Predictive Modeling of Coral Disease Distribution within a Reef System

Gareth J. Williams¹*, Greta S. Aeby², Rebecca O. M. Cowie¹, Simon K. Davy¹*

1 School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand, 2 Hawaii Institute of Marine Biology, Kaneohe, Hawaii, United States of America

Abstract

Diseases often display complex and distinct associations with their environment due to differences in etiology, modes of transmission between hosts, and the shifting balance between pathogen virulence and host resistance. Statistical modeling has been underutilized in coral disease research to explore the spatial patterns that result from this triad of interactions. We tested the hypotheses that: 1) coral diseases show distinct associations with multiple environmental factors, 2) incorporating interactions (synergistic collinearities) among environmental variables is important when predicting coral disease spatial patterns, and 3) modeling overall coral disease prevalence (the prevalence of multiple diseases as a single proportion value) will increase predictive error relative to modeling the same diseases independently. Four coral diseases: Porites growth anomalies (PorGA), Porites tissue loss (PorTL), Porites trematodiasis (PorTrem), and Montipora white syndrome (MWS), and their interactions with 17 predictor variables were modeled using boosted regression trees (BRT) within a reef system in Hawaii. Each disease showed distinct associations with the predictors. Environmental predictors showing the strongest overall associations with the coral diseases were both biotic and abiotic. PorGA was optimally predicted by a negative association with turbidity, PorTL and MWS by declines in butterflyfish and juvenile parrotfish abundance respectively, and PorTrem by a modal relationship with Porites host cover. Incorporating interactions among predictor variables contributed to the predictive power of our models, particularly for PorTrem. Combining diseases (using overall disease prevalence as the model response), led to an average six-fold increase in cross-validation predictive deviance over modeling the diseases individually. We therefore recommend coral diseases to be modeled separately, unless known to have etiologies that respond in a similar manner to particular environmental conditions. Predictive statistical modeling can help to increase our understanding of coral disease ecology worldwide.

Citation: Williams GJ, Aeby GS, Cowie ROM, Davy SK (2010) Predictive Modeling of Coral Disease Distribution within a Reef System. PLoS ONE 5(2): e9264. doi:10.1371/journal.pone.0009264

Editor: John F. Bruno, University of North Carolina at Chapel Hill, United States of America

Received August 19, 2009; Accepted January 29, 2010; Published February 17, 2010

Copyright: © 2010 Williams et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Gareth J. Williams is supported by a Victoria University Vice-Chancellor Strategic Research Scholarship and a New Zealand International Doctoral Research Scholarship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: info@garethjwilliams.net (GJW); simon.davy@vuw.ac.nz (SKD)

Introduction

The notion of a complex web of interactions between a disease and its environment has been postulated for centuries [1] and stems from the fact that diseases often have intricate etiologies [2] and different modes of pathogen transmission between hosts [3]. Furthermore, pathogen virulence can respond positively or negatively to a range of variables, such as temperature, nutrient availability, or habitat quality [4–6]; changes in environmental conditions can promote physiological stress that impairs host immunity [7–9], and there may be differences in disease susceptibility between host genotypes [10,11]. With this in mind, it is easy to envisage how complex associations between a disease, the host, and the environment can become established. For example, cholera in humans is caused by Vibrio cholerae, a bacterium that attaches to zooplankton[12]. Outbreaks of cholera are not the result of changes in a single environmental factor, but instead involve multiple interactions between human host densities, V. cholerae, water temperature, salinity, and copepod abundance, and are generally a result of zooplankton blooms following heavy rainfall [13].

Marine organisms are also subject to the influence of disease. Coral reefs worldwide are in decline [14–16] and the role of marine diseases, in particular coral disease, to this decline is receiving increasing attention [5,17–20]. Coral disease outbreaks can lead to an overall reduction in live coral cover [21] and reduced colony density [22], and in extreme cases initiate community phase-shifts from coral- to algal-dominated communities [23]. Coral diseases can also result in a restructuring of coral populations [24], for example a shift from long-lived slow growing massive reef builders to communities dominated by smaller, shorter-lived corals [25]. As corals act as facilitators for other reef invertebrates [26] and fish [27] their loss threatens coral reef biodiversity and function. Spatial patterns of coral disease are linked to environmental conditions [28]. Significant relationships exist between coral disease prevalence and elevated water temperatures [29–32], a decline in water quality [33–37], vector and host densities [31,38], and intensity of coral bleaching [39,40]. The effects of environmental factors on coral disease prevalence and modes of transmission are likely to be intricate and synergistic [41]. Recently, efforts have shifted towards this multi-factorial concept. For example, scleractinian coral white syndrome outbreaks along the Great Barrier Reef (GBR) require a threshold coral cover of greater than 50% in conjunction with thermal stress events, and the interaction between the two predictors explains a significant amount of the increase in the frequency of outbreaks

[31]. In Kenya, the relationship between massive *Porites* growth anomaly prevalence and 16 environmental parameters including water quality, temperature, intensity of bleaching, and benthic composition were modeled to reveal bleaching intensity as the most important factor in explaining spatial distribution of the disease [40]. However, researchers and monitoring programs are still, on occasion, attempting to understand spatial patterns of overall coral disease prevalence (combining the prevalence of multiple diseases into a single proportion value as the response variable) with the environment. This approach ignores the common-sense notion that diseases with different pathogens and hosts are unlikely to have common spatial/temporal patterns or environmental associations, and therefore should be monitored and analyzed individually unless known to have a similar cause.

