Goniistius vittatus</i>	41	14	49
Diodontidae	Spot-fin porcupinefish	<i>Diodon hystrix</i>	91	1	1
Holocentridae	Speckled squirrelfish	<i>Sargocentron punctatissimum</i>	13	1	8
	Yellowstriped squirrelfish	<i>Neoniphon aurolineatus</i>	23	1	1
Labridae	Yellowstripe coris	<i>Coris flavovittata</i>	65	37	72
	Slingjaw wrasse	<i>Epibulus insidiator</i>	54	15	22
Mullidae	Yellowbarbel goatfish	<i>Parupeneus chrysonemus</i>	22	1	1
Muraenidae	Abbott's moray eel	<i>Gymnothorax eurostus</i>	60	2	2
Oplegnathidae	Barred knifejaw	<i>Oplegnathus fasciatus</i>	80	1	1
Scorpaenidae	Humpback nohu	<i>Sebastapistes coniorta</i>	10	1	5
	Hawaiian lionfish	<i>Dendrochirus barberi</i>	16.5	1	2
<b>Planktivores</b>					
Balistidae	Redtail triggerfish	<i>Xanthichthys mento</i>	30	2	22
Clupeidae	Delicate round herring	<i>Spratelloides delicatulus</i>	7	1	3350
Holocentridae	Yellowfin soldierfish	<i>Myripristis chryseres</i>	25	1	1
Myliobatidae	Giant manta	<i>Manta birostris</i>	800	2	2
Pomacanthidae	Masked angelfish	<i>Genicanthus personatus</i>	25	9	35
Pomacentridae	Midget chromis	<i>Chromis acares</i>	6	1	10
Priacanthidae	Hawaiian bigeye	<i>Priacanthus meeki</i>	33	2	2
Priacanthidae	Glasseye	<i>Heteropriacanthus cruentatus</i>	32	1	1
Serranidae	Hawaiian anthias	<i>Pseudanthias thompsoni</i>	22	1	2
<b>Herbivores</b>					
Acanthuridae	Lined surgeonfish	<i>Acanthurus lineatus</i>	38	3	8
Kyphosidae	Pacific chub	<i>Kyphosus sandwicensis</i>	75	3	60
	Hawaiian chub	<i>Kyphosus hawaiiensis</i>	41	3	4
	Blue seachub	<i>Kyphosus cinerascens</i>	51	2	6
Pomacanthidae	Japanese angelfish	<i>Centropyge interruptus</i>	16	2	4
Scaridae	Yellowbar parrot	<i>Calotomus zonarchus</i>	30	9	15
Tetraodontidae	Valentinni's sharpnosepuffer	<i>Canthigaster valentini</i>	9	2	13
	Brown-lined puffer	<i>Canthigaster rivulata</i>	20	1	1

**Supplemental Table 4.** Species observed on islands with high-, but not low- human disturbance of the Hawaiian Archipelago, their  $L_{\max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed only during second survey period (2011 – 2012).

Family	Common name	Scientific name	$L_{\max}$	# Sites	# Observed
<b>Piscivores</b>					
Muraenidae	Yellow-edged moray	<i>Gymnothorax flavimarginatus</i>	240	7	7
	Undulated moray	<i>Gymnothorax undulatus</i>	150	3	3
	Giant moray	<i>Gymnothorax javanicus</i>	300	2	2
Synodontidae	Variegated lizardfish	<i>Synodus variegatus</i>	24	1	1
<b>Low Carnivores</b>					
Chaetodontidae	Saddle butterflyfish	<i>Chaetodon ephippium</i>	23	2	5
	Lined butterflyfish	<i>Chaetodon lineolatus</i>	30	1	1
Holocentridae	Yellow-striped squirrelfish	<i>Sargocentron ensifer</i>	25	1	1
Labridae	Cigar wrasse	<i>Cheilio inermis</i>	51	7	8
	Sharp-headed wrasse	<i>Cymolutes lecluse</i>	18	1	2
Lutjanidae	Blacktail snapper	<i>Lutjanus fulvus</i>	43	18	50
Muraenidae	Snowflake moray	<i>Echidna nebulosa</i>	75	1	1
	Dwarf moray	<i>Gymnothorax melatremus</i>	26	1	1
Myliobatidae	Spotted eagle ray	<i>Aetobatus narinari</i>	350	1	1
Ostraciidae	Thornback cowfish	<i>Lactoria fornasini</i>	14	1	1
<b>Planktivores</b>					
Apogonidae	Spotted cardinalfish	<i>Apogon maculiferus</i>	14	1	12
Chaetodontidae	Pyramid butterflyfish	<i>Hemitaurichthys polylepis</i>	18	2	25
	Thompson's butterflyfish	<i>Hemitaurichthys thompsoni</i>	21	2	16
Holocentridae	Whitetip soldierfish	<i>Myripristis vittata</i>	20	1	1
Pomacentridae	Whitetail chromis	<i>Chromis leucura</i>	6.5	2	3
Serranidae	Bicolor anthias	<i>Pseudanthias bicolor</i>	13	8	32
<b>Herbivores</b>					
Acanthuridae	Whitespotted surgeonfish	<i>Acanthurus guttatus</i>	29	3	3
Pomacanthidae	Flame angel	<i>Centropyge loricula</i>	10	2	2
Pomacentridae	Rock damselfish	<i>Plectroglyphidodon sindonis</i>	12.5	3	7

**Supplemental Table 5.** Species observed on islands with low-, but not high- human disturbance of the Mariana Archipelago, their  $L_{\max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed during second survey period (2011 – 2012).

Family	Common name	Scientific name	$L_{\max}$	# Sites	# Observed
<b>Piscivores</b>					
Belontiidae	Houndneedlefish	Tylosurus crocodilus	135	1	1
Carangidae	Rainbow runner	Elagatis bipinnulata	180	3	21
Carangidae	Bigeye trevally	Caranx sexfasciatus	100	1	12
Carcharhinidae	Whitetip reef shark	Trienodon obesus	213	9	10
Ginglymostomatid	Tawny nurse shark	Nebrius ferrugineus	320	2	3
Lutjanidae	Black-banded snapper	Lutjanus semicinctus	35	1	1
Serranidae	Sixblotch hind	Cephalopholis sexmaculata	50	7	10
Serranidae	Highfin grouper	Epinephelus maculatus	60	5	7
Serranidae	Spotted soapfish	Pogonoperca punctata	35	2	2
Serranidae	Redmouth grouper	Aethaloperca rogae	60	1	1
Serranidae	Snubnose grouper	Epinephelus macrospilos	51	1	1
Sphyraenidae	Heller's barracuda	Sphyraena helleri	80	1	25
<b>Low Carnivores</b>					
Chaetodontidae	Yellow-crowned butterflyfish	Chaetodon flavocoronatus	12	1	2
Dasyatidae	Blotched fantail ray	Taeniura meyeni	330	1	1
Labridae	Tarry hogfish	Bodianus bilunulatus	55	3	3
Labridae	Cheekspot wrasse	Halichoeres melasmapomus	14	2	4
Labridae	Blackfin hogfish	Bodianus loxozonus	40	1	1
Oplegnathidae	Spotted knifejaw	Oplegnathus punctatus	86	1	1
<b>Planktivores</b>					
Acanthuridae	Gray unicornfish	Naso caesius	62	1	1
Acanthuridae	Elongate surgeonfish	Acanthurus mata	50	1	1
Apogonidae	Short-tooth cardinal	Apogon apogonoides	10	1	15
Apogonidae	Iridescent cardinalfish	Apogon kallopterus	12	1	11
Balistidae	Bluelined triggerfish	Xanthichthys caeruleolineatus	35	2	8
Chaetodontidae	Thompson's butterflyfish	Hemitaurichthys thompsoni	21	4	17
Chaetodontidae	False moorish idol	Heniochus diphreutes	20	1	1
Clupeidae	Delicate round herring	Spratelloides delicatulus	7	12	2782
Holocentridae	Whitetip soldierfish	Myripristis vittata	20	2	30
Labridae	Lyretail hogfish	Bodianus anthioides	21	8	9
Microdesmidae	Helfrich's dartfish	Nemateleotris helfrichi	6.3	1	2
Pomacanthidae	Blackedged angelfish	Genicanthus watanabei	15	1	9
Pomacentridae	Fusilier damselfish	Lepidozygus tapeinosoma	10	8	1206
<b>Herbivores</b>					
Acanthuridae	Whitebar surgeonfish	Acanthurus leucopareius	25	48	452
Acanthuridae	Eyestripe surgeonfish	Acanthurus dussumieri	56	3	10
Ostraciidae	Yellow boxfish	Ostracion cubicus	45	1	1
Pomacanthidae	Orange angelfish	Centropyge fisheri	7.5	1	1
Tetraodontidae	Valentini's sharpnosepuffer	Canthigaster valentini	9	1	1

**Supplemental Table 6.** Species observed on islands with high-, but not low- human disturbance of the Mariana Archipelago, their  $L_{\max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed during second survey period (2011 – 2012).

