I. Conceptual overview: The confluence of biological oceanography and disease ecology

Parasitic worm biomass can exceed the biomass of top predators in temperate estuaries (Kuris et al. 2008). Crustacean parasites reduce the survival (Finley and Forrester 2003, Forrester and Finley 2006), reproduction (Fogelman and Grutter 2008), and recruitment (Artim et al. 2015) of coral reef fishes. Viruses kill ~20–40% of marine bacteria daily – a level of microbial mortality that rivals the effect of grazing by zooplankton (Suttle 2005). These examples indicate that infectious processes are an important structuring force in ocean ecosystems. But what governs the success of these infectious agents? How are their populations regulated? **Can we understand and predict the population and community dynamics – and the resulting ecosystem-level effects – of parasites?**

Previous work by this team of PIs has demonstrated that there are two important predictors of marine parasite abundance and species composition: fishing (Wood et al. 2013, Wood et al. 2014b, Wood and Lafferty 2015) and primary production (Wood et al. 2015). Parasite species with complex life cycles (i.e., those that obligately require more than one host species) decline in the presence of fishing (Wood et al. 2014b, Wood et al. 2015; **Figure 1a**). On the other hand, directly transmitted parasites (i.e., those transmitted among hosts of the same species) can increase dramatically with fishing pressure (Wood et al. 2014b, Wood et al. 2015; **Figure 1a**). Meta-analysis reveals that these outcomes hold for parasites of finfish and invertebrates around the world (Wood and Lafferty 2015). Natural variation in productivity produces similarly strong effects on parasite community composition: increasing productivity tends to increase the abundance of complex life-cycle parasites, but has no detectable effect on directly transmitted parasites (**Figure 1b**). While these patterns are robust, the mechanisms that produce them have not been identified. Infectious processes are an important structuring force in ocean ecosystems, but there is little information with which to predict the distribution of infectious agents across space and time. Our proposal aims to fill that gap.

We hypothesize that the relationships we have observed between fishing and parasitism (**Figure 1a**) and between productivity and parasitism (**Figure 1b**) are driven by an intermediary variable: host biodiversity (**Figure 2**). Interest in the relationship between biodiversity and disease has expanded rapidly in the past decade, driven primarily by work in terrestrial and freshwater ecosystems (Johnson et al. 2015). These studies posit a number of mechanistic pathways by which shifts in host biodiversity may affect parasite dispersal, colonization success, resource availability, and resource heterogeneity (Keesing et al. 2006). Critically, of the 45 empirical studies into the diversity–disease relationship that have been conducted to date, **only three (6.7%) are from marine ecosystems** and none of these have been conducted in the tropics or on coral reefs (Civitello et al. 2015). We suspect that host biodiversity may be as important in shaping the dynamics of marine disease as it has proven to be in other ecosystems.

In the previous studies mentioned above (Wood et al. 2014b, Wood et al. 2015), we did not quantify host biodiversity. Instead, we placed a "black box" around host community composition, assuming that fishing and productivity had consistent effects on host diversity across study replicates. **The central goal of this proposal is to rigorously characterize host biodiversity in our study system – to understand the predictors of host biodiversity and develop taxonomically resolved estimates of biodiversity at each study site – and to use this information to identify potential links between host diversity and parasite transmission.** The work proposed here will constitute the first comprehensive test of leading hypotheses for the relationship between biodiversity and disease in any ecosystem, identifying the conditions under which positive, negative, and neutral diversity–disease relationships are to be expected.

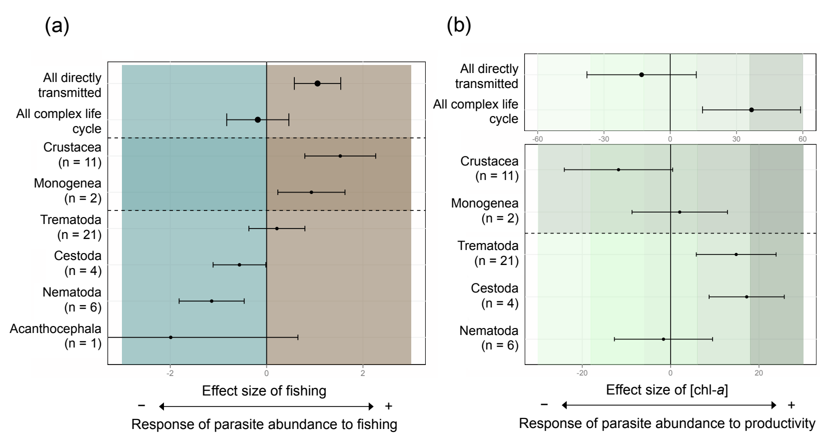


Figure 1. *Mean and 95% CI for the effect of (a) fishing and (b) primary productivity on parasite abundance for various groups of complex life cycle parasites (bottom, lighter shading) and directly transmitted parasites (top, darker shading). Adapted from Wood* et al. *2015.*

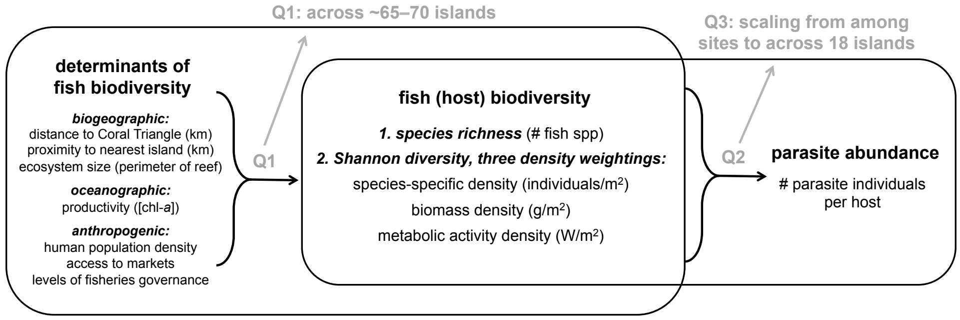


Figure 2. *Conceptual model of the project proposed here. Question 1 (Q1) will address links between environment and coral reef fish diversity at the island level, across all 65–70 islands in the final fish dataset. Question 2 (Q2) will then assess the consequences of variability in several dimensions of island-level host diversity for parasite abundance across >47 parasite species in seven coral reef fish hosts. Question 3 (Q3) will address how this diversity–disease relationship shifts depending upon the spatial scale of observation and the dispersal capacity of the parasite (among sites within islands to across 18 islands in the final parasite dataset).*

Here, we propose research that will **build the foundation for an entirely new sub-field of biological oceanography: the study of links between marine biodiversity and disease** (**Figure 2**). Marine ecosystems are uniquely suited to this work, and have driven substantial progress in community ecology (Poulin et al. 2016), yet marine models are seldom used in disease ecology research. Our proposed project will exploit the exceptional tractability of marine ecosystems (diverse parasite fauna, ease of collection and diagnosis of vertebrate hosts, strong and continuous spatial variability in host biodiversity, broad scope of spatial scales, biologically realistic sampling units) to address long-ignored gaps in disease ecology theory. We will accomplish this by integrating an **existing, fine-resolution dataset** on coral reef fish parasites from six equatorial Pacific islands (Wood et al. 2014b, Wood et al. 2015) with **new sampling from 12 similar, nearby islands**. The focal islands vary in the amount of host biodiversity they support, due to biogeographic, oceanographic, and anthropogenic causes (**Figure 2**). This natural experiment will allow us to explore how parasitism responds to variation in host biodiversity at an unprecedented spatial scale, spanning 18 coral islands of three major archipelagos. Efforts will be cost-leveraged by our existing dataset and an existing, funded field project – the 100 Island Challenge – that will provide easy access to remote field sites. Our project will address three fundamental questions:

Q1. Coral reefs are some of the most speciose ecosystems in the world, but past fish survey efforts have been coarsely resolved in both spatial and taxonomic dimensions, making it difficult to assess drivers of island-level patterns of fish diversity. What are the most **important correlates of fish (i.e., host) diversity** on coral islands of the equatorial Pacific?

Q2. Our preliminary data suggest that fishing and productivity are important drivers of parasite abundance, and we believe that these factors operate through effects on host biodiversity. What is the **shape of the relationship between biodiversity and parasite burden**? Are there non-linearities in that relationship? What factors (e.g., parasite traits like transmission strategy and host specificity) determine variability in its shape?

Q3. Heterogeneity in the spatial scale of disease transmission may influence whether a biodiversity–disease relationship is detectable at a particular spatial scale of observation. How does **spatial scale interact with parasite dispersal capacity** to moderate the effects of host biodiversity on disease?