Exploring coral disease spatial patterns requires a statistical technique that effectively addresses the complexity of disease ecology, in particular the potential for non-linear relationships between the disease, host and environment [42]. One approach is classification and regression tree (CART) modeling [43]. Regression trees have several advantages as a modeling technique, including that various types of predictor and response variables can be analyzed simultaneously rather than in an iterative manner, missing values in data sets can be incorporated and therefore information loss minimized, and complex interactions between predictors can be quantified and modeled in a simple manner [44]. Despite these advantages regression trees are often poor predictors and large trees can be difficult to interpret [44]. Recently these weaknesses have been overcome with the use of boosted regression trees (BRT) [44–49], which incorporate machine learning decision tree methods [50] and boosting, a method for improving model accuracy (reducing predictive error) [46]. BRT differs fundamentally from conventional techniques that aim to fit a single parsimonious model. Instead, the final BRT model is an additive regression model in which individual terms are simple trees, fitted in a forward stage-wise manner [46]. In summary, BRT gives two crucial pieces of information, namely the underlying relationship between the response and each predictor, and the strongest statistical predictor (among the simultaneously tested predictors) of the response in question. These are clearly two different things, and as BRT focuses on building predictive models for theory development, the environmental associations that result can be direct or indirect. Disease-environment relationships revealed by this type of modeling can be used to predict spatial patterns in novel systems and facilitate hypothesis-driven experimental studies. Exploratory and predictive modeling provides an initial step towards understanding spatial patterns and processes and has been underutilized in coral disease research.

In the present study, we used a BRT technique and a reef system with contrasting environmental conditions and a range in coral disease states and prevalence to address the following hypotheses: 1) coral diseases show distinct associations with multiple environmental factors, 2) incorporating interactions (synergistic collinearities) among environmental variables is important when predicting coral disease spatial patterns, and 3) modeling overall coral disease prevalence (the prevalence of multiple diseases as a single proportion value) will increase predictive error relative to modeling the same diseases independently. In addition, to develop the use of BRT modeling in coral disease research we outline the analytical methods of a technique and its novel application.

Materials and Methods

Model System, Host Sensities and Disease Prevalence

In August 2007, pilot surveys were conducted within Coconut Island Marine Reserve (CIMR) (21°26'N, 157°47'W), Kaneohe

Bay, Oahu, Hawaii. The two competitively dominant space holders in the system were *Porites compressa* and *Montipora capitata*. Pocillopora damicornis, P. meandrina and Fungia scutaria were also observed but at low densities. Four disease states affecting Porites and Montipora spp. were documented and CIMR was found to represent an ideal system for modeling coral disease-environment associations due to large variations in host abundance, disease prevalence, and environmental conditions over spatial scales of 100s m (Fig. 1, Table 1).

We conducted surveys over two five-week periods: October – November 2007 (winter), and May – July 2008 (summer). The sampling design was not hierarchical, but instead was designed to maximize variability between observations in both disease prevalence and the environmental predictors. Observations were randomized within 11 specific regions of CIMR known, from preliminary surveys, to display contrasting disease prevalence and environmental conditions. To quantify disease prevalence, 55 belt transects (10×2 m) were surveyed within a depth range of 0.7– 3.1 m in each season (giving 110 observations overall). Lesions on colonies were classified according to gross morphology (growth anomalies, tissue loss, discoloration, trematodiasis) and assigned the host genus and descriptive name [51] (Fig. 1). Porites trematodiasis (PorTrem) was recorded even if a single lesion was found on a colony. The proportion of diseased colonies was calculated for each disease and the overall (total) disease prevalence. To quantify host abundance, every coral colony whose centre fell within the belt transect area was counted and inspected for signs of disease. Percentage cover of live coral was estimated using a point-intercept method at 50-cm increments along the transect line.

Environmental and Biological Predictors

Salinity, turbidity and chlorophyll-a were measured using two $RBR^@$ XR-420 data loggers (www.rbr-global.com) recording every minute over 24 hr periods at the depth of the coral. The chlorophyll-a value is a measure of how much of the suspended material present (turbidity) contains chlorophyll-a. The placement of the loggers was randomized among the 11 CIMR regions throughout each 5-week period. HOBO® Pro temperature data loggers (www.onsetcomp.com) were attached to the reef within each of the 11 regions; these recorded every 10 min from the start of October 2007 to the end of July 2008.

Sedimentation levels were measured as a potential source of stress to the corals which in turn may influence their susceptibility to disease. Within each of the 11 regions, sedimentation was quantified using PVC sediment traps [52]. These were attached to stainless steel poles and placed into, and approximately 30 cm above, the substrate among coral colonies. Sedimentation was measured over 7-day periods, with measurements repeated 5 times per season. To determine the organic carbon fraction of the sediment (a proxy for the level of organics, but not dissolved organics, entering the system), sediment was finely ground, oven dried at 100° C for 10 h and weighed. Samples were then placed in a muffle furnace at 500° C for 12 h to burn off the organic fraction and the remaining inorganic fraction reweighed [53].

Physical injury to the host coral can promote the spread of some coral diseases [54]. Reef fish, such as butterflyfishes, parrotfishes and damselfish, offer a potential source of injury to corals [55–57] and fish are known to be involved in disease transmission [58] and/or promoting the rate of disease spread [59]. Within CIMR, pilot surveys showed butterflyfish to be the dominant coral-feeding fish. Damselfish and adult parrotfish were seldom seen but juvenile parrotfish were abundant and parrotfish feeding scars were seen around CIMR, particularly on P.

Gross description: Focal to multifocal cicular/oblong lesion with distinct edges, smooth margins and an umbonate relief. Generally has a peripheral and medial location on the colony. Lesion tissue (arrow) has a less dense skeletal structure and increased calyx size in comparison to healthy tissue.

Minimum prevalence = 1.06% Maximum prevalence = 56.16% Etiology: Unknown

Gross description: Diffuse irregular tissue loss with indistinct edges and undulating margins. Generally has a central and apical location on colony. Displays tissue thinning and eventual tissue loss with subsequent algal colonization. Patches usually bordered by narrow bleached or mucous band.

Minimum prevalence = 0.0% Maximum prevalence = 35.95% Etiology: Unknown

Gross description: Multifocal circular pink/white swollen (3-5 mm) with distinct edges, smooth margins and umbonate relief. Located on any part of the colony and can be clustered.

Minimum prevalence = 0.0% Maximum prevalence = 73.97% Etiology: Lesions are a result of infection by the larval stage of the digenetic trematode Podocotyloides stenometra.

Gross description: Diffuse irregular lesion with indistinct edges and smooth to undulating margins. Located on any part of the colony. Tissue loss can be gradual or occur in sudden bursts leaving behind intact white skeleton that is colonized by algae. Border between healthy and diseased tissue usually has a band of mucous, bleached tissue, or thin (one polyp deep) layer of white necrotic tissue.