Family	Common name	Scientific name	$L_{\max}$	# Sites	# Observed
<b>Piscivores</b>					
Apogonidae	Large toothed cardinalfish	<i>Cheilodipterus macrodon</i>	20	1	1
Aulostomidae	Chinese trumpetfish	<i>Aulostomus chinensis</i>	80	7	9
Carangidae	Blue trevally	<i>Carangoides ferdau</i>	70	3	12
	Brassy trevally	<i>Caranx papuensis</i>	88	1	4
Lethrinidae	Yellowlip emperor	<i>Lethrinus xanthurus</i>	62	2	2
Muraenidae	Giant moray	<i>Gymnothorax javanicus</i>	300	1	1
Scombridae	Double-lined mackerel	<i>Grammatorcynus bilineatus</i>	100	1	1
Scorpaenidae	False stonefish	<i>Scorpaenopsis diabolus</i>	30	1	1
Serranidae	Honeycomb grouper	<i>Epinephelus merra</i>	33	6	9
	Leopard hind	<i>Cephalopholis leopardus</i>	20	4	7
	Blacksaddled coral grouper	<i>Plectropomus laevis</i>	125	1	2
Synodontidae	Variiegated lizardfish	<i>Synodus variegatus</i>	24	2	3
	Two-spot lizard fish	<i>Synodus binotatus</i>	17	1	1
<b>Low Carnivores</b>					
Apogonidae	Five-lined cardinalfish	<i>Cheilodipterus quinquelineatus</i>	12	4	15
	Wolf cardinalfish	<i>Cheilodipterus artus</i>	18.7	2	6
Blenniidae	Bicolour fangblenny	<i>Plagiotremus laudandus</i>	10	15	61
Chaetodontidae	Chevron butterflyfish	<i>Chaetodon trifascialis</i>	18	6	7
	Singular bannerfish	<i>Heniochus singularis</i>	30	3	5
	Lined butterflyfish	<i>Chaetodon lineolatus</i>	30	2	4
	Sunset butterflyfish	<i>Chaetodon pelewensis</i>	13	1	3
Cirrhitidae	Twospot hawkfish	<i>Amblycirrhitus bimaculatus</i>	9	2	5
Haemulidae	Painted sweetlip	<i>Plectorhynchus picus</i>	85	4	6
Holocentridae	Blackfin squirrelfish	<i>Neoniphon opercularis</i>	35	3	10
	Smallmouth squirrelfish	<i>Sargocentron microstoma</i>	20	3	4
Labridae	Lattice soldierfish	<i>Myripristis violacea</i>	30	2	19
	Tripletail wrasse	<i>Cheilinus trilobatus</i>	45	67	105
	Slingjaw wrasse	<i>Epibulus insidiator</i>	54	58	94
	Blackeye thicklip	<i>Hemigymnus melapterus</i>	60	21	29
	Smalltail wrasse	<i>Pseudojuloides cerasinus</i>	12	16	51
	Redbreast wrasse	<i>Cheilinus fasciatus</i>	36	16	22
	Three-line/Three-	<i>Stethojulis strigiventer</i>	15	10	12
	Humphead wrasse	<i>Cheilinus undulatus</i>	229	5	8
	Cigar wrasse	<i>Cheilio inermis</i>	51	3	3
	Threespot wrasse	<i>Halichoeres trimaculatus</i>	27	2	5
	Palebarred coris	<i>Coris dorsomaculata</i>	20	2	3
	Two-spot wrasse	<i>Oxycheilinus bimaculatus</i>	15	2	2
	Cockerel wrasse	<i>Pteragogus enneacanthus</i>	15	1	1
Lethrinidae	Thumbprint emperor	<i>Lethrinus harak</i>	32	7	17
	Spotcheek emperor	<i>Lethrinus rubrioperculatus</i>	39.5	4	19
Malacanthidae	Bluehead tilefish	<i>Hoplostethus starcki</i>	15	1	4
Monacanthidae	Blacksaddle filefish	<i>Paraluteres prionurus</i>	11	1	2
	Harlequin filefish	<i>Oxymonacanthus longirostris</i>	9	1	1
Mullidae	Dash-and-dot goatfish	<i>Parupeneus barberinus</i>	50	9	10
Myliobatidae	Spotted eagle ray	<i>Aetobatus narinari</i>	350	1	1
Nemipteridae	Striped monocle bream	<i>Scolopsis lineata</i>	20	9	34
Pinguipedidae	Black dotted sand perch	<i>Parapercis millepunctata</i>	19	4	5
Pomacentridae	Staghorn damselfish	<i>Amblyglyphidodon curacao</i>	10.5	1	3
Scorpaenidae	Broadbarred firefish	<i>Pterois antennata</i>	20	3	3
Serranidae	Sixline soapfish	<i>Grammistes sexlineatus</i>	27	3	3
	Arrowhead soapfish	<i>Belonoperca chabanaudi</i>	15	1	1
Tetraodontidae	White-spotted puffer	<i>Arothron hispidus</i>	52	1	1

Planktivores						
Apogonidae	Bridled cardinalfish	<i>Apogon fraenatus</i>	11	1	1	
Blenniidae	Forktail blenny	<i>Meiacanthus atrodorsalis</i>	10	48	236	
	Red-spotted blenny	<i>Blenniella chrysospilos</i>	14	3	26	
Caesionidae	Lunar fusilier	<i>Caesio lunaris</i>	30	1	20	
Carangidae	Bigeye scad	<i>Selar crumenophthalmus</i>	30	1	15	
Chaetodontidae	Sunburst butterflyfish	<i>Chaetodon kleinii</i>	14	2	4	
Gobiidae	Red-banded/Barred	<i>Amblyeleotris fasciata</i>	8	1	3	
Holocentridae	Pinecone soldierfish	<i>Myripristis murdjan</i>	60	1	2	
Labridae	Sixbar wrasse	<i>Thalassoma hardwicke</i>	20	17	75	
	Moon wrasse	<i>Thalassoma lunare</i>	25	1	1	
Microdesmidae	Blue gudgeon	<i>Ptereleotris microlepis</i>	13	2	23	
Pomacentridae	Yellow-speckled chromis	<i>Chromis alpha</i>	8.5	10	148	
	Fire clownfish	<i>Amphiprion melanopus</i>	10.8	2	15	
	Blue green damselfish	<i>Chromis viridis</i>	8	1	78	
Herbivores						
Acanthuridae	Sailfin tang	<i>Zebrasoma veliferum</i>	40	18	29	
	Twotone tang	<i>Zebrasoma scopas</i>	20	3	3	
	Achilles tang	<i>Acanthurus achilles</i>	25	1	2	
	Humpback unicornfish	<i>Naso brachycentron</i>	90	1	1	
Blenniidae	Barred blenny	<i>Cirripectes polyzona</i>	10	1	2	
Chaetodontidae	Dotted butterflyfish	<i>Chaetodon semeion</i>	24	1	1	
Gobiidae	Banded goby	<i>Amblygobius phalaena</i>	14	1	6	
Pomacanthidae	Pearlscale angelfish	<i>Centropyge vrolikii</i>	12	4	6	
	Flame angel	<i>Centropyge loricula</i>	10	1	4	
	Whitebar gregory	<i>Stegastes albifasciatus</i>	11	7	39	
	Twinspot damselfish	<i>Chrysiptera biocellata</i>	7	2	3	
	Singlebar devil	<i>Plectroglyphidodon leucozonus</i>	12	1	21	
	Grey demoiselle	<i>Chrysiptera glauca</i>	8	1	2	
	Dusky farmerfish	<i>Stegastes nigricans</i>	14	1	1	
	Scaridae	Yellowband parrotfish	<i>Scarus schlegeli</i>	38	53	161
		Globehead parrotfish	<i>Scarus globiceps</i>	31	16	22
		Filament-finned parrotfish	<i>Scarus altipinnis</i>	60	13	61
Yellowbarred parrotfish		<i>Scarus dimidiatus</i>	50	8	12	
Darktail parrotfish		<i>Scarus fuscocaudalis</i>	25	7	16	
Festive parrotfish		<i>Scarus festivus</i>	45	5	10	
Pacific longnoseparrotfish		<i>Hipposcarus longiceps</i>	60	5	9	
Blue-barred parrotfish		<i>Scarus ghobban</i>	90	1	1	
Siganidae	Streamlined spinefoot	<i>Siganus argenteus</i>	42	10	31	
	Little spinefoot	<i>Siganus spinus</i>	28	4	4	
	Goldspotted spinefoot	<i>Siganus punctatus</i>	36.5	1	3	
Tetraodontidae	Crowned puffer	<i>Canthigaster coronata</i>	14	1	2	
	Lantern toby	<i>Canthigaster epilampra</i>	11	1	1	

**Supplemental Table 7.** Species observed on islands with low-, but not high- human disturbance of American Samoa, their  $L_{\max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed only during second survey period (2011 – 2012).

Family	Common name	Scientific name	$L_{\max}$	# Sites	# Observed
<b>Piscivores</b>					
Carangidae	Black jack	<i>Caranx lugubris</i>	100	6	42
	Bigeye trevally	<i>Caranx sexfasciatus</i>	100	1	325
Carcharhinidae	Grey reef shark	<i>Carcharhinus amblyrhynchos</i>	240	3	4
	Blacktip reef shark	<i>Carcharhinus melanopterus</i>	200	1	1
Lethrinidae	Yellowlip emperor	<i>Lethrinus xanthochilus</i>	62	2	2
	Longface emperor	<i>Lethrinus olivaceus</i>	100	1	7
Muraenidae	Giant moray	<i>Gymnothorax javanicus</i>	300	2	2
Serranidae	Snubnose grouper	<i>Epinephelus macrospilos</i>	51	4	4
Sphyrnidae	Blackfin barracuda	<i>Sphyrna qenie</i>	170	1	28
<b>Low Carnivores</b>					
Apogonidae	Wolf cardinalfish	<i>Cheilodipterus artus</i>	18.7	1	120
	Five-lined cardinalfish	<i>Cheilodipterus quinquelineatus</i>	12	1	1
Cirrhitidae	Flame hawkfish	<i>Neocirrhites armatus</i>	9	6	16
	Yellow hawkfish	<i>Paracirrhites xanthus</i>	11	1	1
Labridae	Cheekspot wrasse	<i>Halichoeres melasmapomus</i>	14	10	21
<b>Planktivores</b>					
Acanthuridae	Elongate surgeonfish	<i>Acanthurus mata</i>	50	3	3
	Whitemargin unicornfish	<i>Naso annulatus</i>	100	1	3
Balistidae	Gilded triggerfish	<i>Xanthichthys auromarginatus</i>	22	2	4
Carangidae	Mackerel scad	<i>Decapterus macarellus</i>	35	1	24
Holocentridae	Whitetip soldierfish	<i>Myripristis vittata</i>	20	1	4
Pomacentridae	Fusilier damselfish	<i>Lepidozygus tapeinosoma</i>	10	6	1530
	Weber's chromis	<i>Chromis weberi</i>	12	1	6
Serranidae	Olive anthias	<i>Pseudanthias olivaceus</i>	12	4	215
	Whitley's splitfin	<i>Luzonichthys whitleyi</i>	6	1	50
<b>Herbivores</b>					
Acanthuridae	Longnose surgeonfish	<i>Zebrasoma rostratum</i>	21	13	22
Scaridae	Red parrotfish	<i>Scarus xanthopleura</i>	55	7	9

**Supplemental Table 8.** Species observed on islands with high-, but not low- human disturbance of American Samoa, their  $L_{\max}$  (maximum length in cm), the number of sites observed at (# Sites) and total number observed (# Observed) across those sites. Data represent fish observed only during second survey period (2011 – 2012).