This work represents a new frontier for biological oceanography. It will illuminate the drivers of abundance for parasites – a large, taxonomically diverse group of species that are ubiquitous, abundant, and influential and yet poorly understood. It will answer long-standing questions in disease ecology while introducing a new, highly tractable marine study system for the investigation of biodiversity–disease relationships. It also presents a timely opportunity for refining fundamental principles in ecology. Parasites comprise ~40% of Earth's species (Wood and Johnson 2015), yet the bulk of ecological theory has ignored parasites. Our proposed project therefore represents a unique opportunity for deepening our understanding of ecological complexity and significantly expanding ecological theory. Finally, our project will blend research and education by including undergraduates from minority-serving institutions as an integral part of our research team and by developing and disseminating active-learning modules on parasite biodiversity.

II. Background and significance: The diversity–disease relationship in disease ecology theory

Our prior work hints at a role for biodiversity in determining the abundance of parasites. The disease ecology literature is rich in hypotheses for the mechanisms that govern diversity–disease relationships, some of which may explain the patterns we have already observed (McCallum et al. 2004).

The dilution effect hypothesis. A dominant theory in the subset of the community ecology literature that addresses disease, the "dilution effect" hypothesis was developed to explain patterns of Lyme disease (*Borrelia burgdorferi*) prevalence in forest mammals of northeastern North America (van Buskirk and Ostfeld 1995, Norman et al. 1999, Ostfeld and Keesing 2000), but in recent years has been advanced as a general explanation for parasite transmission patterns across ecosystems (Keesing et al. 2010, Keesing and Ostfeld 2012). The theory posits that human impacts on ecosystems should increase parasite transmission by (*i*) reducing biodiversity, increasing the proportion of competent host species in the population, and thereby increasing contacts among competent hosts, or (*ii*) reducing biodiversity, reducing competition and predation for competent host species, and thereby increasing the density of competent hosts. Note that this is a diffuse effect: the subset of biodiversity that is predicted to affect parasite transmission contains both hosts and non-hosts. The theoretical antecedents for the dilution effect hypothesis include elements of biodiversity theory, including two important relationships: that between biodiversity and ecosystem function (BEF) in general (Cardinale et al. 2012) and between biodiversity and invasibility in particular (Fridley et al. 2007). The dilution effect also suffers some of the weaknesses of its predecessors, including questions of (*i*) whether ecosystem services, invasion resistance, and dilution arise from biodiversity *per se* or from the fact that biodiverse assemblages are merely more likely to contain an especially influential species (the "sampling effect"; Cardinale et al. 2006) or to cap the density of an especially influential species (Begon 2008), and (*ii*) whether the results of biodiversity experiments translate to landscape and regional scales (Loreau et al. 2001, Cohen et al. 2016). Nonetheless, since the original studies on Lyme disease (reviewed in Ostfeld 2011), evidence has accumulated to indicate that the dilution effect operates at local scales for a variety of wildlife infections (Clay et al. 2009 a,b, Hall et al. 2009, Haas et al. 2011, Searle et al. 2011, Venesky et al. 2013). A comprehensive meta-analysis revealed a negative relationship between diversity and disease (hereafter, the diversity–disease relationship) across studies (Civitello et al. 2015). If sufficiently general, the dilution effect would produce health benefits as a side effect of conservation – a valuable ecosystem service.

The diversity-begets-diversity hypothesis. In recent years, a competing hypothesis has been advanced, predicting the opposite outcome: a positive diversity–disease relationship. The "host-diversity-begets-parasite-diversity" hypothesis ("diversity-begets-diversity" for short) argues that parasites are indicators of a healthy, functioning ecosystem because parasites depend on primary ecosystem processes like predation to complete their life cycles and because they are often highly host specific or require multiple host species for the completion of their life cycles (Hechinger and Lafferty 2005, Hudson et al. 2006). This hypothesis is philosophically rooted in island biogeography theory (MacArthur and Wilson 1967, Kuris et al. 2010), meta-community theory (Leibold et al. 2004), and the habitat heterogeneity hypothesis (Hutchinson 1959, MacArthur and MacArthur 1961, Rosenzweig 1995). Empirical support for "diversity begetting diversity" is ample (Kamiya et al. 2014, Poulin 2014). Given the contradictory nature of the predictions and supporting empirical evidence for the dilution and diversity-begets-diversity hypotheses, there is a dire need for work that reconciles these two competing hypotheses.

A unifying framework for the diversity–disease relationship. While "dilution" appears superficially to be at odds with "diversity-begets-diversity," these two hypotheses can be reconciled: previous studies of the diversity–disease relationship have often focused on a single parasite at a time (e.g., *B. burgdorferi*), but by considering a cross-section of parasites in an ecosystem, a range of responses to diversity are observed. In our prior work, we assumed that fishing reduced host biodiversity. This work demonstrated that – when screening across a variety of parasites, instead of focusing on a single species – dilution is observed, but other responses can also occur, including increases in transmission in response to decreases in fishing, consistent with the ddiversity-begets-diversity hypothesis (Lafferty and Wood 2013, Wood and Lafferty 2013, Wood et al. 2014a, Wood et al. 2016). We surmise that the contradiction of the evidence supporting these two hypotheses stems, at least in part, from the fact that there is variability among parasites species in the direction of their response to biodiversity – even for parasite species in the same ecosystem – and, further, that within a single parasite species, there may be non-linearity and scale-dependence of the diversity–disease relationship. The research proposed here will test this by (*i*) measuring host biodiversity, thereby facilitating a direct test of the dilution versus diversity-begets-diversity hypotheses across >47 parasite taxa, (*ii*) generating biodiversity–disease data with a highly resolved biodiversity axis, which will permit exploration of the shape of the biodiversity–disease relationship and conditions under which each shape is likely, and (3) expanding our sampling area, which will allow us to understand the influence of spatial scale on this relationship. Predicting the outcome of the biodiversity–disease relationship across parasite species would greatly improve our understanding of how parasite impacts on ecosystems may be distributed across space and time.

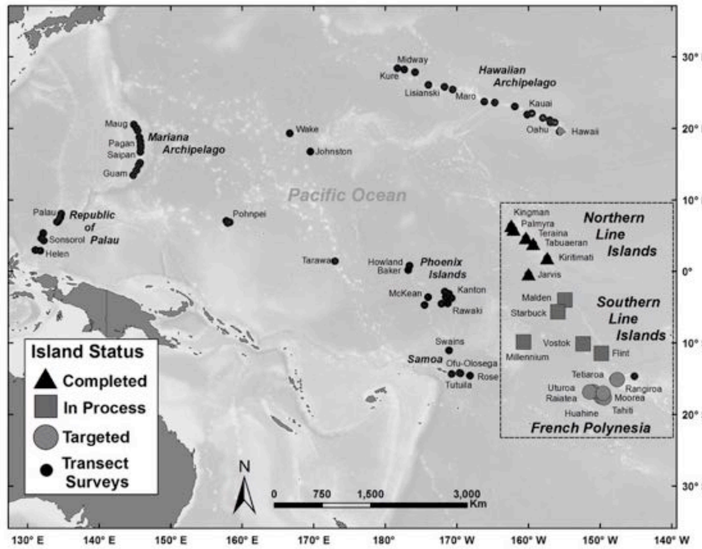
Why does the diversity–disease relationship matter? What is the point of trying to understand the factors that govern the distribution of marine parasites? Parasites are often hidden and can be easy to overlook, but they are important both ecologically and in the opportunities they offer for advancing our understanding of ecological complexity. That parasites are ecologically influential and ubiquitous across ecosystems has been firmly established (Wood and Johnson 2015 and references therein). For example, parasites commonly manipulate the behavior of hosts to facilitate onward transmission, often by increasing the likelihood that the host is preyed upon by the predator species that serves as the parasite's next host (Poulin 2010, Hughes et al. 2012); in this way, parasites effectively shunt energy into higher trophic levels, subsidizing predators (Wood and Johnson 2015). Parasites can also control host populations through negative fitness effects (i.e., disease), including castration (e.g., Decaestecker et al. 2005), re-shape host communities by influencing the outcome of species interactions like competition, predation, and herbivory (e.g., Wood et al. 2007), and can even affect ecosystem nutrient cycling (e.g., Mischler et al. 2016). **This substantial ecological role for parasites suggests that downstream effects of change in parasite abundance are likely to be significant for every marine population from zooplankton to cetaceans.** For example, the complex life-cycle parasites that increase predation rates on their hosts (discussed above) decline substantially in the presence of fishing; we surmise that this loss of manipulative parasites might actually reduce ecosystem-level energy flows from lower to higher trophic levels. But before we can pursue a clearer understanding of how parasites control marine ecosystem processes, we must know when and where parasites are most abundant. The research proposed here addresses this unknown.