Minimum prevalence = 0.0% Maximum prevalence = 29.17% Etiology: Unknown

Figure 1. Gross descriptions of the four coral diseases present at Coconut Island Marine Reserve, Oahu, Hawaii. a) Porites growth anomaly, b) Porites tissue loss, c) Porites trematodiasis, and d) Montipora white syndrome. Minimum and maximum prevalence values between transects are shown.

doi:10.1371/journal.pone.0009264.g001

compressa. Hence, only coral-feeding butterflyfish (facultative and obligate) and juvenile parrotfish were quantified over a 50×6 m area within the vicinity of each disease transect. The observer swam at a speed of approximately 10 m min^{-1} to account for the active nature of these reef fish and 1 m out from the reef-flat edge to detect fish both on the reef flat and slope. Horizontal visibility

limited the width of the transect, with 3 m being the limit at which fish could confidently be identified to species. Butterflyfish species observed were Chaetodon auriga, C. ephippium, C. lineolatus, C. lunula, C. lunulatus (formally C. trifasciatus), C. multicinctus, C. ornatissimus, and C. unimaculatus. Each count was conducted between the daylight hours of 10:00 and 15:00 and replicated a Table 1. Predictor variables used in the analyses with their codes and units.

Min/Max, minimum and maximum predictor values between transects.

doi:10.1371/journal.pone.0009264.t001

minimum of five times, with each count taking place on a different day.

Statistical Analysis

The 110 belt transects (55 from each season) were modeled simultaneously against 17 predictor variables, which included continuous environmental data, count data, and categorical terms (Table 1). Transects were considered independent observations in the models, and not pseudoreplicates, as they were separate from each other in both space and time. We used Boosted Regression Trees (BRT) [46] as the modeling technique. The technique is sometimes referred to as stochastic gradient boosting, as BRT includes an element of stochasticity in order to improve accuracy and reduce overfitting (when a statistical model describes random error or noise instead of the underlying relationship) [60]. BRTs were constructed using the routines *gbm* version 1.5–7 [61] and gbm.step [46] in the R statistical program version 2.6.2 (R Development Core Team, http://www.r-project.org). Prevalence data was log transformed and the few zero disease prevalence counts that did occur removed in order to achieve a normal/ pseudo-normal distribution. The numbers of independent observations were then as follows: *Montipora* white syndrome $(n = 101)$, Porites trematodiasis (n = 86), Porites tissue loss (n = 101), Porites growth anomalies $(n = 110)$, and overall disease $(n = 110)$. Analyses were based on a Gaussian distribution. Due to problems with assigning real probabilities in BRTs (there are no p -values) a key approach is to use validation processes that require a proportion of the data set to be held back. Due to our relatively small data set, we used 10-fold cross-validation (cv) for model development and validation, with the benefit of still using the full data set to fit the final model. The measure of model performance was cv deviance and standard error (se) throughout our study (the lower the value the better the model performance). Within the BRT model, three terms are used to optimize predictive performance: bag-fraction, learning rate, and tree complexity.

The bag-fraction determines the proportion of data to be selected at each step and therefore the model stochasticity; for example a bag fraction of 0.5 means that 50% of the data are drawn at random without replacement. The learning rate (lr) is used to shrink the contribution of each tree as it is added to the model, and tree complexity (tc) determines the number of nodes in a tree and should reflect the true interaction order on the response being modeled [62]. We determined optimal settings for these parameters by examining the cv deviance over tc values 1–5, lr values of 0.05, 0.01 and 0.001, and bag fractions of 0.5 and 0.75. All possible combinations were run, with the optimal number of trees in each case being determined by gbm.step. The combination of the three parameter settings with the lowest cv deviance was then selected to produce the final BRT. Finally, redundant predictor variables may degrade model accuracy by increasing variance, particularly in small data sets. We therefore achieved optimization to create a balance between statistical performance, parsimony, and usefulness of the model using the routine gbm.simplify, a method analogous to backwards selection in regression [46]. Both season and reef type (categorical predictors) were found to exert no influence upon predicting the prevalence of any disease and were removed during optimization before the creation of the final BRTs.

As part of the final model, BRT assesses the relative importance (or contribution) of each predictor. This measure is based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees [45,46]. A higher relative importance of a predictor indicates a stronger influence on the response (disease) in question. Partial dependency plots were used for interpretation and to quantify the relationship between each predictor variable and the disease, after accounting for the average effect of all other predictor variables in the model. To quantify interaction effects between predictors (the collinearity and synergistic effect upon predicting the response in question) we used the routine gbm.interactions [46]. The relative strength of interaction fitted by BRT is quantified by the residual variance from a linear model, and the value indicates the relative degree of departure from a purely additive effect, with zero indicating no interaction effects fitted. The interaction value can also be thought of as the relative contribution of the interaction between the two predictors towards the overall predictive performance of the individual model (the cv deviance value). We defined a threshold interaction value and reported the interactions with values ≥ 0.1 . We performed the above analyses for individual diseases and for the calculation of overall disease prevalence.

Results

Environmental Associations and Strongest Predictors

Porites growth anomalies (PorGA). Two relationships contributed most strongly to predicting PorGA prevalence (Fig.2), namely negative relationships with both turbidity and depth. PorGA prevalence was highest in clearer waters within 1 m of the surface. In addition, prevalence peaked when there was an overall coral cover of 40–70%, increased chlorophyll-a concentration within any suspended material, lower juvenile parrotfish abundance, and an increase in Porites cover. Turbidity offered the largest contribution (i.e. it was the strongest predictor) with a relative importance of 33.9% (Fig.2). Model cv deviance equalled 0.391, with second order interactions present between predictors (Table 2). The largest interaction (collinearity and synergistic effect) involved Porites cover and total coral cover (Table 3).