Family	Common name	Scientific name	$L_{\max}$	# Sites	# Observed
<b>Piscivores</b>					
Aulostomidae	Chinese trumpetfish	<i>Aulostomus chinensis</i>	80	3	4
Carangidae	Giant trevally	<i>Caranx ignobilis</i>	165	1	1
Serranidae	Yellow-edged lyretail	<i>Variola louti</i>	83	23	27
	Blacksaddled coralgrouper	<i>Plectropomus laevis</i>	125	8	10
	White-edged lyretail	<i>Variola albimarginata</i>	65	4	4
	Squaretail coralgrouper	<i>Plectropomus areolatus</i>	73	3	3
	Slender grouper	<i>Anyperodon leucogrammicus</i>	60	1	2
	Greasy grouper	<i>Epinephelus tauvina</i>	75	1	1
	Honeycomb grouper	<i>Epinephelus merra</i>	33	1	1
	Camouflage grouper	<i>Epinephelus polyphemadion</i>	75	1	1
<b>Low Carnivores</b>					
Balistidae	Clown triggerfish	<i>Balistooides conspicillum</i>	50	2	2
Bleniidae	Bicolour fangblenny	<i>Plagiotremus laudandus</i>	10	3	12
Bothidae	Flowery flounder	<i>Bothus mancus</i>	45	1	2
Chaetodontidae	Speckled butterflyfish	<i>Chaetodon citrinellus</i>	13	33	100
	Oval butterflyfish	<i>Chaetodon lunulatus</i>	15	26	48
	Chevron butterflyfish	<i>Chaetodon trifascialis</i>	18	24	39
	Vagabond butterflyfish	<i>Chaetodon vagabundus</i>	23	19	34
	Horned bannerfish	<i>Heniochus varius</i>	19	15	25
	Singular bannerfish	<i>Heniochus singularis</i>	30	1	1
Cirrhitidae	Dwarf hawkfish	<i>Cirrhitichthys falco</i>	9	10	15
Ephippidae	Orbicular batfish	<i>Platax orbicularis</i>	50	1	5
Gobiidae	Blueband goby	<i>Valenciennea strigata</i>	16	16	49
Haemulidae	Indian Ocean	<i>Plectorhinchus vittatus</i>	85	12	16
	Harry hotlips	<i>Plectorhinchus gibbosus</i>	75	1	1
Holocentridae	Sammara squirrelfish	<i>Neoniphon sammara</i>	32	8	12
	Crown squirrelfish	<i>Sargocentron diadema</i>	17	2	2
	Lattice soldierfish	<i>Myripristis violacea</i>	30	1	1
Labridae	Tripletail wrasse	<i>Cheilinus trilobatus</i>	45	9	9
	Blackeye thicklip	<i>Hemigymnus melapterus</i>	60	8	10
	Red-lined wrasse	<i>Halichoeres biocellatus</i>	12	7	16
	Three-line/Three-	<i>Stethojulis strigiventer</i>	15	7	11
	Pastel ringwrasse	<i>Hologymnosus doliatus</i>	40	5	7
	Tubelip wrasse	<i>Labrichthys unilineatus</i>	16	4	5
	Twotone wrasse	<i>Halichoeres prosopeion</i>	16	2	5
	Rockmover wrasse	<i>Novaculichthys taeniourus</i>	30	2	2
	Spotted wrasse	<i>Anampses meleagrides</i>	22	1	1
	Diana's hogfish	<i>Bodianus diana</i>	25	1	1
Malacanthidae	Blue blanquillo	<i>Malacanthus latovittatus</i>	45	5	9
	Quakerfish	<i>Malacanthus brevisrostris</i>	30	2	5
Monacanthidae	Broom filefish	<i>Amanses scopas</i>	20	12	18
	Blackbar filefish	<i>Pervagor janthinosoma</i>	13.5	1	1
Mullidae	Yellowfin goatfish	<i>Mulloidichthys vanicolensis</i>	38	16	100
	Doublebar goatfish	<i>Parupeneus crassilabris</i>	35	5	8
	Yellowstripe goatfish	<i>Mulloidichthys flavolineatus</i>	40	3	16
	Dash-and-dot goatfish	<i>Parupeneus barberinus</i>	50	1	1
	Finstripe goatfish	<i>Upeneus taeniopterus</i>	33	1	1
Pinguipedidae	Latticed sandperch	<i>Parapercis clathrata</i>	18	31	59
	Black dotted sand perch	<i>Parapercis millepunctata</i>	19	2	2
Pomacanthidae	Threespot angelfish	<i>Apolemichthys trimaculatus</i>	25	13	20
Tetraodontidae	Guinea fowl puffer	<i>Arothron meleagris</i>	50	4	5
	Blackspotted puffer	<i>Arothron nigropunctatus</i>	33	2	2
	Map puffer	<i>Arothron mappa</i>	65	1	1

Planktivores					
Balistidae	Redtoothed triggerfish	<i>Odonus niger</i>	40	8	375
Blenniidae	Bicolor blenny	<i>Ecsenius bicolor</i>	10	15	65
	Forktail blenny	<i>Meiacanthus atrodorsalis</i>	10	11	123
Caesionidae	Dark-banded fusilier	<i>Pterocaesio tile</i>	30	18	662
	Blue and gold fusilier	<i>Caesio caeruleaurea</i>	35	10	139
	Marr's fusilier	<i>Pterocaesio marri</i>	35	3	68
Chaetodontidae	Pyramid butterflyfish	<i>Hemitaenichthys polylepis</i>	18	6	41
	Sunburst butterflyfish	<i>Chaetodon kleinii</i>	14	1	1
Clupeidae	Delicate round herring	<i>Spratelloides delicatulus</i>	7	1	60
Holocentridae	Shadowfin soldierfish	<i>Myripristis adusta</i>	32	1	1
Labridae	Sixbar wrasse	<i>Thalassoma hardwicke</i>	20	16	29
	Red shoulder wrasse	<i>Stethojulis bandanensis</i>	16	12	24
Microdesmidae	Blacktail goby	<i>Ptereleotris heteroptera</i>	12	7	59
	Chinese zebra goby	<i>Ptereleotris zebra</i>	12	4	106
Pomacentridae	Threespot dascyllus	<i>Dascyllus trimaculatus</i>	14	11	133
	Ternate chromis	<i>Chromis ternatensis</i>	8	5	21
	Indo-Pacific sergeant	<i>Abudefduf vaigiensis</i>	20	5	7
	Charcoal damsel	<i>Pomacentrus brachialis</i>	10	4	115
	Scissortail sergeant	<i>Abudefduf sexfasciatus</i>	15	2	21
	Pink anemonefish	<i>Amphiprion perideraion</i>	10	2	10
	Smokey chromis	<i>Chromis fumea</i>	10	1	50
	Sapphire devil	<i>Chrysiptera cyanea</i>	8	1	2
	Fire clownfish	<i>Amphiprion melanopus</i>	10.8	1	1
Herbivores					
Acanthuridae	Twospot surgeonfish	<i>Ctenochaetus binotatus</i>	22	8	13
	Bulbnose unicornfish	<i>Naso tonganus</i>	63	2	2
	Palelipped surgeonfish	<i>Acanthurus leucocheilus</i>	45	1	1
Blenniidae	Barred blenny	<i>Cirripectes polyzona</i>	10	4	19
	Red-streaked blenny	<i>Cirripectes stigmaticus</i>	10	2	20
Chaetodontidae	Atoll butterflyfish	<i>Chaetodon mertensii</i>	13	5	10
Kyphosidae	Blue seachub	<i>Kyphosus cinerascens</i>	51	6	14
Ostraciidae	Yellow boxfish	<i>Ostracion cubicus</i>	45	1	1
Pomacanthidae	Bicolor angelfish	<i>Centropyge bicolor</i>	15	20	62
	Twinspot damselfish	<i>Chrysiptera biocellata</i>	7	3	66
	Ambon chromis	<i>Chromis amboinensis</i>	8	2	62
	Banded sergeant	<i>Abudefduf septemfasciatus</i>	23	2	3
	Dusky farmerfish	<i>Stegastes nigricans</i>	14	1	18
	Whitebar gregory	<i>Stegastes albifasciatus</i>	11	1	3
Scaridae	Common parrotfish	<i>Scarus psittacus</i>	33	19	72
	Bicolour parrotfish	<i>Cetoscarus ocellatus</i>	90	9	11
	Dusky parrotfish	<i>Scarus niger</i>	43	8	9
	Filament-finned parrotfish	<i>Scarus altipinnis</i>	60	5	6
	Festive parrotfish	<i>Scarus festivus</i>	45	1	1
Siganidae	Streamlined spinefoot	<i>Siganus argenteus</i>	42	2	3
Tetraodontidae	Valentini's sharpnosepuffer	<i>Canthigaster valentini</i>	9	1	2



## Appendix E: Testing the Robustness of Human Disturbance Categories

Human population density varies considerably around the most heavily populated of the surveyed islands, such that at the smallest scale (within island) scale,  $\beta$ -diversity could reflect differences between sites that are very heavily disturbed (i.e. near the population centre) and sites with lower human disturbance (i.e. on the same islands, but far from population centre), instead of reflecting the  $\beta$ -diversity of fish communities across sites that are equally disturbed. To test this, I re-calculated  $\beta_{RC}$  and  $\beta_W$  at the smallest scale for the most heavily populated islands in the three island groups (Maui, Oahu, Guam, and Tutuila), using only those sites with human populations exceeding 50,000 within a 20 km radius of the surveyed site. These calculations yielded generally similar  $\beta$ -diversity values for those islands compared to using all sites on those islands (Supplemental Table 9). Similarly, the proportions of variability in community composition (using either incidence based or  $\alpha$ -constrained abundance based  $\beta$ -deviation metrics of site-site dissimilarity,  $\beta_{RC}$  or  $\beta_W$  respectively) explained by environmental predictors did not differ in a consistent manner from other sites on each respective island ( $p > 0.05$ ; Supplemental Table 10). Therefore, I suggest that using only sites near population centres would yield similar results throughout my thesis with regards to the effect of human disturbance.