III. Coral reefs of equatorial Pacific islands: A test bed for the diversity–disease relationship

Disease ecology theory has been developed primarily with reference to terrestrial and freshwater ecosystems, but marine systems are an equally appropriate – and in some cases, more tractable – test bed for these hypotheses (McCallum et al. 2004, Poulin et al. 2016). Our study system contains discrete replicates (i.e., islands) that are disjunct from one another and large, thereby balancing statistical independence, biological realism, and scope for exploring the effects of spatial scale. Because most marine parasites have a pelagic larval stage, we can explicitly discriminate among long- and short-distance transmission in a way that is not possible on land or in freshwater. This gives us the ability to explore the influence of dispersal capacity on infectious processes. A final major advantage of our natural experiment is its potential for expansion; there are thousands of islands in the Polynesian/Micronesian region of the central equatorial Pacific, including dozens in the vicinity of our original six islands. Here, we propose to build on our existing six-island dataset with 12 new islands (**Figure 3**), but there is ample scope for further expansion in the future.

Over the past 10 years, our team has gained extensive experience in sampling fish from coral reefs of Pacific islands, and we have been collaborating for the past six years to characterize the parasite assemblages of these islands. Presently, our database contains records of 1,839 fish hosts of seven species from six islands (the Northern Line Islands), which yielded 988,071 individual parasites from >47 taxa, including monogeneans, trematodes, tapeworms, nematodes, acanthocephalans, and crustaceans. We currently possess an additional 875 unprocessed fish of these seven species collected from the five islands of the Southern Line Islands archipelago, and we propose additional collections at seven islands in French Polynesia (**Figure 3**). We also currently possess a 56-island dataset of fish visual survey data, including records of more than three million individual fish, which will allow us to make site-level, species-level estimates of fish density and to thereby calculate fish diversity. All existing data and the data that we propose to collect will be formally cataloged and made publicly available through BCO-DMO. These data offer a unique opportunity to critically test the underpinnings of the relationship between diversity and disease.

__Figure 3.__ Map of the study region, including the six islands in our existing dataset of parasite burden (black triangles, Northern Line Islands), five islands where fish collection is complete and support is requested for sample processing (grey squares, Southern Line Islands), and seven islands where support is requested for both fish collection and sample processing (grey circles, French Polynesia). Also shown are additional islands currently included in our existing dataset of species-level fish density estimates (black dots indicating transect surveys).



IV. Research questions, preliminary data, and methods

The central goal of this proposal is to rigorously characterize host biodiversity in our study system – to understand the predictors of host biodiversity and develop taxonomically resolved estimates of host biodiversity at each study site – and to use this information to identify potential links between host diversity and parasite transmission. We will proceed using existing data and data from new sampling proposed below.

Q1: What are the most important correlates of fish (i.e., host) diversity on coral islands of the equatorial Pacific?

In previous work on the diversity–disease relationship, defining, delimiting, and measuring host biodiversity has presented a substantial challenge (Wood and Johnson 2016). The dilution effect and the diversity-begets-diversity effect are functions of both species richness and relative abundance (Keesing et al. 2006), but despite this, most studies of these phenomena focus exclusively on species richness (Civitello et al. 2015). Furthermore, as the product of species interactions, we might expect these phenomena to manifest primarily at fine spatial resolutions, where biotic interactions can exert the greatest influence on ecological processes, rather than at coarse resolutions, where larger-scale processes like climate wield the most influence (Fridley et al. 2007, Vellend 2016), but few disease ecology studies consider the influence of spatial scale on the likelihood of detecting dilution or diversity-begets-diversity effects (Cohen et al. 2016). Finally, most studies of the diversity–disease hypothesis consider only a truncated range of biodiversity (e.g., 1–6 species in Johnson et al. 2013). We propose to overcome these challenges by rigorously characterizing the diversity of hosts (i.e., coral reef fishes) at a fine spatial resolution across a broad geographic expanse and a wide range of diversity values (**Figure 4**). In addition to providing necessary host diversity estimates as inputs for Questions 2 and 3, this effort will yield data that answer outstanding questions about the spatial distribution of coral reef fish diversity.

What are the determinants of coral reef fish biodiversity? Coral reefs have higher vertebrate biodiversity than any other ecosystem on the planet (Reaka-Kudla et al. 1996, Knowlton 2001), but among-reef variability in diversity is poorly understood (Williams 1982, Mora et al. 2003). Substantial research effort has been invested in mapping species ranges (Sale 1977, Randall 1998) and in explaining distributional patterns of reef fish assemblages (Bellwood and Hughes 2001, Bellwood and Wainwright 2002, Hughes et al. 2002, Bellwood et al. 2005, Sandin et al. 2008a, Hobbs et al. 2012, Bowen et al. 2013). These studies indicate that proximity to the Coral Triangle, distance to “stepping-stone” neighbors, and island size predict high coral reef fish species richness. But even islands that possess similar total species richness can exhibit substantial differences in species relative abundance, linked in many cases to differences in trophic structure (McClanahan 1994, Jennings et al. 1995, Stevenson et al. 2007, DeMartini et al. 2008). Although patterns of relative species abundance and associated diversity metrics (e.g., area-controlled richness, Shannon diversity) have been reported from a handful of sites (Williams and Hatcher 1983, Watson et al. 1996, Ohman and Rajasuriya 1998, Tittensor et al. 2007, Halpern and Floeter 2008), empirical data across broad spatial scales are challenging to generate: the logistical demands of such assessments are substantial, requiring detailed, fine-scale, high-taxonomic-resolution information on fish replicated across an expansive geographical scope.

Nonetheless, several efforts have characterized relative levels of fish biodiversity, considering the drivers and consequences of variation across regional (100s–1000s of km) scales. To address the need for broad spatial coverage, such efforts often combine data from multiple research groups that use variable methods of fish enumeration (Mora et al. 2001, 2003, Cinner et al. 2016). In order to account for potential biases in comparing data across methodologies, these studies have depended upon summarized metrics of fish diversity, such as family-level or functional-group richness (Tittensor et al. 2007, Mora et al. 2011). Functional group analyses can reveal ecological patterns linked, for instance, to energy flow and trophic dynamics (Bellwood et al 2004), but lack the resolution necessary for understanding topologically structured ecological dynamics (e.g., abundance of host-specific parasites). Further, with growing evidence suggesting that even congeneric and co-occurring reef fishes can have surprisingly different feeding behaviors (McMahon et al. 2016), there is reason to believe that assessments of relative diversity at finer taxonomic resolution are needed to understand community functioning (Wright et al. 2006).

We have a unique opportunity to overcome these empirical challenges, through the use of a large and growing dataset collected across a broad swath of the central Pacific by our research group and close collaborators (Brainard et al. 2005, DeMartini et al. 2008, Sandin et al. 2008b, Williams et al. 2011, Zgliczynski and Sandin 2017). The data have been collected by a limited subset of trained observers employing directly comparable methods with high taxonomic resolution. We will combine these data with geographic information to assess the importance of various correlates for coral reef fish diversity. Where previous efforts have considered only the drivers of coral reef fish richness, we have the ability to simultaneously assess effects on fish relative abundance. Where previous efforts have compromised spatial resolution, taxonomic resolution, or both, we will compromise neither. This unprecedented dataset gives us the ability to rigorously address the question: what are the most important correlates of fish biodiversity on coral islands of the central Pacific? Our effort will not only provide accurate estimates of host diversity for understanding disease dynamics (Questions 2 and 3), but will stand on its own as a fundamental contribution to coral reef ecology.

Fish survey data – To assess correlates of fish biodiversity, we will use and build on an existing dataset of fish survey data from the tropical equatorial Pacific. Our existing database contains results of size- and species-specific surveys collected over the past 10 years by members of the Sandin lab and the National Oceanic and Atmospheric Administration's Coral Reef Ecosystem Program (part of the Pacific Islands Fisheries Science Center). The existing data include surveys from 56 islands and atolls from five archipelagos or geographic regions – the Hawaiian Archipelago, Commonwealth of the Northern Marianas, US Pacific Remote Islands Marine National Monument, American Samoa, and Micronesia. We will complement these existing data with additional surveys from more islands in Micronesia (including islands within the Federated States of Micronesia and the Republic of Palau, data collection supported by the 100 Island Challenge; see *Facilities, Equipment, and Other Resources*) and French Polynesia (data collection supported by this proposal; see *Letter of Collaboration* from Serge Planes, CRIOBE, Moorea, French Polynesia). The final dataset will include 65–70 islands with island-specific survey effort ranging from 7 to 50 transects located between 7 and 15 m depth on the leeward forereef of each surveyed island. All fish survey data have been and will continue to be collected using the same methodology, with methods described in detail in Friedlander et al. (2010) and Zgliczynski and Sandin (2017).