Porites tissue loss (PorTL). Three relationships contributed most strongly to predicting PorTL prevalence (Fig.2): a negative correlation with butterflyfish abundance, a positive correlation with temperature, and a negative correlation with turbidity. Prevalence peaked in areas with few butterflyfish, where temperatures reached above 27° C, and low turbidity environments. Butterflyfish abundance was the strongest predictor with a relative importance of 47.5% (Fig.2). Model cv deviance equalled 0.350, with second order interactions present between predictors (Table 2). The largest interaction involved butterflyfish (the strongest predictor) and turbidity (Table 3).

Porites trematodiasis (PorTrem). Four relationships contributed most strongly to predicting PorTrem prevalence (Fig.2). A modal relationship occurred with Porites cover, with a peak in prevalence at approximately 50% cover, and a positive correlation existed with overall colony density, reaching an asymptote at approximately 9 colonies m^{2-1} . There was a positive correlation with butterflyfish abundance (peaking above 4 fish 300 m²⁻¹), and a weak negative correlation with depth (Fig.2). Porites cover was the strongest predictor with a relative importance of 31.2% (Fig.2). Model cv deviance equalled 1.182, the highest deviance for any of the individual models, with third order interactions present between predictors (Table 2). The largest interaction involved *Porites* cover and overall colony density (the two strongest predictors). This was the largest interaction value (2.02) seen within any of the models (Table 3).

Montipora white syndrome (MWS). Two relationships contributed most strongly to predicting MWS prevalence, namely a negative correlation with juvenile parrotfish abundance and a positive correlation with chlorophyll-a concentration (Fig.2). In addition, a positive correlation existed with temperature, with peak prevalence occurring above 27°C. Juvenile parrotfish abundance was the strongest predictor with a relative importance of 53.6% (Fig.2). Model cv deviance equalled 0.213, the lowest deviance (best fit) for any of the models, with second order interactions present between predictors (Table 2). The single

interaction involved chlorophyll-a with temperature (Table 3). This was the only model where the strongest predictor (juvenile parrotfish abundance) did not interact with another predictor variable above the defined interaction threshold.

Combining disease states (overall prevalence). Combined modeling of the diseases led to a loss in predictive performance (increased predictive error) of the model, with an approximate six-fold increase in cv deviance above the average cv deviance for all four diseases analysed individually (Table 2).

Discussion

Coral diseases, like most diseases, can logically be expected to display complex associations with their environment due to the intricate nature of the host -environment-pathogen triad of disease causation [2], and the inherent multi-collinearity present between biotic and abiotic variables in any ecological system. With the use of boosted regression tree (BRT) modeling, this study has shown that different coral diseases do indeed show complex associations with a range of environmental variables and that these associations are distinct between diseases. We determined the environmental associations, and of these, the strongest statistical predictors of four individual coral diseases within a reef system in Hawaii from a suite of 17 predictor variables. The environmental conditions showing the strongest overall associations (direct or indirect) with coral disease prevalence in our model system were fish abundance, host availability, temperature, water quality (turbidity and chlorophyll-a concentration), and depth.

Biotic, Abiotic and Physical Associations with Disease

Within our study the relative importance of disease associations (direct or indirect) with biotic, abiotic and physical parameters differed across coral disease states. Porites growth anomalies (PorGA) were optimally predicted by turbidity (abiotic), Porites tissue loss (PorTL) and Montipora white syndrome (MWS) by a decline in reef fish abundances (biotic), whilst spatial patterns of Porites trematodiasis (PorTrem) were optimally predicted by Porites host cover (biotic). The ecological mechanisms behind these disease-environment patterns are likely to be complex. Reef fish could be involved in regulating the disease dynamics of PorTL and MWS directly by offering a mechanism for diseased tissue removal via predation that could lead to individual host recovery [63]. Conversely, the association could equally be indirect and overall conditions which have negative effects on butterflyfish and juvenile parrotfish abundance may favor PorTL and MWS prevalence. In the Philippines, negative relationships between coral disease prevalence and fish taxonomic diversity exist inside and outside of marine protected areas [64], and in Palau increased prevalence of skeletal eroding band disease is associated with a reduction in the richness of a fish species targeted by fishers [65]. Further research is needed to tease apart the direct and indirect associations between coral disease prevalence and reef fish.

In addition to reef fish, we found strong links between the spatial patterns of PorTrem and a further biotic predictor, namely host abundance. The relationship between disease prevalence and host abundance is central to the theory of disease ecology [3]. As transmission is a key process in host-pathogen interactions, increased host density can increase horizontal transmission of a disease [66]. Hence, to a degree, host availability can determine how many infected individuals are observed in a defined area [67], regulated by both density dependent [68–70] and frequency dependent processes [71–73]. Scleractinian coral white syndrome outbreaks along the Great Barrier Reef require, in part, an overall

Figure 2. Boosted regression tree (BRT) analyses relating prevalence of four coral diseases to environment. Models are developed and validated using 10-fold cross-validation on 86–110 independent observations for each disease and 17 predictor variables. The 8 most influential predictors to the model are shown. Their relative importance is shown as a % in parentheses. The deciles of the distribution of the predictors are indicated by tick marks along the top of each plot. Predictor variable codes and units are as per Table 1. doi:10.1371/journal.pone.0009264.g002

Table 2. Optimal settings and predictive performance of boosted regression tree (BRT) analyses relating prevalence of four coral diseases to environment.

lr, learning rate; tc, tree complexity. Cross-validation (cv) deviance and standard error (se) is shown as the measure of model performance (the lower the value the better the model performance).

doi:10.1371/journal.pone.0009264.t002

coral cover in excess of 50% [31]. Positive correlations between coral cover and prevalence of black band, yellow band and white band disease were reported at reefs in Dubai [74], and positive relationships between overall disease and Porites cover in the Philippines [41]. PorTrem is caused by a digenetic trematode that relies on trophic transmission for completion of its multi-host (fish, mollusc, coral) life cycle [75]. Infected coral polyps are fed upon by coral-feeding fish, such as butterflyfish, which then become infected with the adult worm. Transmission of PorTrem across the reef occurs when the parasite's eggs are shed with the fish host feces. It therefore follows that as Porites cover and colony densities increase the chance of infected feces landing on the Porites host also increases, hence the positive relationship. The reason for the drop at higher levels of Porites cover is unclear and has been found across the entire Kaneohe Bay area [38]. Additionally, PorTrem is unable to establish without the full compliment of intermediate hosts and therefore the positive relationship with butterflyfish

abundance is not surprising. Increased butterflyfish abundance leads to more infected polyps being fed upon and in turn more infected feces dropping over the reef.