**Supplemental Table 9.** Within island  $\beta_{RC}$  and  $\beta_W$  for heavily populated islands calculated using subsets of sites that had human populations of greater than 50000 individuals living within a 20 km radius.

	$\beta_{RC}$		$\beta_W$	
	All sites	Subset	All sites	Subset
Maui	3.8	2.9	-0.65	-0.71
Oahu	3	2.9	-0.78	-0.80
Guam	9	8.6	-0.48	-0.46
Tutuila	9.8	9.5	-0.41	-0.45

**Supplemental Table 10.** Mean proportion of variability of community composition within islands explained by interannual variability of sea surface temperature ( $SST_v$ ), mean annual sea surface temperature ( $SST_m$ ), reef rugosity (rug.), live coral cover (cover), and a combination of all four (Env.). Variance explained was derived from the mean of 100 distance-based redundancy analysis (dbRDA) performed separately on 10 sites from all sites on each island (All) or only sites with a human populations of greater than 50000 individuals living within a 20 km radius (High Pop. Subset). Raup-Crick dissimilarity and  $\alpha$ -constraining  $\beta$ -deviations ( $\beta_w$ ) based dissimilarity were each used for separate dbRDA.

Island	Sites Used	Env.	$SST_v$	$SST_m$	Rug.	Cover
<b>Incidence based (Raup-Crick) site-site dissimilarity</b>						
Maui	All	0.473	0.093	0.082	0.166	0.128
	High Pop. Subset	0.376	0.089	0.082	0.082	0.098
Oahu	All	0.432	0.097	0.087	0.117	0.093
	High Pop. Subset	0.458	0.090	0.087	0.100	0.157
Guam	All	0.72	0.112	0.116	0.127	0.142
	High Pop. Subset	0.456	0.110	0.116	0.109	0.131
Tutuila	All	0.446	0.112	0.110	0.129	0.129
	High Pop. Subset	0.453	0.101	0.110	0.114	0.114
<b>Abundance based (<math>\alpha</math>-constrained <math>\beta</math>-deviation (<math>\beta_w</math>)) site-site dissimilarity</b>						
Maui	All	0.570	0.107	0.141	0.188	0.233
	High Pop. Subset	0.589	0.183	0.141	0.185	0.160
Oahu	All	0.548	0.095	0.142	0.244	0.163
	High Pop. Subset	0.442	0.129	0.142	0.094	0.132
Guam	All	0.537	0.118	0.126	0.174	0.186
	High Pop. Subset	0.475	0.107	0.126	0.121	0.140
Tutuila	All	0.484	0.128	0.114	0.145	0.116
	High Pop. Subset	0.457	0.115	0.114	0.111	0.124

## Appendix F: $\beta$ -diversity Models Without Live Coral Complexity

Models for  $\beta_{RC}$  and  $\beta_W$  at all three scales were built without the inclusion of live coral complexity as an explanatory variable, and outputs shown below. Because human disturbance on coral reefs is not restricted to direct removal of fishes by fishing, habitat rugosity or live coral cover could be altered due to human disturbance. Indeed, habitat rugosity was 13% lower on islands with high human disturbance (island mean  $\pm$  SD =  $2.22 \pm 0.37$ ) compared to islands with low human disturbance (island mean  $\pm$  SD =  $2.56 \pm 0.45$ ; student's t-test:  $p = 0.024$ ). This difference was in part driven by islands with high human disturbance having sites with very low habitat rugosity (mean  $\pm$  SD of lowest site rugosity per islands with high and low human disturbance =  $1.04 \pm 0.14$  and  $1.24 \pm 0.34$  respectively; student's t-test:  $p = 0.032$ ). In contrast, there was no difference in the highest site rugosity (mean  $\pm$  SD of highest site rugosity on islands with high and low human disturbance =  $4.31 \pm 0.77$  and  $4.29 \pm 0.84$  respectively; student's t-test:  $p = 0.932$ ), the mean range of site rugosity (mean  $\pm$  SD of range of site rugosity on islands with high and low human disturbance =  $3.27 \pm 0.80$  and  $3.05 \pm 1.00$  respectively; student's t-test:  $p = 0.500$ ), or the variance in the site rugosity among islands of the two disturbance categories (mean  $\pm$  SD of variance in site rugosity on islands with high and low human disturbance =  $0.91 \pm 0.27$  and  $0.94 \pm 0.27$  respectively; student's t-test:  $p = 0.734$ ). As such, modelling  $\beta$ -diversity with both human disturbance and live coral complexity (a measure that incorporates habitat rugosity and live coral complexity) could detect the effect of human disturbance in two separate parameters, and therein result in

erratic behaviour. I therefore modelled  $\beta$ -diversity without a habitat variable as any difference in habitat variables among islands with low or high human disturbance likely reflects differences due to human disturbance corals. In these models, the estimated effect of the human disturbance parameter is similar at all scales to those models that include both human disturbance and live coral complexity.

**Supplemental Table 11.** Candidate models of  $\beta_{RC}$  where live coral complexity was not included as an explanatory variable at all three scales: within island  $\beta_{RC,1}$ ; within island group  $\beta_{RC,2}$ ; and across island group  $\beta_{RC,3}$ . AIC weights ( $AIC_w$ ) were calculated to sum to 1 over candidate models with  $\Delta AIC \leq 2$ . Pseudo  $R^2$  indicates the proportion of deviation explained by model relative to deviation in null (intercept only) model.  $SST_v$  and  $SST_v$  denote mean and interannual variability in sea surface temperature.

Model	$AIC_w$	$\Delta AIC$	Pseudo $R^2$
<b>Within island <math>\beta_{RC,1}</math></b>			
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Reef Area}$	0.181	0	0.443
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SSTv} + \text{Reef Area}$	0.181	0.001	0.509
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group}$	0.125	0.738	0.394
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Reef Area}$	0.115	0.907	0.351
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SSTv}$	0.095	1.28	0.456
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{SSTv} + \text{Reef Area}$	0.083	1.554	0.451
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{Reef Area}$	0.08	1.639	0.450
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{SSTm} + \text{Reef Area}$	0.073	1.804	0.447
$\beta_{RC,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SSTv} + \text{SSTm} + \text{Reef Area}$	0.068	1.954	0.510
<b>Within island group <math>\beta_{RC,2}</math></b>			
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Wave Energy} + \text{SSTv} + \text{SSTm}$	0.407	0	0.643
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Wave Energy} + \text{SSTv}$	0.272	0.81	0.603
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Wave Energy} + \text{SSTv} + \text{SSTm} + \text{Reef Area}$	0.164	1.815	0.645
$\beta_{RC,2} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SSTv} + \text{SSTm}$	0.157	1.91	0.644
<b>Across island group <math>\beta_{RC,3}</math></b>			
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SSTv}$	0.279	0	0.745
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SSTv} + \text{Reef Area}$	0.262	0.121	0.760
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SSTv} + \text{SSTm}$	0.238	0.319	0.758
$\beta_{RC,3} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + \text{SSTv} + \text{SSTm} + \text{Reef Area}$	0.222	0.459	0.772

**Supplemental Table 12.** Model weighted parameter estimation for incidence based Raup-Crick  $\beta$ -diversity,  $\beta_{RC}$ , where live coral complexity was not included as a potential explanatory variable. Weighting using all models within 2  $\Delta$ AIC of best model. Intercept represents an island in the Hawaiian Archipelago. Unconditional variance (Variance) as per Buckland et al. (1997); # Models = the number of models within 2  $\Delta$ AIC that each parameter was included in.

	Estimate	Variance	# Models
Within Island $\beta_{RC,1}$ Models			
(Intercept)	-0.796615469	0.005333737	9
Humans	0.176596991	0.00299572	9
Mariana Archipelago	0.116149032	0.01195237	8
American Samoa	0.001799575	0.012157315	8
Reef Area	-0.03364933	0.00085652	7
Wave Energy	0.050064601	0.005667206	4
SSTv	-0.027089139	0.001622495	4
SSTm	-0.0006243	4.05E-05	2
Within Island Group $\beta_{RC,2}$ Models			
(Intercept)	-0.577583191	0.000968377	4
Humans	0.213770506	0.002339215	4
Wave Energy	0.141824724	0.003166148	4
SSTv	-0.120587499	0.001494186	4
SSTm	0.035773063	0.001227049	3
Mariana Archipelago	-0.004180609	0.00038705	1
Reef Area	-0.001474708	2.81E-05	1
Across Island Group $\beta_{RC,3}$ Models			
(Intercept)	1.061653837	0.004730621	4
Humans	-0.227309743	0.002337768	4
Mariana Archipelago	-0.343595359	0.009847519	4
American Samoa	-0.406383561	0.012885512	4
Wave Energy	-0.182488351	0.005270743	4
SSTv	0.104000877	0.001490319	4
SSTm	-0.017767678	0.000775288	2
Reef Area	0.013265173	0.000389347	2

**Supplemental Table 13.** Candidate models of  $\beta_w$  where live coral complexity was not included as an explanatory variable at all three scales: within island  $\beta_{w,1}$ ; within island group  $\beta_{w,2}$ ; and across island group  $\beta_{w,3}$ . AIC weights ( $AIC_w$ ) were calculated to sum to 1 over candidate models with  $\Delta AIC \leq 2$ . Pseudo  $R^2$  indicates the proportion of deviation explained by model relative to deviation in null (intercept only) model.  $SST_v$  and  $SST_v$  denote mean and interannual variability in sea surface temperature.