Predictor variables – We will consider the influence of biogeographic, oceanographic, and anthropogenic factors as correlates of fish host diversity. <u>*Biogeographic*</u> – Distance to the Coral Triangle, distance to the nearest neighbor island, and island size may be important determinants of island-level coral reef fish richness (Hughes et al. 2002, Sandin et al. 2008a); the influence of such factors Shannon diversity is less well understood. <u>*Oceanographic*</u> – Allochthonous resource delivery has direct influence on the trophic opportunities and overall productivity of coral reef communities. High-productivity systems might support more feeding niches and hence more diversity per unit area than low-productivity systems; alternately, high production could lead to hyper-abundance of a subset of opportunistic taxa (Wright 1983, Waide et al. 1999, Gaston 2000). <u>*Anthropogenic*</u> – Fishing has been shown to shift trophic structure, reduce standing stock biomass, and shift size structure of coral reef fish communities toward small body sizes (Dulvy et al. 2004, DeMartini et al. 2008, Sandin et al. 2008b). The influence of harvest on patterns of diversity is less clear, though preliminary insights point to declines in trophic diversity (Tittensor et al. 2007) and taxonomic diversity (**Figure 4**) with exploitation.

Estimates of *biogeographic*, *oceanographic*, and *anthropogenic* context will be calculated at the island scale, as follows. *Biogeographic* context in the insular Pacific will be estimated with island perimeter, proximity to nearby large islands (*sensu* island biogeographic models; MacArthur and Wilson 1967, Sandin et al. 2008a), and proximity to the biodiversity "hotspot" of the Coral Triangle (Bellwood and Hughes 2001, Reaka et al. 2008). *Oceanographic* context will be estimated with nearshore productivity, using appropriately area-scaled and masked grid cells from satellite-derived ocean color data (using MODIS maps estimating surface chlorophyll-*a* concentrations; methodological details in Gove et al. 2013). Briefly, long-term means of chl-*a* (mg/m^3) will be estimated by sampling the eight-day 0.0417^o (~4-km) spatial resolution product for targeted islands over a 10-year period (2006–2016). *Anthropogenic* context will be estimated with metrics of demonstrated descriptive power, including human population density, access to markets, and levels of governance (Cinner et al. 2009, Cinner et al. 2016).

Response variables – The fish survey data provide estimates of fish diversity at each survey site, at a spatial grain of 100 m^2 for large fish (≥20 cm in total length) and 50 m^2 for small fish (<20 cm). We will use these raw data to develop estimates of fish biodiversity that account for probabilities of encounter. As such, we will calculate metrics of total richness at the island scale, one of the "go-to" metrics for coral reef fishes (Bellwood and Hughes 2001, Sandin et al. 2008a, Parravicini et al. 2013). But we will also develop metrics of density-weighted Shannon diversity. These metrics provide information on the expected number of species to be found within bounded areas and the probability of finding a new species when randomly encountering new individuals. Such metrics have value both intrinsically for understanding biodiversity and mechanistically for considering linkages with associated parasites.

Most applications of Shannon diversity are density-weighted based upon relative numbers of individuals in a sample. However, number of individuals may not describe functional elements of fish assemblages, such as spawning potential, trophic impact, and energy flow. Patterns of relative diversity at the site level may, for example, be influenced by patterns of competition for food or relative impact of parasites. As such, we follow the advice of Wilhm (1968) to investigate patterns of Shannon diversity using three forms of density weightings – species-specific density (individuals/m^2), biomass density (g/m^2), and metabolic activity density (W/m^2). Density of individuals will be estimated directly from fish survey data. Biomass density will be estimated based upon raw estimates of individual-specific length scaled with species-specific allometric scaling constants (length–weight parameters available from multiple sources, most linked through FishBase; Froese and Pauly 2010). Metabolic activity density will be estimated based upon allometric scaling approximations (Brown et al. 2004). The three forms of weighting lie on a continuum of emphasis on individual-specific size, from no emphasis (density) to moderate (metabolic activity) and linear (biomass density) emphasis. The results of replicated modeling of Shannon diversity across metrics will reveal novel linkages with environmental factors (Wilhm 1968). Further, we will use the diversity outputs across metrics to explore candidate models linking parasite diversity to competing functional interpretations of host diversity.

Statistical analyses – We will employ a linear-model fitting approach to explore the relative contribution of the defined biogeographic, oceanographic, and anthropogenic parameters (after removal of collinear predictors) in describing variation in site-specific fish *species richness* and *density-weighted diversity* (i.e., Shannon diversity). Given the design of the fish survey data, we have the opportunity to consider patterns of relative diversity at an island scale with uniquely high statistical power (65–70 islands). Response data will be nested within island, given that multiple surveys are conducted on each island (7–30 surveys per island). Analyses will be constrained to first-order contributions of predictor variables (to avoid pitfalls of over-fitting), and link functions will be selected based upon distributions of response data.

Preliminary results – While we might expect independent and interacting effects of biogeography, oceanography, and human activities on the patterns of relative diversity of reef fishes (see *Predictor variables,* above), preliminary analyses reveal at least one common correlate across islands: oceanic productivity (**Figure 4**). The relationship between productivity and Shannon diversity might arise because factors that lead to increases in density of fishes (e.g., increases in food availability) should lead to an increase in the area-specific diversity (Gotelli and Colwell 2001). In contrast, independent factors that shift the fish assemblage toward faster dynamics and turnover, such as anthropogenic disturbances that mimic shifts to early successional systems (Odum 1969, Sandin and Sala 2012), are expected to lead to increasing numerical dominance of a subset of opportunistic species; "faster" systems will be expected to have lower Shannon diversity relative to "slower" systems (e.g., see negative deviation of inhabited islands in **Figure 4b**). In sum, alterations of trophic flow hold potential to shift the relative diversity of fishes; such shifts will systematically alter the availability of hosts for associated guilds of parasites.

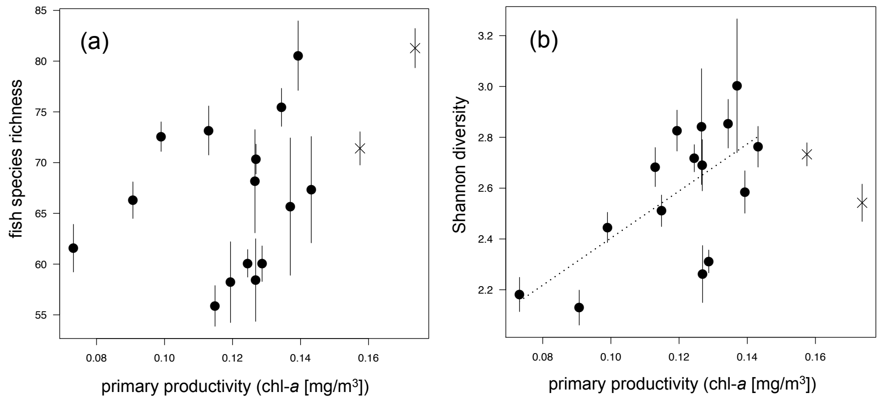


Figure 4. *Across 15 uninhabited islands (black dots) of the Line and Phoenix Islands archipelagos, oceanic productivity is (a) not correlated with coral reef fish species richness but is (b) a strong positive predictor of coral reef fish Shannon diversity (species-specific density weighted). Two inhabited (and fished) islands are also indicated (with x symbols). Estimates of primary production were derived using the method explained above. Points indicate means and whiskers indicate standard errors.*

Q2: What is the shape of the relationship between diversity and disease?
Previous studies on the diversity–disease relationship have bookended the axis of diversity and made untested assumptions about the relationship between disturbance and diversity, grouping replicates into discrete categories: “fished” versus “unfished” (e.g., Wood et al. 2014b, Wood et al. 2015), “de-forested” versus “forested” (e.g., Bauch et al. 2015), and “large wildlife removed” versus “large wildlife retained” (e.g., Young et al. 2015). This approach has powerfully demonstrated that disturbances like fishing can drive change in parasite assemblages. However, it masks the functional shape of the relationship between biodiversity change and parasitism by putting host biodiversity in a “black box”. Does reducing biodiversity cause a decline in parasite transmission, as the dilution effect predicts (**Figure 5a[i]**)? Might transmission be highest at intermediate levels of biodiversity, where competent hosts have the maximum joint likelihood of being (1) present and (2) abundant, in the absence of competitors or predators (**Figure 5a[ii]**)? Or might increasing biodiversity increase disease risk, either linearly (**Figure 5a[iii]**), with saturation at high levels of biodiversity (**Figure 5a[iv]**), or with a peak at high levels of biodiversity, where intensifying species interactions begin to exert control on the density of any single host, as the dilution effect predicts (**Figure 5a[v]**)? Tests of existing theory depend upon a complete understanding of the trajectory of disease change – not just selective subsets of host diversity (e.g., dashed box in **Figure 5a**). Our dataset will allow us to explore the effects of a broad range of host richness (a range of ~55–80 species; **Figure 4**) and Shannon diversity (a range of ~2.2–3.0; **Figure 4**) values.