Disease spatial patterns are often predicted by abiotic as well as biotic parameters. Among our four coral diseases we found PorGA and PorTL were both associated with reduced water turbidity, PorTL was positively associated with temperature, and MWS was positively associated with water chlorophyll-a concentration. For PorGA, water turbidity and depth (the sole physical parameter) were superior to any of the biotic parameters in predicting the prevalence of the disease, with prevalence peaking in shallow, less turbid waters. Little is known about PorGA ecology but it has been speculated that growth anomaly formation in corals could be linked to increased sensitivity to ultraviolet radiation (UVR) [76,77]. Improved water clarity and shallow depths (with subsequent low light attenuation) could lead to increased levels of UVR reaching the coral surfaces [78]. In addition, shallower

Table 3. Pairwise interactions between predictor variables used to relate prevalence of four coral diseases to environment.

Interactions displayed are those that exceeded an interaction value of ≥ 0.1 and involved the 8 predictors offering the highest contribution to the model displayed in Figure 2. Interaction value indicates the relative degree of departure from a purely additive effect, with a value of zero indicating that no interaction is present. A summary description is given for the association of the peak in disease prevalence and the pairwise interactions for those predictor variables showing a clear relationship (for example positive, negative, or modal) with the disease in Figure 2. doi:10.1371/journal.pone.0009264.t003

depths are often associated with higher variations in environmental stressors such as temperature and salinity and these fluctuations may be affecting PorGA prevalence.

A positive association between disease prevalence and temperature, as seen with PorTL and to a lesser degree with MWS in our study, is common in disease ecology. Increased temperature, like any environmental stressor, can shift the balance in favour of either the host or pathogen [10]. Compromized hosts may be more susceptible to ubiquitous pathogens and/or the stressor may increase pathogen virulence [5,7,79]. For example, malaria prevalence is often associated with temperature. At higher temperatures the parasite development time inside the mosquito vector shortens and so mosquitoes become infectious sooner and transmission rates increase [80]. For corals, increased temperatures can lead to loss of the symbiotic algae (bleaching) and possible mortality [81]. Higher water temperatures can also promote bacterial growth [5]. For bacterial diseases, the combined effect of temperature stress on the coral host and enhanced bacterial growth may ultimately result in disease occurrence. This was recently found in the Virgin Islands where coral bleaching led to a lethal white plague disease outbreak [82]. Many coral diseases show positive associations with temperature, such as black band disease in the Caribbean [83,84], the Florida Keys [85], the GBR [86], and Venezuela [32]; white plague in Puerto Rico [87]; atramentous necrosis in Australia [29], and white syndrome along the GBR [31]. Of these diseases, three have been identified as caused by a bacterial pathogen resulting in chronic or acute tissue loss: white plague Type II [88], black band disease [89], and white syndrome [90]. The emergent pattern suggests coral diseases that produce progressive tissue loss are responding to seawater temperature whereas those displaying disease signs other than tissue loss are not, or at least not in the same manner. Similarly, we found that the two diseases within CIMR that displayed a positive association with temperature were PorTL and MWS (both tissue loss diseases). Importantly, as only two of our four diseases showed distinct associations with temperature we emphasize that temperature should not be assumed to universally display specific relationships with coral disease prevalence.

A further environmental stressor for coral is reduced water quality, as measured by increased nutrients and chlorophyll-a concentration. Reduced water quality has been linked to increased severity and prevalence of aspergillosis in sea fans [33,35,37], increased prevalence of yellow band disease [35], and increased black band disease prevalence and progression, respectively [34,91]. In our study the only diseases to show strong positive associations with increasing chlorophyll-a concentration were MWS and, to a lesser degree of predictive importance, PorGA. Consistent with this, MWS prevalence across Kaneohe Bay, an area with historically poor water quality, was found to be four times higher than in other areas of the Main Hawaiian Islands [92].

Predictor Interactions and Combining Diseases

Researchers often view collinearity between predictor variables as a problem in ecological modeling and remove predictor variables that are highly collineated prior to model formation. However, providing the collinearity between predictors can be

References

identified, quantified and built into the model their synergistic effect may improve model predictive capability. Incorporating interactions between predictor variables increased the predictive power of our models, particularly for Porites trematodiasis (PorTrem). When predicting the prevalence of PorTrem, Porites cover and overall colony density (the two strongest predictors) were also the two variables showing the highest interaction level (highest degree of departure from a purely additive effect) and together explained the largest amount of variation in the disease occurrence. The number and higher values of the interactions present for PorTrem probably reflects the complex multi-host relations required for this disease to occur. Significant interaction terms between predictors have also been reported for scleractinian coral white syndrome outbreaks in Australia [31] and researchers have started to adopt a more multi-factorial approach to understanding coral disease-environment associations [31,40]. Our results, in conjunction with these studies, emphasize the need for considering multiple environmental predictors and their respective collinearity for coral disease-environment modeling.

Modeling combined diseases (the prevalence of multiple diseases as a single proportion value as the model response), led to an average six-fold increase in cross-validation deviance (reduction in predictive accuracy) over modeling the diseases individually. This is to be expected. For example, environmental modeling of human cholera (caused by an intestinal bacterium), and measles (a viral disease), even though they affect the same host, would most likely produce confusing results due to their differing etiologies and modes of transmission [13,71]. However, when disease etiologies are known and their ecological similarities recognized then combined disease modelling may be appropriate. For example, dengue fever and dengue haemorrhagic fever, two diseases both transmitted by mosquitoes within the genus Aedes, were modeled together successfully within Thailand [93]. Importantly, the authors were not modeling a combined proportion value of both diseases as the response variable. Modeling overall coral disease prevalence, multiple diseases each with a possibly distinct etiology, seems inappropriate. We recommend coral diseases to be modeled individually, unless they are known to have etiologies that respond in a uniform manner to particular environmental conditions. Predictive statistical modeling forms an important stage in the understanding of coral disease patterns and in conjunction with biomedical techniques, field observations and laboratory manipulations, can increase our understanding of coral disease ecology worldwide.