Model	$AIC_w$	$\Delta AIC$	Pseudo $R^2$
<b>Within island <math>\beta_{w,1}</math></b>			
$\beta_{w,1} \sim \text{Intercept} + \text{Island Group}$	0.254	0	0.837
$\beta_{w,1} \sim \text{Intercept} + \text{Island Group} + \text{Reef Area}$	0.18	0.691	0.844
$\beta_{w,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group}$	0.154	1.005	0.842
$\beta_{w,1} \sim \text{Intercept} + \text{Island Group} + SSTm$	0.127	1.38	0.840
$\beta_{w,1} \sim \text{Intercept} + \text{Island Group} + \text{Wave Energy}$	0.097	1.925	0.837
$\beta_{w,1} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Reef Area}$	0.095	1.955	0.847
$\beta_{w,1} \sim \text{Intercept} + \text{Island Group} + SSTv$	0.093	1.997	0.837
<b>Within island group <math>\beta_{w,2}</math></b>			
$\beta_{w,2} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + SSTv$	0.443	0	0.981
$\beta_{w,2} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + SSTv + \text{Reef Area}$	0.2	1.591	0.981
$\beta_{w,2} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + SSTv + SSTm$	0.186	1.73	0.981
$\beta_{w,2} \sim \text{Intercept} + \text{Humans} + \text{Island Group} + \text{Wave Energy} + SSTv$	0.171	1.899	0.981
<b>Across island group <math>\beta_{w,3}</math></b>			
$\beta_{w,3} \sim \text{Intercept} + \text{Humans} + SSTv$	0.337	0	0.542
$\beta_{w,3} \sim \text{Intercept} + \text{Humans} + SSTv + \text{Wave Energy}$	0.249	0.61	0.561
$\beta_{w,3} \sim \text{Intercept} + \text{Humans} + SSTv \text{ Reef Area}$	0.226	0.8	0.559
$\beta_{w,3} \sim \text{Intercept} + \text{Humans} + SSTv + SSTm$	0.189	1.15	0.554

**Supplemental Table 14.** Model weighted parameter estimation for  $\alpha$ -constraining abundance-based null model  $\beta$ -diversity,  $\beta_w$ , where live coral complexity was not included as a potential explanatory variable. Weighting using all models within 2  $\Delta$ AIC of best model. Intercept represents an island in the Hawaiian Archipelago. Unconditional variance (Variance) as per Buckland et al. (1997); # Models = the number of models within 2  $\Delta$ AIC that each parameter was included in.

	Estimate	Variance	# Models
Within Island $\beta_{w,1}$ Models			
(Intercept)	3.516581556	0.193402555	7
Mariana Archipelago	5.900807206	0.425370911	7
Island GroupSAMOA	4.844181819	0.712703278	7
Humans	-0.107775537	0.051768977	2
Reef Area	-0.082394896	0.02538749	2
SSTv	-0.001318627	0.000745251	1
Wave Energy	0.011378864	0.003125939	1
SSTm	-0.038710819	0.009223904	1
Within Island Group $\beta_{w,2}$ Models			
(Intercept)	5.65625684	0.031409162	4
Humans	-1.594337608	0.042264336	4
Mariana Archipelago	5.273603202	0.066976186	4
SSTv	-0.429977471	0.013573967	4
Wave Energy	0.015121245	0.004438407	1
SSTm	0.012167144	0.001415122	1
Reef Area	-0.011987311	0.001049181	1
Across Island Group $\beta_{w,3}$ Models			
(Intercept)	8.468162478	0.012432364	4
Humans	-0.55367442	0.037689694	4
SSTv	0.374389593	0.013545238	4
SSTm	0.015354951	0.001162692	1
Reef Area	-0.021100784	0.001803552	1
Wave Energy	-0.040679087	0.006101191	1

## Bibliography

- Acosta-González, G., F. A. Rodríguez-Zaragoza, R. C. Hernández-Landa, and J. E. Arias-González. 2013. Additive diversity partitioning of fish in a Caribbean coral reef undergoing shift transition. *PLoS ONE* 8:e65665.
- Allen, G. R. 2008. Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems* 18:541–556.
- Allen, G. R., and M. Adrim. 2003. Coral reef fishes of Indonesia. *Zoological Studies-Taipei* 42(1): 1-72.
- Anderson, M. J., K. E. Ellingsen, and B. H. McArdle. 2006. Multivariate dispersion as a measure of beta diversity. *Ecology Letters* 9:683–693.
- Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, A. L. Freestone, N. J. Sanders, H. V. Cornell, L. S. Comita, K. F. Davies, S. P. Harrison, N. J. B. Kraft, J. C. Stegen, and N. G. Swenson. 2011. Navigating the multiple meanings of  $\beta$  diversity: a roadmap for the practicing ecologist. *Ecology Letters* 14:19–28.
- Anderson, T. W., M. H. Carr, and M. A. Hixon. 2007. Patterns and mechanisms of variable settlement and recruitment of a coral reef damselfish, *Chromis cyanea*. *Marine Ecology Progress Series* 350:109–116.
- Arias-González, J., P. Legendre, and F. Rodríguez-Zaragoza. 2008. Scaling up beta diversity on Caribbean coral reefs. *Journal of Experimental Marine Biology and Ecology*.
- Ayotte, P., K. McCoy, I. Williams, and J. Zamzow. 2011. Coral Reef Ecosystem Division standard operating procedures: data collection for rapid ecological assessment fish surveys. Pages 1–30.
- Becking, L. E., D. F. R. Cleary, N. J. Voogd, W. Renema, M. Beer, R. W. M. Soest, and B. W. Hoeksema. 2006. Beta diversity of tropical marine benthic assemblages in the Spermonde Archipelago, Indonesia. *Marine Ecology* 27:76–88.
- Bell, J. D., and R. Galzin. 1984. Influence of live coral cover on coral-reef fish communities. *Mar Ecol Prog Ser* 15:265–274.
- Bellwood, D. R., and T. P. Hughes. 2001. Regional-scale assembly rules and biodiversity of coral reefs. *Science* 292:1532–1534.
- Bellwood, D. R., T. P. Hughes, C. Folke, and M. Nyström. 2004. Confronting the coral reef crisis. *Nature* 429:827–833.
- Bellwood, D., P. Wainwright, C. Fulton, and A. Hoey. 2006a. Functional versatility supports coral reef biodiversity. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 273:101-107.
- Bellwood, D., T. Hughes, and A. Hoey. 2006b. Sleeping functional group drives coral-reef recovery. *Current Biology* 16:2434-2439.
- Bellwood, D., T. Hughes, S. Connolly, and J. Tanner. 2005. Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. *Ecology Letters* 8:643-651.
- Belmaker, J., Y. Ziv, N. Shashar, and S. R. Connolly. 2008. Regional variation in the hierarchical partitioning of diversity in coral-dwelling fishes. *Ecology* 89:2829–2840.



- Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* 153:51-68.
- Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73:1045–1055.
- Brandt, M. E., N. Zurcher, A. Acosta, J. S. Ault, J. A. Bohnsack, M. W. Feeley, D. E. Harper, J. H. Hunt, T. Kellison, D. B. McClellan, M. E. Patterson, and S. G. Smith. 2009. A cooperative multi-agency reef fish monitoring protocol for the Florida Keys coral reef ecosystem:1–128.
- Bray, J. R., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern wisconsin. *Ecological Monographs* 27:325–349.
- Bruno, J. F., L. E. Petes, C. Drew Harvell, and A. Hettinger. 2003. Nutrient enrichment can increase the severity of coral diseases. *Ecology Letters* 6:1056–1061.
- Buckland, S. T., K. P. Burnham, and N. H. Augustin. 1997. Model selection: an integral part of inference. *Biometrics* 53:603–618.
- Burke, L., K. Reytar, M. Spalding, and A. Perry. 2011. *Reefs at Risk: Revisited*.
- Burman, S. G., R. B. Aronson, and R. van Woelik. 2012. Biotic homogenization of coral assemblages along the Florida reef tract. *Marine Ecology Progress Series* 467:89-96.
- Burnham, K. P., and D. R. Anderson. 1998. *Model selection and multimodel inference - a practical information-theoretic approach*, 2nd edition. Springer, New York.
- Cadotte, M. W., K. Carscadden, and N. Mirotchnick. 2011. Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology* 48:1079–1087.
- Carpenter, K. E., and V. G. Springer. 2005. The center of the center of marine shore fish biodiversity: the Philippine Islands. *Environmental Biology of Fishes* 72:467–480.
- Casey, K. S., T. B. Brandon, P. Cornillon, and R. Evans. 2010. The Past, Present, and Future of the AVHRR Pathfinder SST Program. Pages 273–287 *in* *Oceanography from Space*. Springer Netherlands, Dordrecht.
- Cassey, P., J. L. Lockwood, T. M. Blackburn, and J. D. Olden. 2007. Spatial scale and evolutionary history determine the degree of taxonomic homogenization across island bird assemblages. *Diversity and Distributions* 13:458–466.
- Chao, A., R. L. Chazdon, R. K. Colwell, and T. J. Shen. 2006. Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics* 62:361–371.
- Chase, J. M. 2007. Drought mediates the importance of stochastic community assembly. *Proceedings of the National Academy of Sciences of the United States of America* 104:17430–17434.
- Chase, J. M. 2010. Stochastic community assembly causes higher biodiversity in more productive environments. *Science* 328:1388–1391.
- Chase, J. M., and J. A. Myers. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B* 366L2351-2363.
- Chase, J. M., E. G. Biro, W. A. Ryberg, and K. G. Smith. 2009. Predators temper the relative importance of stochastic processes in the assembly of prey metacommunities. *Ecology Letters* 12:1210–1218.
- Cinner, J., T. R. McClanahan, T. M. Daw, N. A. Graham, J. Maina, S. K. Wilson, and T.