There are a few unique properties of the relationship between diversity and disease. First, parasites obligately require hosts. Consequently, when biodiversity equals zero, parasite abundance must also equal zero (Wood et al. 2016; **Figure 5a**). This constraint obviates the possibility of convex relationships, but a diversity of concave relationships might result. In studies contrasting disease outcomes as a function of categorical biodiversity states (e.g., high biodiversity versus low biodiversity), it has often been assumed that a linear relationship connects the bookend categories. But depending on the magnitude of the biodiversity value of each category, this assumption could lead to faulty interpretations – even the opposite of the actual, underlying pattern (e.g., dashed box in **Figure 5a**).

To explore the shape of the diversity–disease relationship across a broad and well-resolved scope of host biodiversity values, we will build on our existing dataset (**Figure 3**). Rather than categorizing these islands as “fished” or “unfished”, we will use four continuous metrics for characterizing the biodiversity of hosts that occur in the ecosystems of equatorial Pacific coral reefs (see Q1, above). We will then examine the diversity–disease relationship for each of an expected >47 parasite species (i.e., at least as many species as occur in our current, six-island dataset), providing a comprehensive portrayal of the biodiversity–disease relationship across substantial variability in host biodiversity, parasite taxonomic diversity, parasite transmission strategies, and host specificities (Wood et al. 2014b, Wood et al. 2015).

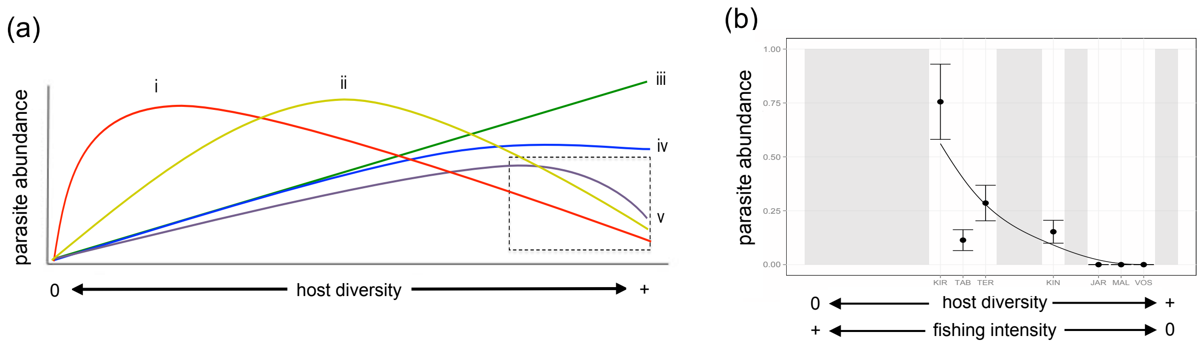


Figure 5. *(a) Hypotheses for the relationship between host diversity and parasite abundance. Each line represents the relationship between abundance of a single parasite species and site-level host diversity. (i) The dilution effect, (ii) an intermediate hypothesis, where dilution mechanisms prevail at high host diversity and diversity-begets-diversity mechanisms prevail at low host diversity, (iii) diversity-begets-diversity, (iv) diversity-begets-diversity with saturation at high levels of host diversity, and (v) diversity-begets-diversity that shifts to dilution at high levels of biodiversity. The dashed box indicates how a selective frame of reference can lead to spurious conclusions. Figure adapted from (Wood* et al. *2016). (b) From our existing dataset, abundance of a Cucullanidae sp. nematode in the whitecheek surgeonfish* Acanthurus nigricans, *at five islands sampled in 2010 (KIR = Kiritimati, TAB = Tabuaeran, TER = Teraina, KIN = Kingman, JAR = Jarvis) and two additional islands sampled in 2013 (MAL = Malden, VOS = Vostok). Islands are arranged along the x-axis in order of predicted host diversity (proxied by fishing intensity). Gray areas indicate host diversity space that will be sampled in the proposed project. Line indicates LOESS smoothing regression on means for each island, span = 2.*

We will synthesize data on the parasite burden of seven host species from 18 islands of the central equatorial Pacific (**Figure 3**), including the six Northern Line Islands already included in our database, five Southern Line Islands where we have already collected 875 fish but request support for processing these samples, and seven French Polynesian islands where we propose additional fish collections and processing. Our existing database (held by PI Wood) currently reports counts of all metazoan parasites detected in seven host species from 58 geo-referenced sites at six islands (Jarvis, Kingman, Palmyra, Teraina, Tabuaeran, Kiritimati). Presently, the database contains records of 1,839 fish hosts that yielded 988,071 individual parasites from >47 taxa, including monogeneans, trematodes, tapeworms, nematodes, acanthocephalans, and crustaceans. The seven host species (*Cephalopholis urodeta, Acanthurus nigricans, Ctenochaetus marginatus, Paracirrhites arcatus, Stegastes aureus, Chromis margaritifer,* and *Pseudanthias bartlettorum*) were chosen to span a variety of body sizes, to represent important trophic and taxonomic groupings, and to include only broadly distributed species that would be present at most reefs in the region (Wood et al. 2014b, Wood et al. 2015).

Fish host sampling – Lethal fish sampling will be conducted at each of the seven French Polynesian islands where we have not already conducted fish sampling (**Figure 3**), will include individuals of the seven species targeted in our previous efforts, and will match fish survey locations (see *Letter of Collaboration* from Serge Planes, CRIOBE, Moorea, French Polynesia). The resulting spatially extensive dataset spanning 18 islands is necessary for two reasons: (*i*) to expand and resolve the axis of host diversity (**Figure 5b**) and (*ii*) to expand the scope of spatial scales available for investigation in Question 3 (see below). Fish will be captured by scuba divers using three-pronged spears (for fish >10 cm) and hand nets (for fish <10 cm) and will be euthanized immediately upon capture by pithing (as approved by IACUC, see attached documentation). At least 25 individuals will be collected per species per island, a level of replication that represents a balance between, on one hand, attaining sufficiently high replication to accurately assess the abundance of parasites (which are typically over-dispersed in hosts) and, on the other hand, minimizing ecological impacts on reefs and the number of animals that experience discomfort, pain, and injury. Sampling will be conducted at 3–5 sites within each island, with sites chosen to match as closely as possible those where fish surveys are conducted (see above).

All sampling will occur between June and November of year 1 of the grant (2017), to minimize the amount of time that has passed since prior collections (2010 for Northern Line Islands, 2013 for Southern Line Islands) and biodiversity surveys (2005 for Northern Line Islands, 2009 for Southern Line Islands). The amount of time between sampling bouts raises the risk that short-term fluctuations might decouple relationships between parasite abundance and host biodiversity. Given the relative longevity of parasites in their hosts (Lafferty and Kuris 2002), the aseasonality of these ecosystems, and the strong inter-island variability we have thus far observed in both parasite abundance and host biodiversity, we view the potential for significant temporal variability in these collections to be low, but time will be included in all statistical models to account for its potential effects.

Fish will be processed in PI Wood's parasite ecology lab at the University of Washington. At collection, each fish will be deposited in a separate ziplock bag with a tag individually identifying the specimen. Total length, standard length, fork length, and mass will be recorded. After measurements are taken, fish will be frozen as quickly as possible and shipped on dry ice to PI Wood's facility, where they will be kept frozen until processing. In addition to performing parasite sampling, we will also photograph each fish and archive fish tissues (see *Data Management Plan*).