Acknowledgments

We thank the Hawaiian Institute of Marine Biology for use of their facilities and permission to work within Coconut Island Marine Reserve. We thank Barrett Wolfe, Chris Runyon, Ann Farrell, and Ku'ulei Rodgers for field and laboratory assistance. We also thank Kevin Lafferty, Thierry Work, Joanne Davy and Ingrid Knapp for stimulating discussions and feedback on early drafts of the manuscript.

Author Contributions

Conceived and designed the experiments: GJW GSA. Performed the experiments: GJW ROMC. Analyzed the data: GJW. Wrote the paper: GJW GSA SKD.

1. Ostfeld RS, Keesing F, Eviner VT (2008) Infectious disease ecology: The effects of ecosystems on disease and of disease on ecosystems. New Jersey, USA: Princeton University Press. 506 p.

^{2.} Work TM, Richardson LL, Reynolds TL, Willis BL (2008) Biomedical and veterinary science can increase our understanding of coral disease. Journal of Experimental Marine Biology and Ecology 362: 63–70.

- 3. Lloyd-Smith JO, Cross PC, Briggs CJ, Daugherty M, Getz WM, et al. (2005) Should we expect population thresholds for wildlife disease? Trends in Ecology & Evolution 20: 511–519.
- 4. Lafferty KD, Kuris AM (1999) How environmental stress affects the impacts of parasites. Limnology and Oceanography 44: 925–931.
- 5. Harvell CD, Mitchell CE, Ward JR, Altizer S, Dobson AP, et al. (2002) Ecology - Climate warming and disease risks for terrestrial and marine biota. Science 296: 2158–2162.
- 6. Harvell D, Jordan-Dahlgren E, Merkel S, Rosenberg E, Raymundo L, et al. (2007) Coral disease, environmental drivers, and the balance between coral microbial associates. Oceanography 20: 172–195.
- 7. Fitt WK, Brown BE, Warner ME, Dunne RP (2001) Coral bleaching: interpretation of thermal tolerance limits and thermal thresholds in tropical corals. Coral Reefs 20: 51–65.
- 8. Jackson JA, Tinsley RC (2002) Effects of environmental temperature on the susceptibility of Xenopus laevis and X-wittei (Anura) to Protopolystoma xenopodis (Monogenea). Parasitology Research 88: 632–638.
- 9. Lafferty KD, Holt RD (2003) How should environmental stress affect the population dynamics of disease? Ecology Letters 6: 654–664.
- 10. Blanford S, Thomas MB, Pugh C, Pell JK (2003) Temperature checks the Red Queen? Resistance and virulence in a fluctuating environment. Ecology Letters 6: 2–5.
- 11. Vollmer SV, Kline DI (2008) Natural Disease Resistance in Threatened Staghorn Corals. PLoS ONE 3: e3718.
- 12. Kirn TJ, Jude BA, Taylor RK (2005) A colonization factor links Vibrio cholerae environmental survival and human infection. Nature 438: 863–866.
- 13. Colwell RR (2004) Infectious disease and environment: cholera as a paradigm for waterborne disease. International Microbiology 7: 285–289.
- 14. Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, et al. (2003) Climate change, human impacts, and the resilience of coral reefs. Science 301: 929–933.
- 15. Bellwood DR, Hughes TP, Folke C, Nystrom M (2004) Confronting the coral reef crisis. Nature 429: 827–833.
- 16. Bruno JF, Selig ER (2007) Regional Decline of Coral Cover in the Indo-Pacific: Timing, Extent, and Subregional Comparisons. PLoS ONE 2: e711.
- 17. Harvell CD, Kim K, Burkholder JM, Colwell RR, Epstein PR, et al. (1999) Review: Marine ecology - Emerging marine diseases - Climate links and anthropogenic factors. Science 285: 1505–1510.
- 18. Sutherland KP, Porter JW, Torres C (2004) Disease and immunity in Caribbean and Indo-Pacific zooxanthellate corals. Marine Ecology-Progress Series 266: 273–302.
- 19. Ward JR, Lafferty KD (2004) The elusive baseline of marine disease: Are diseases in ocean ecosystems increasing? Plos Biology 2: 542–547.
- 20. Weil E, Smith G, Gil-Agudelo DL (2006) Status and progress in coral reef disease research. Diseases of Aquatic Organisms 69: 1–7.
- 21. Nugues MM (2002) Impact of a coral disease outbreak on coral communities in St. Lucia: What and how much has been lost? Marine Ecology-Progress Series 229: 61–71.
- 22. Richardson LL, Voss JD (2005) Changes in a coral population on reefs of the northern Florida Keys following a coral disease epizootic. Marine Ecology-Progress Series 297: 147–156.
- 23. Aronson RB, Precht WF (2001) White-band disease and the changing face of Caribbean coral reefs. Hydrobiologia 460: 25–38.
- 24. Borger JL (2005) Scleractinian coral diseases in south Florida: incidence, species susceptibility, and mortality. Diseases of Aquatic Organisms 67: 249–258.
- 25. Bruckner AW, Bruckner RJ (2006) Consequences of yellow band disease (YBD) on *Montastraea annularis* (species complex) populations on remote reefs off Mona
Island, Puerto Rico. Diseases of Aquatic Organisms 69: 67–73.
- 26. Idjadi JA, Edmunds PJ (2006) Scleractinian corals as facilitators for other invertebrates on a Caribbean reef. Marine Ecology-Progress Series 319: 117–127.
- 27. Jones GP, McCormick MI, Srinivasan M, Eagle JV (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.
- 28. Kaczmarsky LT (2006) Coral disease dynamics in the central Philippines. Diseases of Aquatic Organisms 69: 9–21.
- 29. Jones RJ, Bowyer J, Hoegh-Guldberg O, Blackall LL (2004) Dynamics of a temperature-related coral disease outbreak. Marine Ecology-Progress Series 281: 63–77.
- 30. Bally M, Garrabou J (2007) Thermodependent bacterial pathogens and mass mortalities in temperate benthic communities: a new case of emerging disease linked to climate change. Global Change Biology 13: 2078–2088.