- P. Hughes. 2009. Linking social and ecological systems to sustain coral reef fisheries. *Current Biology* 19:206–212.
- Clarke, R. D. 1977. Habitat distribution and species diversity of chaetodontid and pomacentrid fishes near Bimini, Bahamas. *Marine Biology* 40:277–289.
- Coles, S. L., R. C. DeFelice, L. G. Eldredge, and J. T. Carlton. 1999. Historical and recent introductions of non-indigenous marine species into Pearl Harbor, Oahu, Hawaiian Islands. *Marine Biology* 135:147–158.
- Condit, R., R. Condit, N. Pitman, E. G. Leigh Jr, J. E. Chave, J. Terborgh, R. B. Foster, P. Nunez V, S. Aguilar, R. Valencia, G. Villa, H. C. Muller-Landau, E. Losos, and S. P. Hubbell. 2002. Beta-diversity in tropical forest trees. *Science* 295:666–669.
- Connell, J. H. 1978. Diversity in tropical rain forests and coral reefs. *Science* 199:1302–1310.
- Connell, S. D., and M. J. Kingsford. 1998. Spatial, temporal and habitat-related variation in the abundance of large predatory fish at One Tree Reef, Australia. *Coral Reefs* 17:49–57.
- Connolly, S. R., T. P. Hughes, D. R. Bellwood, and R. H. Karlson. 2005. Community structure of corals and reef fishes at multiple scales. *Science* 309:1363–1365.
- Connolly, S., and D. Bellwood. 2003. Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. *Ecology* 84:2007–2020.
- Connor, E. F., and D. Simberloff. 1979. The assembly of species communities: chance or competition? *Ecology* 60:1132–1140.
- Cornell, H. V., R. H. Karlson, and T. P. Hughes. 2007. Scale-dependent variation in coral community similarity across sites, islands, and island groups. *Ecology* 88:1707–1715.
- Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*.
- Cowen, R. K., C. B. Paris, and A. Srinivasan. 2006. Scaling of connectivity in marine populations. *Science* 311:522–527.
- Crist, T. O., J. A. Veech, J. C. Gering, and K. S. Summerville. 2003. Partitioning species diversity across landscapes and regions: a hierarchical analysis of  $\alpha$ ,  $\beta$ , and  $\gamma$  diversity. *The American Naturalist* 162:734–743.
- DeMartini, E., A. Friedlander, S. Sandin, and E. Sala. 2008. Differences in fish-assemblage structure between fished and unfished atolls in the northern Line Islands, central Pacific. *Marine Ecology Progress Series* 365:199–215.
- Dexter, K. G., J. W. Terborgh, and C. W. Cunningham. 2012. Historical effects on beta diversity and community assembly in Amazonian trees. *Proceedings of the National Academy of Sciences of the United States of America* 109:7787–7792.
- Diamond, J. M., and M. E. Gilpin. 1982. Examination of the "null" model of Connor and Simberloff for species co-occurrences on Islands. *Oecologia* 52:64–74.
- Dobson, A. J. 2003. *An Introduction to Generalized Linear Models*, 2nd edition. Chapman & Hall/CRC.
- Doherty, P., and T. Fowler. 1994. An empirical test of recruitment limitation in a coral reef fish. *Science* 263:935–939.
- Dollar, S. J. 1982. Wave stress and coral community structure in Hawaii. *Coral Reefs* 1:71–81.
- Donohue, I., A. L. Jackson, M. T. Pusch, and K. Irvine. 2009. Nutrient enrichment

- homogenizes lake benthic assemblages at local and regional scales. *Ecology* 90:3470–3477.
- Dornelas, M., S. R. Connolly, and T. P. Hughes. 2006. Coral reef diversity refutes the neutral theory of biodiversity. *Nature* 440:80–82.
- Dulvy, N. K., N. Polunin, and A. C. Mill. 2004a. Size structural change in lightly exploited coral reef fish communities: evidence for weak indirect effects. *Canadian Journal of Fisheries and Aquatic Sciences* 61:466–475.
- Dulvy, N. K., R. P. Freckleton, and N. V. C. Polunin. 2004b. Coral reef cascades and the indirect effects of predator removal by exploitation. *Ecology Letters* 7:410–416.
- Dulvy, N. K., Y. Sadovy, and J. D. Reynolds. 2003. Extinction vulnerability in marine populations. *Fish and Fisheries* 4:25–64.
- Edwards, C. B., A. M. Friedlander, A. G. Green, M. J. Hardt, E. Sala, H. P. Sweatman, I. D. Williams, B. Zgliczynski, S. A. Sandin, and J. E. Smith. 2013. Global assessment of the status of coral reef herbivorous fishes: evidence for fishing effects. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 281:20131835.
- Figueira, W., and D. Booth. 2010. Increasing ocean temperatures allow tropical fishes to survive overwinter in temperate waters. *Global Change Biology* 16:506–516.
- Forrester, G. E. 1995. Strong density-dependent survival and recruitment regulate the abundance of a coral reef fish. *Oecologia* 103:275–282.
- Fox, R., T. Sunderland, A. Hoey, and D. Bellwood. 2009. Estimating ecosystem function: contrasting roles of closely related herbivorous rabbitfishes (Siganidae) on coral reefs. *Marine Ecology Progress Series* 385:261–269.
- Francisco-Ramos, V., and J. E. Arias-González. 2013. Additive partitioning of coral reef fish diversity across hierarchical spatial scales throughout the caribbean. *PLoS ONE* 8(10):e78761.
- Friedlander, A. M., and E. E. DeMartini. 2002. Contrasts in density, size, and biomass of reef fishes between the northwestern and the main Hawaiian islands: the effects of fishing down apex predators. *Marine Ecology Progress Series* 230:253–264.
- Friedlander, A. M., and J. D. Parrish. 1998. Habitat characteristics affecting fish assemblages on a Hawaiian coral reef. *Journal of Experimental Marine Biology and Ecology* 224:1–30.
- Fulton, C. J., and D. R. Bellwood. 2005. Wave-induced water motion and the functional implications for coral reef fish assemblages. *Limnology and Oceanography* 50:255–264.
- Fulton, C. J., D. R. Bellwood, and P. C. Wainwright. 2005. Wave Energy and Swimming Performance Shape Coral Reef Fish Assemblages. *Proceedings of the Royal Society B* 272:827–832.
- Gardner, T. A., I. M. Côté, J. A. Gill, A. Grant, and A. R. Watkinson. 2003. Long-term region-wide declines in Caribbean corals. *Science* 301:958–960.
- Gaston, K. J. 2000. Global patterns in biodiversity. *Nature* 405:220–227.
- Gleason, H. A. 1920. Some applications of the quadrat method. *Bulletin of the Torrey Botanical Club* 47:21–33.
- Gotelli, N. J., and G. R. Graves. 1996. *Null Models in Ecology*. Smithsonian Institution Press, Washington.
- Gove, J. M., G. J. Williams, M. A. McManus, S. F. Heron, S. A. Sandin, O. J. Vetter, and

- D. G. Foley. 2013. Quantifying climatological ranges and anomalies for pacific coral reef ecosystems. *PLoS ONE* 8(4):e61974.
- 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 of the United States of America* 103:8425–8429.
- Graham, N., and R. D. Evans. 2003. The effects of marine reserve protection on the trophic relationships of reef fishes on the Great Barrier Reef. *Environmental Conservation* 30(2):200-208.
- Guillemot, N., M. Kulbicki, P. Chabanet, and L. Vigliola. 2011. Functional redundancy patterns reveal non-random assembly rules in a species-rich marine assemblage. *PLoS ONE* 6:e26735.
- Haapkylä, J., R. K. F. Unsworth, M. Flavell, D. G. Bourne, B. Schaffelke, and B. L. Willis. 2011. Seasonal Rainfall and Runoff Promote Coral Disease on an Inshore Reef. *PLoS ONE* 6:e16893.
- Harborne, A., P. Mumby, K. Zychaluk, J. Hedley, and P. Blackwell. 2006. Modeling the beta diversity of coral reefs. *Ecology* 87:2871–2881.
- Harvell, C. D. 2002. Climate warming and disease risks for terrestrial and marine biota. *Science* 296:2158–2162.
- Harvell, C. D., K. Kim, J. M. Burkholder, R. R. Colwell, P. R. Epstein, D. J. Grimes, E. E. Hofmann, E. K. Lipp, A. D. M. E. Osterhaus, R. M. Overstreet, J. W. Porter, G. W. Smith, and G. R. Vasta. 1999. Emerging marine diseases - climate links and anthropogenic factors. *Science* 285:1505–1510.
- Hearn, C. J. 1999. Wave-breaking hydrodynamics within coral reef systems and the effect of changing relative sea level. *Journal of Geophysical Research* 104:30007-30019.
- Hewitt, J. E., S. F. Thrush, J. Halliday, and C. Duffy. 2005. The importance of small-scale habitat structure for maintaining beta diversity. *Ecology* 86:1619–1626.
- Hiatt, R. W., and D. W. Strasburg. 1960. Ecological relationships of the fish fauna on coral reefs of the Marshall Islands. *Ecological Monographs* 30:65–127.
- Hixon, M. A. 2011. 60 years of coral reef fish ecology: past, present, future. *Bulletin of Marine Science* 87(4):727-765.
- Hoegh-Guldberg, O. 1999. Climate change, coral bleaching and the future of the world's coral reefs. *Marine and Freshwater Research* 50:839.
- Hoegh-Guldberg, O., P. J. Mumby, A. J. Hooten, R. S. Steneck, P. Greenfield, E. Gomez, C. D. Harvell, P. F. Sale, A. J. Edwards, K. Caldeira, N. Knowlton, C. M. Eakin, R. Iglesias-Prieto, N. Muthiga, R. H. Bradbury, A. Dubi, and M. E. Hatziolos. 2007. Coral reefs under rapid climate change and ocean acidification. *Science* 318:1737–1742.
- Hoey, A. S., and D. R. Bellwood. 2009. Limited functional redundancy in a high diversity system: single species dominates key ecological process on coral reefs. *Ecosystems* 12:1316–1328.
- Hubbell, S. P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.
- Hughes, T. P. 1994. Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. *Science* 265:1547–1551.
- Hughes, T. P., D. R. Bellwood, and S. R. Connolly. 2002. Biodiversity hotspots, centres