Parasite sampling – A comprehensive parasitological examination will be performed for each fish, designed to detect all metazoan parasites. The following organs will be examined individually under a stereomicroscope: fins, gills, eyes, heart, liver, spleen, gonad, gills, muscle, skin, and intestines (after Wood et al. 2014b). Photographs of each parasite taxon and detailed images of diagnostic morphological features will be linked to our database (see *Data Management Plan*). Parasites will be identified to the lowest possible taxonomic level using general guides (Yamaguti 1963, Schultz 1969, Skryabin 1991, Khalil et al. 1994, Gibson et al. 2002, Kabata 2003, Bray et al. 2005, 2008). For data on parasite life history traits (e.g., host specificity), we will build on an existing database (see Appendix F in Wood et al. 2014b). We will preserve parasite vouchers in ethanol for genetic analyses (Question 3) and as an archived resource for other researchers to access and use (see *Data Management Plan*).

Analyses – We will begin by plotting, for each parasite species detected and each of our four metrics of host diversity, the relationship between diversity and mean abundance of parasites (number of parasites per host; e.g., **Figure 5b**) at the island level (i.e., island-level host diversity versus island-level parasite abundance; after Wood et al. 2014b, Wood et al. 2015). Both the dilution effect hypothesis (Keesing et al. 2006) and the diversity-begets-diversity hypothesis (Hechinger and Lafferty 2005) predict a relationship between diversity "writ large" (i.e., of both hosts and non-hosts) and parasite abundance. A generalized linear mixed-effects model (GLMM) with negative binomial error structure will be used to assess whether there is a significant linear relationship between host diversity and parasite abundance. Models will include fixed effects of latitude (to control for covariance in hosts and parasites not related to their direct relationship, but instead related to some third variable correlated with latitude) and host body size (to control for positive effects of host size on parasite burden), and random effects of fish ID nested within site nested within island. Those models with non-random residual structure indicating possible non-linearity in the host diversity–parasite abundance relationship will be assessed for change points. Three techniques for change point detection will be used to ensure that change point values are robust (Aguilar et al. 2003, Gsell et al. 2016): piecewise linear regression (Muggeo 2003), the non-parametric Pettitt rank test (Pettitt 1979), and the sequential regime shift detector (SRSD) method (Rodionov 2004, 2015). For parasites that display change points, we will split the data into lower-biodiversity-than-change-point and higher-biodiversity-than-change-point sections, and compare their mean abundances with *t*-tests to derive a coefficient (standardized regression coefficient for the effect of "high" versus "low" biodiversity) that describes the direction and magnitude of the biodiversity–disease relationship for each parasite and is comparable among linear and non-linear responses to host biodiversity. Finally, we will synthesize these diversity–disease relationships for all >47 parasite species using meta-analytic statistical techniques (as in Wood et al. 2014b) to test whether the direction and magnitude of the relationship and the value of any change points are: (**Q2a**) different from zero when averaged across all parasite taxa, (**Q2b**) consistent across parasite transmission strategies (i.e., does the shape of the diversity–disease relationship vary depending on whether the parasite is directly transmitted species or a complex life-cycle species?) and (**Q2c**) related to host specificity (i.e., degree of host specialization).

Anticipated results – We expect that these data will support both the dilution effect hypothesis (**Figure 5a[i]**) and the diversity-begets-diversity hypothesis (**Figure 5a[iii], 5a[iv], 5a[v]**), because we predict that there will be significant variation in the shape and strength of the host diversity–parasite abundance relationship among parasite species. We also expect diverse change points among parasite taxa (with low change points indicating resilience to host biodiversity change and high change points indicating susceptibility), that dilution will dominate among directly transmitted monogenean and crustacean parasites and those parasites with low host specificity, and that diversity-begets-diversity will dominate among trophically transmitted trematode, cestode, nematode, and acanthocephalan parasites and those parasites with high host specificity. For one parasite (a nematode in the family Cucullanidae) in our existing dataset, dilution (as in **Figure 5a[i]**) is supported, but significant gaps in the axis of host diversity remain (**Figure 5c**). Our proposed sampling will fill those gaps.

Q3: How does spatial scale interact with parasite dispersal capacity to moderate the strength of the relationship between biodiversity and disease?

Ecological pattern tends to be governed by regional forces like climate at coarse spatial resolutions and local forces like species interactions at fine spatial resolutions (e.g., Fridley et al. 2007, Menge et al. 2015). Layered on these scale-dependent processes are cross-scale (i.e., meta-ecosystem) flows of organisms, nutrients, and energy (Loreau et al. 2003, Loreau and Holt 2004, Menge et al. 2015). But the same spatial grain is experienced in vastly different ways by different organisms; for example, disturbance to a 1-ha area might be catastrophic for a sedentary species and meaningless to a vagile species. Like many marine organisms, parasites produce pelagic larvae whose dispersal distances depend on the pelagic larval duration (PLD; i.e., the amount of time during which the planktonic parasite larva is competent). In addition to the dispersal capacity inherent in the pelagic larval phase, parasites can also disperse across space by infecting a vagile host (Blasco-Costa and Poulin 2013). As a result, parasites like monogenean flatworms, which produce weakly motile larvae with short PLDs, may disperse only tens or hundreds of meters (**Figure 6a**), while cestode flatworms, which produce larval stages with long PLDs, use zooplankton as intermediate hosts, and infect highly vagile elasmobranch final hosts, probably disperse thousands of kilometers in one turn of the life cycle (**Figure 6b**). The genetic structure of parasite populations is strongly influenced by transmission strategy among parasitic flatworms (Davies et al. 1999, Criscione and Blouin 2004, Blasco-Costa and Poulin 2013), parasitic nematodes (Blouin et al. 1999), lice (Johnson et al. 2002), and ticks (McCoy et al. 2005) and parasite populations are often more genetically subdivided than their hosts' populations (Criscione et al. 2006). This suggests that dispersal capacity has important consequences for structuring parasite populations, yet few studies have directly addressed whether a parasite's dispersal capacity mediates its response to biodiversity (but see categorical comparisons in Wood et al. 2013, Johnson et al. 2016). This problem has clear parallels in island biogeography (MacArthur and Wilson 1967), meta-community dynamics (Leibold et al. 2004), and other bodies of theory; here, we have the opportunity to test whether these foundational ecological theories hold among parasitic organisms.

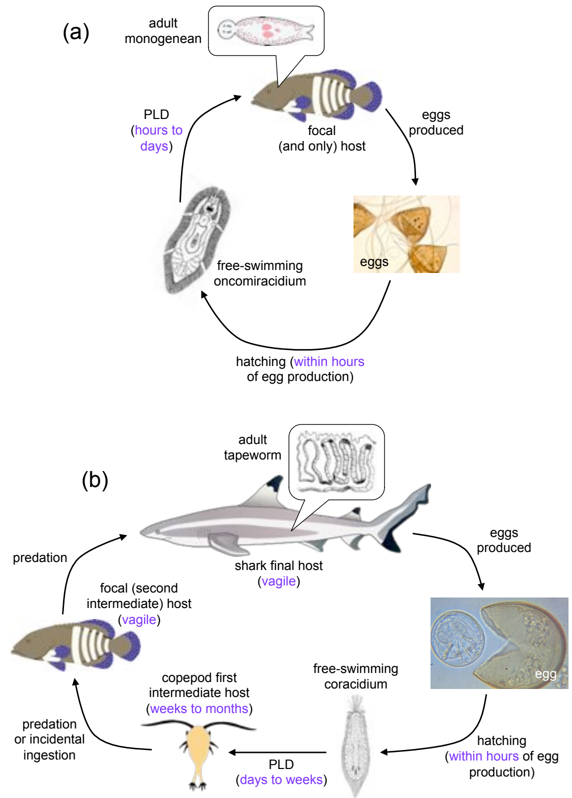


Figure 6. *Life cycles of (a) low dispersal-capacity monogenean parasites and (2) high dispersal-capacity tapeworm parasites. Factors that are likely to influence the distance traveled over the course of one life cycle are indicated in violet.*

Without considering parasite dispersal capacity, a few studies have investigated the role of scale in the diversity–disease relationship. In an analysis of three disease agents (amphibian chytrid fungus, Lyme disease, West Nile virus), Cohen et al. (2016) showed that the dilution effect was observable only at local scales (~1,000–10,000 km^2), while climate explained disease prevalence at larger scales (>100,000 km^2). This is consistent with the prediction from community ecology that biotic interactions should influence species distributions only at the small scales where those interactions manifest, whereas environmental conditions like climate should dominate at larger spatial scales (Huston 1999, Davies et al. 2005, Fridley et al. 2007). Using data on the distribution of amphibian parasites across >300 California ponds, Johnson et al. (2016) demonstrated that the slope of the relationship between host diversity and parasite diversity increased with increasing spatial grain, probably because the species–area curves of parasites saturate at a coarser spatial grain than do those of their hosts (Wood and Johnson 2016). Together, these existing studies suggest an important role for spatial scale in the diversity–disease relationship, but they are missing a key consideration: the diversity among parasites in their dispersal capabilities. In part, this may be due to the difficulty of quantifying the dispersal capacity of parasites (Nathan 2001), which usually cannot be marked and recaptured, GPS-tagged, or induced to disperse in a meso-cosm or other laboratory apparatus. One of the main advantages of using marine organisms in disease ecology studies is the extreme inter-specific variance in their dispersal capacity (Strathmann 1990), which can range from a few meters to thousands of kilometers; this variance in dispersal capacity can be estimated with molecular techniques. Parasite transmission is a spatial process in all ecosystems, but it is exceptionally tractable to study in the oceans.