- 31. Bruno JF, Selig ER, Casey KS, Page CA, Willis BL, et al. (2007) Thermal stress and coral cover as drivers of coral disease outbreaks. Plos Biology 5: 1220–1227.
- 32. Rodriguez S, Croquer A (2008) Dynamics of Black Band Disease in a Diploria strigosa population subjected to annual upwelling on the northeastern coast of Venezuela. Coral Reefs 27: 381–388.
- 33. Kim K, Harvell CD (2002) Aspergillosis of sea fan corals: dynamics in the Florida Keys. In: Porter JW, Porter KG, eds. The Everglades, Florida Bay, and coral reefs of the Florida Keys: an ecosystem sourcebook. Boca Raton: CRC. pp 813–824.
- 34. Kuta KG, Richardson LL (2002) Ecological aspects of black band disease of corals: relationships between disease incidence and environmental factors. Coral Reefs 21: 393–398.
- 35. Bruno JF, Petes LE, Harvell CD, Hettinger A (2003) Nutrient enrichment can increase the severity of coral diseases. Ecology Letters 6: 1056–1061.
- 36. Voss JD, Richardson LL (2006) Nutrient enrichment enhances black band disease progression in corals. Coral Reefs 25: 569–576.
- 37. Baker DM, MacAvoy SE, Kim K (2007) Relationship between water quality, delta N-15, and aspergillosis of Caribbean sea fan corals. Marine Ecology-Progress Series 343: 123–130.
- 38. Aeby GS (2007) Spatial and temporal patterns of Porites trematodiasis on the reefs of Kaneohe Bay, Oahu, Hawaii. Bulletin of Marine Science 80: 209–218.
- 39. Muller EM, Rogers CS, Spitzack AS, van Woesik R (2008) Bleaching increases likelihood of disease on Acropora palmata (Lamarck) in Hawksnest Bay, St John, US Virgin Islands. Coral Reefs 27: 191–195.
- 40. McClanahan TR, Weil E, Maina J (2009) Strong relationship between coral bleaching and growth anomalies in massive *Porites*. Global Change Biology 15: 1804–1816.
- 41. Raymundo LJ, Rosell KB, Reboton CT, Kaczmarsky L (2005) Coral diseases on Philippine reefs: genus *Porites* is a dominant host. Diseases of Aquatic Organisms 64: 181–191.
- 42. Sokolow SH, Foley P, Foley JE, Hastings A, Richardson LL (2009) Disease dynamics in marine metapopulations: modelling infectious diseases on coral reefs. Journal of Applied Ecology 46: 621–631.
- 43. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and Regression Trees. California: Wadsworth International Group.
- 44. De'ath G (2007) Boosted trees for ecological modeling and prediction. Ecology 88: 243–251.
- 45. Friedman JH, Meulman JJ (2003) Multiple additive regression trees with application in epidemiology. Statistics in Medicine 22: 1365–1381.
- 46. Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. Journal of Animal Ecology 77: 802–813.
- 47. Ge Y, He JZ, Zhu YG, Zhang JB, Xu ZH, et al. (2008) Differences in soil bacterial diversity: driven by contemporary disturbances or historical contingencies? Isme Journal 2: 254–264.
- 48. Leathwick JR, Elith J, Chadderton WL, Rowe D, Hastie T (2008) Dispersal, disturbance and the contrasting biogeographies of New Zealand's diadromous and non-diadromous fish species. Journal of Biogeography 35: 1481–1497.
- 49. Carslaw DC, Taylor PJ (2009) Analysis of air pollution data at a mixed source location using boosted regression trees. Atmospheric Environment 43: 3563–3570.
- 50. Breiman L (2001) Statistical modeling: The two cultures. Statistical Science 16: 199–215.
- 51. Work TM, Aeby GS (2006) Systematically describing gross lesions in corals. Diseases of Aquatic Organisms 70: 155–160.
- 52. Bloesch J (1994) A review of methods used to measure sediment resuspension. Hydrobiologia 284: 13–18.
- 53. Craft CB, Seneca ED, Broome SW (1991) Loss on ignition and kjeldahl digestion for estimating organic-carbon and total nitrogen in estuarine marsh soils - calibration with dry combustion. Estuaries 14: 175–179.
- 54. Page CA, Willis BL (2008) Epidemiology of skeletal eroding band on the Great Barrier Reef and the role of injury in the initiation of this widespread coral disease. Coral Reefs 27: 257–272.
- 55. Cox EF (1986) The effects of a selective corallivore on growth-rates and competition for space between 2 species of Hawaiian corals. Journal of Experimental Marine Biology and Ecology 101: 161–174.
- 56. Bruckner AW, Bruckner RJ (1998) Destruction of coral by Sparisoma viride. Coral Reefs 17: 350–350.
- 57. Miller MW, Hay ME (1998) Effects of fish predation and seaweed competition on the survival and growth of corals. Oecologia 113: 231–238.
- 58. Lu YN, Yu QG, Zamzow JP, Wang Y, Losey GS, et al. (2000) Detection of green turtle herpesviral sequence in saddleback wrasse Thalassoma duperrey: A possible mode of transmission of green turtle fibropapilloma. Journal of Aquatic Animal Health 12: 58–63.
- 59. Aeby GS, Santavy DL (2006) Factors affecting susceptibility of the coral Montastraea faveolata to black-band disease. Marine Ecology-Progress Series 318: 103–110.
- 60. Friedman JH (2002) Stochastic gradient boosting. Computational Statistics & Data Analysis 38: 367–378.
- 61. Ridgeway G (2006) Generalized boosted regression models. Documentation on the R package ''gbm'', version. pp 1.5–7.
- 62. Friedman JH (2001) Greedy function approximation: A gradient boosting machine. Annals of Statistics 29: 1189–1232.
- 63. Cole AJ, Pratchett MS, Jones GP, Seng KMC (2009) Coral-feeding fishes slow progression of black-band disease. Coral Reefs 28: 965–965.
- 64. Raymundo LJ, Halford AR, Maypa AP, Kerr AM (2009) Functionally diverse reef-fish communities ameliorate coral disease. Proceedings of the National Academy of Sciences of the United States of America 106: 17067–17070.