- of endemism, and the conservation of coral reefs. *Ecology Letters* 5:775–784.
- Hughes, T., A. Baird, D. Bellwood, M. Card, T. P. Hughes, A. H. Baird, D. R. Bellwood, M. Card, S. R. Connolly, C. Folke, R. Grosberg, O. Hoegh-Guldberg, J. B. C. Jackson, J. Kleypas, J. M. Lough, P. Marshall, M. Nyström, S. R. Palumbi, J. M. Pandolfi, B. Rosen, and J. Roughgarden. 2003. Climate change, human impacts, and the resilience of coral reefs. *Science* 301:929–933.
- Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22:415–427.
- Hutchinson, G. E. 1959. Homage to Santa Rosalia or why are there so many kinds of animals? *American Naturalist* 93:145–159.
- Jaccard, P. 1912. The Distribution of the flora in the alpine zone. *New Phytologist* 11:37–50.
- Jackson, J. B. C., M. X. Kirby, W. H. Berger, K. A. Bjorndal, L. W. Botsford, B. J. Bourque, R. H. Bradbury, R. Cooke, J. Erlandson, J. A. Estes, T. P. Hughes, S. Kidwell, C. B. Lange, H. S. Lenihan, J. M. Pandolfi, C. H. Peterson, R. S. Steneck, M. J. Tegner, and R. R. Warner. 2001. Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293:629–638.
- Jennings, S., and N. V. C. Polunin. 1997. Impacts of predator depletion by fishing on the biomass and diversity of non-target reef fish communities. *Coral Reefs* 16:71–82.
- Jennings, S., D. P. Boull, and N. V. C. Polunin. 1996. Habitat correlates of the distribution and biomass of Seychelles' reef fishes. *Environmental Biology of Fishes* 46:15–25.
- Jones, G. P., M. I. McCormick, M. Srinivasan, and J. V. Eagle. 2004. Coral decline threatens fish biodiversity in marine reserves. *Proceedings of the National Academy of Sciences of the United States of America* 101:8251–8253.
- Jones, G. P., M. J. Milicich, M. J. Emslie, and C. Lunow. 1999. Self-recruitment in a coral reef fish population. *Nature* 402:802–804.
- Jones, G. P., S. Planes, and S. R. Thorrold. 2005. Coral reef fish larvae settle close to home. *Current Biology* 15:1314–1318.
- Jones, R. S., and P. Helfrich. 1967. Ecological relationships in Hawaiian and Johnston Island Acanthuridae (Surgeonfishes) with emphasis on food and feeding habits. University of Hawaii.
- Karlson, R. H., H. V. Cornell, and T. P. Hughes. 2007. Aggregation influences coral species richness at multiple spatial scales. *Ecology* 88:170–177.
- Karp, D. S., A. J. Rominger, J. Zook, J. Ranganathan, P. R. Ehrlich, and G. C. Daily. 2012. Intensive agriculture erodes  $\beta$ -diversity at large scales. *Ecology Letters* 15:963–970.
- Kingsford, M. J., A. J. Underwood, and S. J. Kennelly. 1991. Humans as predators on rocky reefs in New South Wales, Australia. *Marine Ecology Progress Series* 72:1–14.
- Kinlan, B. P., and S. D. Gaines. 2003. Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology* 84:2007–2020.
- Knowlton, N. 2001. The future of coral reefs. *Proceedings of the National Academy of Sciences of the United States of America* 98:5419–5425.
- Kolasa, J., L. L. Manne, and S. N. Pandit. 2011. Species–area relationships arise from interaction of habitat heterogeneity and species pool. *Hydrobiologia* 685:135–144.
- Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence-

- absence data. *Journal of Animal Ecology* 72:367–382.
- Kraft, N. J. B., L. S. Comita, J. M. Chase, N. J. Sanders, N. G. Swenson, T. O. Crist, J. C. Stegen, M. Vellend, B. Boyle, M. J. Anderson, H. V. Cornell, K. F. Davies, A. L. Freestone, B. D. Inouye, S. P. Harrison, and J. A. Myers. 2011. Disentangling the drivers of beta diversity along latitudinal and elevational gradients. *Science* 333:1755–1758.
- Kramer, D. L., and M. R. Chapman. 1999. Implications of fish home range size and relocation for marine reserve function. *Environmental Biology of Fishes* 55:65–79.
- Kulbicki, M., Parravicini, V., Bellwood, D.R., Arias-Gonzalez, E., Chabanet, P., Floeter, S.R., Friedlander, A., McPherson, J., Myers, R.E., Vigliola, L., and D. Mouillot. 2013. Global biogeography of reef fishes: a hierarchical quantitative delineation of regions. *PLoS ONE* 8(12):e81847.
- Legendre, P., and L. Legendre. 2012. *Numerical Ecology*, 3rd edition. Elsevier.
- Legendre, P., and M. De Cáceres. 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters* 16:951–963.
- Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs* 75:435–450.
- Legendre, P., D. Borcard, and P. R. Peres-Neto. 2008. Analyzing or explaining beta diversity? Comment. *Ecology* 89:3238–3244.
- Levin, S. A. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology* 73:1943–1967.
- MacArthur, R., and E. Wilson. 1967. *The theory of island biogeography*. Princeton University Press, Princeton, N.J.
- Macneil, A. M., N. A. J. Graham, N. V. C. Polunin, M. Kulbicki, R. Galzin, M. Harmelin-Vivien, and S. P. Rushton. 2009. Hierarchical drivers of reef-fish metacommunity structure. *Ecology* 90:252–264.
- Maillefer, A. 1929. Le Coefficient générique de P. Jacard et sa signification. *Memoires de la societe vaudoise des sciences naturelles* 3(4):113-183.
- Mantyka, C., and D. Bellwood. 2007. Macroalgal grazing selectivity among herbivorous coral reef fishes. *Marine Ecology Progress Series* 352:177–185.
- Marchetti, M. P., J. L. Lockwood, and T. Light. 2006. Effects of urbanization on California's fish diversity: Differentiation, homogenization and the influence of spatial scale. *Biological Conservation* 127:310–318.
- Martins, G. M., F. Arenas, A. I. Neto, and S. R. Jenkins. 2012. Effects of fishing and regional species pool on the functional diversity of fish communities. *PLoS ONE* 7:e44297.
- McAllister, D. 1995. Status of the World Ocean and its Biodiversity. *Sea Wind*.
- McClanahan, T. 2005. Recovery trajectories of coral reef fish assemblages within Kenyan marine protected areas. *Marine Ecology Progress Series* 294:241-248.
- McKinney, M. L., and J. L. Lockwood. 1999. Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology & Evolution* 14:450–453.
- Mellin, C., C. Bradshaw, M. Meekan, and M. Caley. 2010. Environmental and spatial predictors of species richness and abundance in coral reef fishes. *Global Ecology and Biodiversity* 19:212-222.

- Mellin, C., J. Ferraris, R. Galzin, M. Harmelin-Vivien, M. Kulbicki, and T. L. de Loma. 2008. Natural and anthropogenic influences on the diversity structure of reef fish communities in the Tuamotu Archipelago (French Polynesia). *Ecological Modelling* 218:182–187.
- Messmer, V., G. P. Jones, P. L. Munday, S. J. Holbrook, R. J. Schmitt, and A. J. Brooks. 2011. Habitat biodiversity as a determinant of fish community structure on coral reefs. *Ecology* 92:2285–2298.
- Micheli, F., and B. S. Halpern. 2005. Low functional redundancy in coastal marine assemblages. *Ecology Letters* 8:391–400.
- Miller, J. R., and R. J. Hobbs. 2002. Conservation where people live and work. *Conservation Biology* 16:330–337.
- Mora, C., and P. F. Sale. 2002. Are populations of coral reef fish open or closed? *Trends in Ecology & Evolution* 17:422–428.
- Mora, C., P. M. Chittaro, P. F. Sale, J. P. Kritzer, and S. A. Ludsin. 2003. Patterns and processes in reef fish diversity. *Nature* 421:933–936.
- Myers, J. A., J. M. Chase, I. Jiménez, P. M. Jørgensen, A. Araujo-Murakami, N. Paniagua-Zambrana, and R. Seidel. 2012. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. *Ecology Letters* 16:ele.12021.
- Nadon, M. O., J. K. Baum, I. D. Williams, J. M. McPherson, B. J. Zgliczynski, B. L. Richards, and R. E. Schroeder. 2012. Re-creating missing population baselines for Pacific reef sharks. *Conservation Biology*:1–12.
- Nunes, F. L. D., R. D. Norris, and N. Knowlton. 2011. Long distance dispersal and connectivity in amphi-Atlantic corals at regional and basin Scales. *PLoS ONE* 6:e22298.
- Nyström, M. 2006. Redundancy and response diversity of functional groups: implications for the resilience of coral reefs. *AMBIO: A Journal of the Human Environment* 35:30–35.
- Olden, J. D. 2006. Biotic homogenization: a new research agenda for conservation biogeography. *Journal of Biogeography* 33:2027–2039.
- Olden, J. D., and N. L. Poff. 2003. Toward a mechanistic understanding and prediction of biotic homogenization. *The American Naturalist* 162.
- Olden, J. D., M. J. Kennard, and B. J. Pusey. 2007. Species invasions and the changing biogeography of Australian freshwater fishes. *Global Ecology and Biogeography* 17:25–37.
- Olden, J. D., N. L. Poff, and M. L. McKinney. 2006. Forecasting faunal and floral homogenization associated with human population geography in North America. *Biological Conservation* 127:261–271.
- Olden, J., N. Poff, M. Douglas, M. Douglas, and K. Fausch. 2004. Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution* 19:18–24.
- Pandolfi, J., J. Jackson, N. Baron, R. Bradbury, H. Guzman, T. Hughes, C. Kappel, F. Micheli, J. Ogden, H. Possingham, and E. Sala. 2005. Are U.S. coral reefs on the slippery slope to slime? *Science* 307:1725–1726.
- Pandolfi, J., R. Bradbury, E. Sala, T. Hughes, K. Bjorndal, R. Cooke, D. McArdle, L. McClenachan, M. Newman, G. Paredes, R. Warner, and J. Jackson. 2003. Global