Analyses – As in Question 2, for each parasite taxon a generalized linear mixed-effects model (GLMM) with negative binomial error structure will be used to assess whether there is a significant linear relationship between host diversity and parasite abundance. Also as in Question 2, models will include fixed effects of latitude (to control for covariance in hosts and parasites not related to their direct relationship, but instead related to some third variable correlated with latitude) and host body size (to control for positive effects of host size on parasite burden), and random effects of fish ID nested within site nested within island. This analysis will differ from that in Question 2 in that it will be repeated across spatial scales. To estimate both fish diversity and parasite abundance, will use a neighborhood-based approach that sequentially joins spatial units into progressively larger groups, tracking how spatial grain of observation (as a continuous variable) influences the diversity–disease relationship. The continuously varying spatial grain will begin at a finer resolution than that used in Question 2 (collection-site level) and end at a coarser resolution (island groups). We will then use meta-regression (as in Wood et al. 2014b) to explore the relationship between sensitivity of the diversity–disease relationship to spatial scale and dispersal capacity of each parasite taxon (**Figure 7b**).

Preliminary data and anticipated results – We predict that, at any given spatial grain, the slope of the diversity–disease relationship will weaken with increasing parasite dispersal capacity (**Figure 7a**). We also predict a hump-shaped relationship between slope of the diversity–disease relationship and spatial grain, where the peak of the hump increases with decreasing parasite dispersal capacity (**Figure 7b**). The hump-shaped relationship is expected because (*i*) fine spatial grains may reflect random heterogeneity in parasite abundance, particularly for high-dispersal parasites and (*ii*) parasite abundance at large spatial grains may be governed by large-scale forces (e.g., latitudinal productivity gradients) rather than local-scale forces, like host diversity. We expect that the peak of the hump will occur at coarser spatial grains for parasites with higher dispersal capacity, presumably peaking at a point where scale is coarse enough to reflect the large-scale dispersal process but fine enough that large-scale environmental factors (e.g., productivity) have not yet swamped the signal of dispersal. The peak will probably be lower for high-dispersal parasites, due to the swamping effect of large-scale factors (e.g., productivity) at the scales relevant for high parasite dispersal (**Figure 7b**).

Preliminary data from the six islands currently in our dataset indicate a negative relationship between parasite dispersal capacity and responsiveness to fishing pressure (which we assume here is a proxy for host biodiversity; **Figure 7c**). We used meta-regression to assess the responsiveness of parasite taxa to fishing pressure (standardized regression coefficients from Wood et al. [2014b]) as a function of parasite dispersal capacity and found that – consistent with our predictions – the greater the

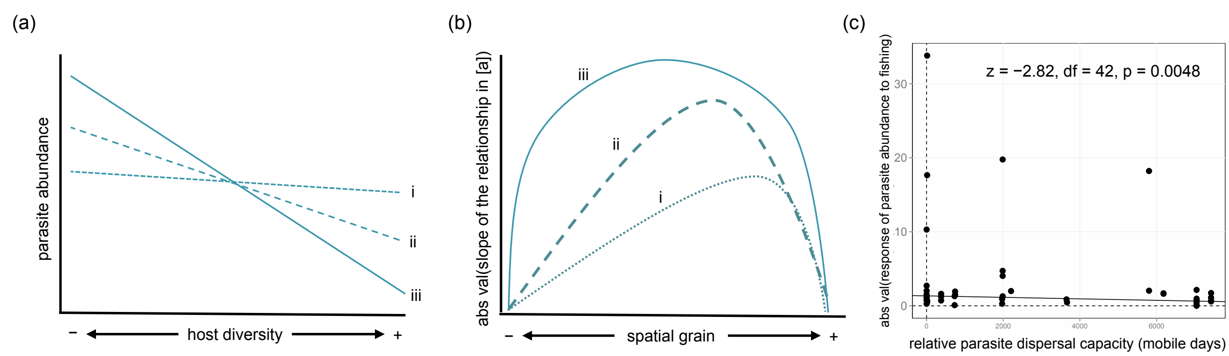


Figure 7. *We predict (a) that the slope of the relationship between host diversity and parasite abundance will shallow with increasing dispersal capacity of parasites and (b) that slope will be greatest in magnitude at intermediate spatial grains, with peaks at coarser spatial grains for higher dispersal parasites. Shown are [i] high-dispersal, [ii] medium-dispersal, and [iii] low-dispersal parasites. (c) We have already observed that the slope of the relationship between fishing and parasite abundance shallows with increasing parasite dispersal capacity.*

dispersal capacity of the parasite, the less it responded to island-level fishing pressure (**Figure 7c**). The next step in this analysis is to expand the spatial scope of our dataset, so we can explore how this relationship varies across spatial scales (**Figure 7b**) and to refine our estimates of parasite dispersal capacity using molecular techniques.

Estimating parasite dispersal capacity – We will obtain information about parasite dispersal capacity in two ways: first, by collating information on the life cycle and natural history of each parasite from the literature, and second, by using molecular data to estimate dispersal for a subset of parasite taxa.

We will begin by building a matrix of information on the life cycle and natural history of each identified parasite taxon (>47 taxa). This matrix will be dynamically linked to the database of parasite abundance, and will build on an existing dataset that was constructed for parasites from the six islands already sampled (see Appendix F in Wood et al. 2014b). Each parasite will be classified according to its higher-order taxonomic group, transmission strategy, the identity of all known intermediate and definitive hosts, and the presence of any free-living infectious stages. We will use the literature to estimate both the average distance over which each definitive and intermediate host might disperse over its lifetime and the length of the pelagic larval duration for any free-living infectious stages. This information will be collated and parasites will be ranked with respect to their dispersal capacity. Where parasite taxa are poorly known, we will surmise life history traits based on membership in higher-order taxonomic groups. While this is a coarse approach, many life history traits are phylogenetically conserved within these higher-order taxonomic groups (Sasal et al. 1998, Mouillot et al. 2006, Wood et al. 2015).

To complement this literature-based means of estimating relative dispersal capacity for all parasite taxa detected, we will also estimate absolute dispersal capacity using molecular techniques for a subset of parasite taxa. Because assessing genetic structure across all >47 parasite taxa would create unreasonable costs, we will choose six taxa that are common across the 18 islands: two taxa that have potential for high dispersal (e.g., nematodes, cestodes), two that are expected to have low dispersal (e.g., monogeneans), and two with intermediate dispersal capacities (e.g., crustaceans, trematodes). We will use tissue samples from ~25 genetically independent individuals of each of these six parasite taxa at each island to characterize overall genetic structure of each parasite population at the mtDNA COI locus. For each of the six parasite taxa, we will begin by running samples taken from distant islands and will add islands if there is genetic heterogeneity at these initial, coarse spatial scales; given the high dispersal potential of some of these parasites, we expect to run ~2300 samples in total, a number that we can reasonably expect six undergraduate technicians and Collaborator Haupt to complete in 13 weeks of lab work (see below). Genomic DNA will be extracted from individual parasites using Nucleospin extraction kits (BD Biosciences, San Jose, CA) and the mtDNA COI locus will be amplified through polymerase chain reaction. The mtDNA COI locus has been used in the past for assessing genetic connectivity and estimating dispersal among marine species (Haupt et al. 2013). Primers specific to each parasite taxon at the COI locus have been identified (Hansen et al. 2003, Hu and Gasser 2006, Yazawa et al. 2008, Razo-Mendivil et al. 2010, Waeschenbach et al. 2012, Makouloutou et al. 2014). After amplification, PCR products will be sent out for sequencing in both forward and reverse directions. Sequences will be aligned and visually scored in Sequencher (Gene Codes Corporation). These sequences will then be exported in FASTA files and analyzed using Arlequin (Excoffier et al. 2007) and R. We will calculate global, group, and pairwise *F*-statistics to estimate population subdivision, measures of genetic diversity, and haplotype networks. Pairwise *F*-statistics will be used to calculate isolation by distance (IBD) to estimate dispersal potential of species (Rousset 1997). We hypothesize that parasite taxa with potential for high dispersal will have little to no genetic structure and that taxa with low dispersal will possess high levels of genetic structure throughout the sampled range. Because of this, we might not be able to calculate dispersal distance for all six parasite taxa, and for those with homogeneous genetic structure across the study region, we will make the conservative assessment that dispersal is greater than or equal to the distance between the most distant sites. The results from this molecular analysis will be compared to expectations based on our literature survey (see paragraph above). The population genetic analyses will be led by Collaborator Haupt (California State University Monterey Bay; CSUMB) and a team of CSUMB undergraduates. The team's estimates of isolation by distance will become input for analyses designed to answer Question 3, but the data they produce will also answer other questions in the form of independent undergraduate research projects (see *Broader Impacts*). Collaborator Haupt will work closely with students from DNA extraction through data analyses and supervise them to ensure quality control throughout the project.