- 65. Page CA, Baker DM, Harvell CD, Golbuu Y, Raymundo L, et al. (2009) Influence of marine reserves on coral disease prevalence. Diseases of Aquatic Organisms 87: 135–150.
- 66. Altizer SM, Augustine DJ (1997) Interactions between frequency-dependent and vertical transmission in host-parasite systems. Proceedings of the Royal Society of London Series B-Biological Sciences 264: 807–814.
- 67. McCallum H, Barlow N, Hone J (2001) How should pathogen transmission be modelled? Trends in Ecology & Evolution 16: 295–300.
- 68. Brown CR, Brown MB (2004) Empirical measurement of parasite transmission between groups in a colonial bird. Ecology 85: 1619–1626.
- 69. Ramsey D, Spencer N, Caley P, Efford M, Hansen K, et al. (2002) The effects of reducing population density on contact rates between brushtail possums: implications for transmission of bovine tuberculosis. Journal of Applied Ecology 39: 806–818.
- 70. Berthier K, Langlais M, Auger P, Pontier D (2000) Dynamics of a feline virus with two transmission modes within exponentially growing host populations. Proceedings of the Royal Society of London Series B-Biological Sciences 267: 2049–2056.
- 71. Bjornstad ON, Finkenstadt BF, Grenfell BT (2002) Dynamics of measles epidemics: Estimating scaling of transmission rates using a Time series SIR model. Ecological Monographs 72: 169–184.
- 72. Begon M, Feore SM, Brown K, Chantrey J, Jones T, et al. (1998) Population and transmission dynamics of cowpox in bank voles: testing fundamental assumptions. Ecology Letters 1: 82–86.
- 73. Begon M, Hazel SM, Baxby D, Bown K, Cavanagh R, et al. (1999) Transmission dynamics of a zoonotic pathogen within and between wildlife host species. Proceedings of the Royal Society of London Series B-Biological Sciences 266: 1939–1945.
- 74. 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.
- 75. Aeby GS (1998) A digenean metacercaria from the reef coral, Porites compressa, experimentally identified as Podocotyloides stenometra. Journal of Parasitology 84: 1259–1261.
- 76. Peters EC, Halas JC, McCarty HB (1986) Caliblastic neoplasms in Acroporapalmata, with a review of reports on anomalies of growth and form in corals. Journal of the National Cancer Institute 76: 895–912.
- 77. Coles SL, Seapy DG (1998) Ultraviolet absorbing compounds and tumorous growths on acroporid corals from Bandar Khayran, Gulf of Oman, Indian Ocean. Coral Reefs 17: 195–198.
- 78. Tedetti M, Sempere R (2006) Penetration of ultraviolet radiation in the marine environment. A review. Photochemistry and Photobiology 82: 389–397.
- 79. Ward JR, Kim K, Harvell CD (2007) Temperature affects coral disease resistance and pathogen growth. Marine Ecology-Progress Series 329: 115–121.
- 80. Patz JA, Olson SH (2006) Malaria risk and temperature: Influences from global climate change and local land use practices. Proceedings of the National Academy of Sciences of the United States of America 103: 5635–5636.
- 81. Hoegh-Guldberg O, Mumby PJ, Hooten AJ, Steneck RS, Greenfield P, et al. (2007) Coral reefs under rapid climate change and ocean acidification. Science 318: 1737–1742.
- 82. Miller J, Waara R, Muller E, Rogers C (2006) Coral bleaching and disease combine to cause extensive mortality on reefs in US Virgin Islands. Coral Reefs 25: 418–418.
- 83. Edmunds PJ (1991) Extent and effect of Black Band Disease on a Caribbean reef. Coral Reefs 10: 161–165.
- 84. Bruckner AW, Bruckner RJ (1997) The persistence of black band disease in Jamaica: impact on community structure. Proceedings of the 8th International Coral Reef Symposium, Smithsonian Tropical Research Institute, Panama 1: 601–606.
- 85. Kuta KG, Richardson LL (1996) Abundance and distribution of black band disease on coral reefs in the northern Florida Keys. Coral Reefs 15: 219–223.
- 86. Boyett HV, Bourne DG, Willis BL (2007) Elevated temperature and light enhance progression and spread of black band disease on staghorn corals of the Great Barrier Reef. Marine Biology 151: 1711–1720.
- 87. Bruckner AW, Bruckner RJ (1997) Outbreak of coral disease in Puerto Rico. Coral Reefs 16: 260.
- 88. Denner EBM, Smith GW, Busse HJ, Schumann P, Narzt T, et al. (2003) Aurantimonas coralicida gen. nov., sp nov., the causative agent of white plague type II on Caribbean scleractinian corals. International Journal of Systematic and Evolutionary Microbiology 53: 1115–1122.
- 89. Richardson LL (2004) Black band disease. In: Rosenberg E, Loya Y, eds. Coral Health and Disease. Berlin: Springer-Verlag. pp 325–336. 90. Sussman M, Willis BL, Victor S, Bourne DG (2008) Coral Pathogens Identified
- for White Syndrome (WS) Epizootics in the Indo-Pacific. PLoS ONE 3: e2393.
- 91. Voss JD, Richardson LL (2006) Coral diseases near Lee Stocking Island, Bahamas: Patterns and potential drivers. Diseases of Aquatic Organisms 69: 33–40.
- 92. Friedlander AM, Aeby G, Brainard R, Brown E, Chaston K, et al. (2008) The state of coral reef ecosystems of the Main Hawaiian Islands. In: Waddell J, ed. The State of Coral Reef Ecosystems of the United States and Pacific Freely Associated States: NOAA Technical Memorandum NOS NCCOS 11. NOAA/ NCCOS Center for coastal monitoring and assessment's biogeography team. pp 219–261.
- 93. Nakhapakorn K, Tripathi NK (2005) An information value based analysis of physical and climatic factors affecting dengue fever and dengue haemorrhagic fever incidence. International Journal of Health Geographics 4: 13.