- Trajectories of the long-term decline of coral reef ecosystems. *Science* 301:955–958.
- Parker, K. A., M. J. Anderson, P. F. Jenkins, and D. H. Brunton. 2012. The effects of translocation-induced isolation and fragmentation on the cultural evolution of bird song. *Ecology Letters* 15:778–785.
- Polunin, N., and C. M. Roberts. 1993. Greater biomass and value of target coral-reef fishes in two small Caribbean marine reserves. *Marine Ecology Progress Series* 100:167–167.
- Porter, J. W., P. Dustan, W. C. Jaap, K. L. Patterson, V. Kosmynin, O. W. Meier, M. E. Patterson, and M. Parsons. 2001. Patterns of spread of coral disease in the Florida Keys. *Hydrobiologia* 460:1–24.
- Pratchett, M. S., A. S. Hoey, S. K. Wilson, V. Messmer, and N. A. J. Graham. 2011. Changes in biodiversity and functioning of reef fish assemblages following coral bleaching and coral loss. *Diversity* 3:424–452.
- Rahel, F. J. 2000. Homogenization of fish faunas across the United States. *Science* 288:854–856.
- Randall, J. E. 1987. Introductions of marine fishes to the Hawaiian Islands. *Bulletin of Marine Science* 41:490–502.
- Raup, D. M., and R. E. Crick. 1979. Measurement of faunal similarity in paleontology. *Journal of Paleontology*:1213–1227.
- Reaka-Kudla, M. L. 1997. The global biodiversity of coral reefs: a comparison with rain forests. Pages 83–108 *in* M. L. Reaka-Kudla, D. E. Wilson, and E. O. Wilson, editors. *Biodiversity II: Understanding and Protecting Our Biological Resource*. Joseph Henry Press, Washington, DC.
- Riegl, B. 2002. Effects of the 1996 and 1998 positive sea-surface temperature anomalies on corals, coral diseases and fish in the Arabian Gulf (Dubai, UAE). *Marine Biology* 140:29–40.
- Rocha, L. A., and B. W. Bowen. 2008. Speciation in coral-reef fishes. *Journal of Fish Biology* 72:1101–1121.
- Rodríguez-Zaragoza, F. A. 2008. Additive diversity partitioning of reef fishes across multiple spatial scales. *Caribbean Journal of Science* 44:90–101.
- Rodríguez-Zaragoza, F. A., A. L. Cupul-Magaña, C. M. Galván-Villa, E. Ríos-Jara, M. Ortiz, E. G. Robles-Jarero, E. López-Uriarte, and J. E. Arias-González. 2011. Additive partitioning of reef fish diversity variation: a promising marine biodiversity management tool. *Biodiversity and Conservation* 20:1655–1675.
- Rogers, C. S., and J. Miller. 2006. Permanent “phase shifts” or reversible declines in coral cover? Lack of recovery of two coral reefs in St. John, US Virgin Islands. *Marine Ecology Progress Series* 306:103–114.
- Rohmann, S. O., J. J. Hayes, R. C. Newhall, M. E. Monaco, and R. W. Grigg. 2005. The area of potential shallow-water tropical and subtropical coral ecosystems in the United States. *Coral Reefs* 24:370–383.
- Rosen, B. R. 1988. Progress, problems and patterns in the biogeography of reef corals and other tropical marine organisms. *Helgoländer Meeresuntersuchungen* 42:269–301.
- Sale, P. F. 1977. Maintenance of high diversity in coral reef fish communities. *The American Naturalist* 111:337–359.
- Sale, P. F. (Ed.). 2006. *Coral Reef Fishes: Dynamics and Diversity in a Complex*



- Ecosystem, 1st edition. Elsevier, Burlington, MA, USA.
- Sale, P. F., and W. A. Douglas. 1984. Temporal variability in the community structure of fish on coral patch reefs and the relation of community structure to reef structure. *Ecology* 65:409–422.
- Sanciango, J. C., K. E. Carpenter, P. J. Etnoyer, and F. Moretzsohn. 2013. Habitat availability and heterogeneity and the Indo-Pacific warm pool as predictors of marine species richness in the tropical Indo-Pacific. *PLoS ONE* 8:e56245.
- Sandin, S. A., J. E. Smith, E. E. DeMartini, E. A. Dinsdale, S. D. Donner, A. M. Friedlander, T. Konotchick, M. Malay, J. E. Maragos, D. Obura, O. Pantos, G. Paulay, M. Richie, F. Rohwer, R. E. Schroeder, S. Walsh, J. B. C. Jackson, N. Knowlton, and E. Sala. 2008. Baselines and degradation of coral reefs in the Northern Line Islands. *PLoS ONE* 3:e1548.
- Sbrocco, E. J., and P. H. Barber. 2013. MARSPEC: ocean climate layers for marine spatial ecology. *Ecology* 94:979–979.
- Selig, E. R., K. S. Casey, and J. F. Bruno. 2010. New insights into global patterns of ocean temperature anomalies: implications for coral reef health and management. *Global Ecology and Biogeography* 19:397–411.
- Shanks, A. L. 2009. Pelagic larval duration and dispersal distance revisited. *The Biological Bulletin* 216:373–385.
- Smart, S. M., K. Thompson, R. H. Marrs, M. G. Le Duc, L. C. Maskell, and L. G. Firbank. 2006. Biotic homogenization and changes in species diversity across human-modified ecosystems. *Proceedings of the Royal Society B* 273:2659–2665.
- Smith, C. L. 1973. Small rotenone stations : a tool for studying coral reef fish communities. *American Museum Novitates* 2512:1–21.
- Socioeconomic Data and Applications Center (SEDAC). 2004. Available at <http://sedac.ciesin.columbia.edu/> from SEDAC, Palisades, NY. Accessed 2012.
- Spalding, M. D., and A. M. Grenfell. 1997. New estimates of global and regional coral reef areas. *Coral Reefs* 16:225–230.
- Stehli, F. G., and J. W. Wells. 1971. Diversity and Age Patterns in Hermatypic Corals. *Systematic Biology* 20:115–126.
- Stevenson, C., L. S. Katz, F. Micheli, B. Block, K. W. Heiman, C. Perle, K. Weng, R. Dunbar, and J. Witting. 2006. High apex predator biomass on remote Pacific islands. *Coral Reefs* 26:47–51.
- Storlazzi, C. D., E. K. Brown, M. E. Field, K. Rodgers, and P. L. Jokiel. 2004. A model for wave control on coral breakage and species distribution in the Hawaiian Islands. *Coral Reefs* 24:43–55.
- Syms, C., and G. P. Jones. 2000. Disturbance, habitat structure, and the dynamics of a coral-reef fish community. *Ecology* 81:2714–2729.
- Taylor, E. B. 2004. An analysis of homogenization and differentiation of Canadian freshwater fish faunas with an emphasis on British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* 61:68–79.
- Tittensor, D. P., C. Mora, W. Jetz, H. K. Lotze, D. Ricard, E. V. Berghe, and B. Worm. 2010. Global patterns and predictors of marine biodiversity across taxa. *Nature* 466:1098–1101.
- Tittensor, D., F. Micheli, and M. Nyström. 2007. Human impacts on the species–area relationship in reef fish assemblages. *Ecology Letters* 10:760–772.

- Ulrich, W., and N. J. Gotelli. 2010. Null model analysis of species associations using abundance data. *Ecology* 91:3384–3397.
- Vellend, M. 2001. Do commonly used indices of  $\beta$ -diversity measure species turnover? *Journal of Vegetation Science* 12:545–552.
- Vellend, M. 2004. Parallel effects of land-use history on species diversity and genetic diversity of forest herbs. *Ecology* 85:3043–3055.
- Vellend, M. 2010. Conceptual synthesis in community ecology. *The Quarterly Review of Biology* 85:183–206.
- Vellend, M., K. Verheyen, K. M. Flinn, H. Jacquemyn, A. Kolb, H. van Calster, G. Peterken, B. J. Graae, J. Bellemare, O. Honnay, J. Brunet, M. Wulf, F. Gerhardt, and M. Hermy. 2007. Homogenization of forest plant communities and weakening of species-environment relationships via agricultural land use. *Journal of Ecology* 95:565–573.
- Villegier, S., S. Blanchet, O. Beauchard, T. Oberdorff, and S. Brosse. 2011. Homogenization patterns of the world's freshwater fish faunas. *Proceedings of the National Academy of Sciences of the United States of America* 108:18003–18008.
- Volterra, V. 1926. Variations and fluctuations of the number of individuals in animal species living together. *Journal du Conseil Permanent International pour l'Exploration de la Mer* 3:3-51.
- White, J. W., J. F. Samhuri, A. C. Stier, C. L. Wormald, S. L. Hamilton, and S. A. Sandin. 2010. Synthesizing mechanisms of density dependence in reef fishes: behavior, habitat configuration, and observational scale. *Ecology* 91:1949–1961.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30:279–338.
- Whittaker, R. H. 1972. Evolution and measurement of species diversity. *Taxon* 21:213–251.
- Williams, G. J., J. E. Smith, E. J. Conklin, J. M. Gove, E. Sala, and S. A. Sandin. 2013. Benthic communities at two remote Pacific coral reefs: effects of reef habitat, depth, and wave energy gradients on spatial patterns. *PeerJ* 1:e81.
- Williams, I. D., B. L. Richards, S. A. Sandin, J. K. Baum, R. E. Schroeder, M. O. Nadon, B. J. Zgliczynski, P. Craig, J. L. McIlwain, and R. E. Brainard. 2010. Differences in reef fish assemblages between populated and remote reefs spanning multiple archipelagos across the central and western Pacific. *Journal of Marine Biology* 2011.
- Wilson, M., and B. R. Rosen. 1998. Implications of paucity of corals in the Paleogene of SE Asia: plate tectonics or centre of origin? *Biogeography and Geological Evolution of SE Asia* 165-195.
- Wilson, S. 2001. Multiscale habitat associations of detritivorous blennies (Blenniidae: Salariaiini). *Coral Reefs* 20:245–251.
- Wilson, S. K., A. M. Dolman, A. J. Cheal, M. J. Emslie, M. S. Pratchett, and H. P. A. Sweatman. 2009. Maintenance of fish diversity on disturbed coral reefs. *Coral Reefs* 28:3–14.
- Wilson, S. K., N. A. Graham, M. S. Pratchett, G. P. Jones, and N. V. Polunin. 2006. Multiple disturbances and the global degradation of coral reefs: are reef fishes at risk or resilient? *Global Change Biology* 12:2220–2234.
- Wilson, S. K., S. C. Burgess, A. J. Cheal, M. Emslie, R. Fisher, I. Miller, N. V. C. Polunin, and H. P. A. Sweatman. 2008. Habitat utilization by coral reef fish:

implications for specialists vs. generalists in a changing environment. *Journal of Animal Ecology* 77:220–228.