V. Broader impacts

Integrating research and education – Our proposed project will intimately intermingle education with research by placing a group of underrepresented undergraduates in a central research role: performing the molecular analyses required to estimate parasite dispersal distance (Question 3). Collaborator Haupt, an Assistant Professor at the minority-serving California State University, Monterey Bay (CSUMB), will lead a group of six undergraduates in a 13-week research experience, the Research Internship in Molecular Ecology (RIME). CSUMB is a federally classified Hispanic Serving Institution where Latinx students comprise 36% of the student body. The university serves a three-county region where nearly 50% of families are low-income, including many Hispanic migrant farm workers.

The RIME program will be administered with the logistical support, infrastructure, and instructional resources of CSUMB's Undergraduate Research Opportunities Center (UROC; see attached *Letter of Collaboration* from John Banks, UROC). UROC is a cross-campus center that trains, supports, and engages students in undergraduate research. RIME positions will be paid (see *Budget*) and will include no-cost campus housing provided by UROC, as financial support increases participation of students from low-income backgrounds (National Academy of Sciences 2011). The five CSUMB undergraduates to be supported by this grant will be selected through a competitive application process, and we will specifically recruit students from groups underrepresented in the sciences (i.e., women, individuals with disabilities, underrepresented minorities). We will leverage the CSUMB REU program to recruit a sixth student, who will be fully supported by the REU program, including research funds (~$1,000) and funds for conference attendance (see *Letter of Collaboration* from Corey Garza, CSUMB REU Program). These students will spend 13 weeks performing DNA sequencing of parasite vouchers in the Molecular Ecology Lab at CSUMB's Chapman Science Center. They will receive a one-week "crash course" in molecular ecology taught by Collaborator Haupt over spring break, and will begin working part-time (five hours/week) to optimize primers in the final five weeks of the spring semester, followed by seven weeks of part-time (30 hours/week) sequencing and analysis work and part-time (10 hours/week) professional development over the summer. Professional development activities will be administered by CSUMB's UROC/REU Summer Research Program and will include workshops on responsible conduct of research, presentation skills, writing skills, professionalism, and graduate school preparedness.

In addition to learning valuable laboratory and professional skills, students will each lead an independent research project that uses molecular data generated from parasite vouchers to address questions in parasite ecology (e.g., identifying barriers to dispersal, assessing variation in genetic structure among parasite taxa, estimating effective parasite population sizes). RIME students will receive direct one-on-one mentorship from Collaborator Haupt and either PI Wood or co-PI Sandin, as appropriate to the topic of the student's chosen independent project. The end of the internship will be timed to coincide with the annual in-person meeting of the collaborators (to take place in Monterey, CA in the final year of the grant), where RIME undergraduates will present and discuss their findings with postdocs, graduate students, and PIs from the Wood and Sandin Labs. All RIME students will also present a poster on their respective projects at the CSUMB Undergraduate Summer Research Poster Symposium. Students will receive authorship on all manuscripts that make use of the molecular data they produce and will be encouraged and supported to carry their research forward into undergraduate honors theses, to publish their work in peer-reviewed journals, and to present their work at the annual meeting of the Western Society of Naturalists in the fall semester after their internship ends. By integrating these students into the heart of our project, we will provide an unparalleled undergraduate research experience, one that allows students the opportunity to work with mentors from primarily undergraduate institutions (CSUMB) as well as large R1s (UW and Scripps/UC San Diego). Empirical data show that authentic research experiences like the one proposed here are high-impact practices for teaching (National Academy of Sciences 2011, McNair and Albertine 2012). We will work closely with UROC to assess the effectiveness of the RIME program by administering pre- and post-research experience surveys and comparing responses to a control group of CSUMB students. Outcomes to be assessed will include peer and faculty interaction, higher-order learning, academic gains, and professional development.

Our proposal will also support the training of two graduate students, two postdoctoral scholars, and several University of Washington undergraduate "capstone" (i.e., honors thesis) students. All students involved in the project will participate in project meetings and receive one-on-one mentorship (see *Postdoctoral Researcher Mentoring Plan*).

***Enhancing infrastructure for research and education* –** As our dataset continues to grow, it is imperative that we permanently archive it and make it publicly available. All of the data collected for this project – including fish data, parasite data, meta-data on sampling sites and hosts, accession info for sequences deposited to GenBank, and physical specimens – will be documented and made freely available through BCO-DMO (see *Data Management Plan*).

Facilitating participation of women, persons with disabilities, and underrepresented minorities in science – All of the new positions created by this proposal will be filled by women, persons with disabilities, or underrepresented minorities, where possible. PI Wood is currently conducting purposeful recruiting for such candidates by scouting for students at the annual meetings of the Western Society of Naturalists and the Society for Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS). In the selection process for the CSUMB RIME program, preference will be given to women, individuals with disabilities, and members of underrepresented minorities.

***Improving biodiversity science education and increasing public engagement with biodiversity science* –** We will develop and disseminate a vital resource that is in extremely short supply: quality educational tools for teaching about parasite biodiversity. In collaboration with the Network of Conservation Educators and Practitioners (NCEP), a program of the American Museum of Natural History's Center for Biodiversity and Conservation, we will develop an open-access learning module (see *Letter of Collaboration* from Ana Luz Porzecanski, AMNH). Like all modules developed by NCEP, ours will be peer-reviewed, designed for the university and professional level (but adaptable to other audiences), and will target educational outcomes central to conservation practice. These materials can be used by a diversity of educators, but are intended primarily for undergraduate faculty, especially early career faculty designing new courses. Parasites comprise more than half of Earth's species, but are rarely addressed in conservation science education (Nichols and Gómez 2010). Our module will explore the sometimes-hidden biodiversity of marine parasites, couching marine parasitology as the study of an unseen but vibrant world that operates just below the surface of familiar seascapes. Our module will consist of a Powerpoint introduction to the system, an exercise with active learning techniques, and teacher resources, designed to foster critical thinking skills and expand student understanding of parasite ecology by posing the question: what is the shape of the relationship between host and parasite biodiversity? Students will learn the competing hypotheses that address this question and use real data to discriminate between those hypotheses. Our efforts will make a high-quality educational resource widely available, with the broad international reach and high volume of web traffic generated by North America's premier museum of natural history. Assessment information will be gathered by tracking downloads of the module, distributing surveys to users of the module, and including a student pre- and post-assessment in module materials, which teachers can upload to our online system.

VI. Potential for transformative insights

The research proposed herein will provide the world's first insights into the direction, magnitude, and shape of the biodiversity–disease relationship across a diversity of parasite taxa, host taxa, and spatial scales, and will comprehensively identify conditions under which biodiversity is likely to be important in determining the abundance of disease agents – a fundamental contribution that can be readily extrapolated from our marine focal system to freshwater and terrestrial ecosystems. **No prior study has attempted such a comprehensive investigation of how biodiversity–disease relationships play out in real ecosystems**, and this will contribute substantially to a literature that currently relies on empirical evidence from a mere handful of parasite species. For marine ecosystems in particular, our work will indicate **when and where disease processes are likely to exert important ecosystem-level effects**. In addition to its unique suitability for its scientific purpose, this project leverages existing research infrastructure – the 100 Island Challenge – to perform field collections that would otherwise necessitate a multi-million dollar investment. Our efforts will not only produce major advancements in disease ecology and biological oceanography, but will also bring authentic research experiences to undergraduates from underrepresented groups, produce and disseminate active-learning modules on parasite biodiversity, and significantly enrich fundamental ecological theory.

VIII. Results from prior NSF support

PI Wood, co-PI Sandin, and Collaborator Haupt: No NSF funding in the past five years.