Evolving Coral Reef Conservation with Genetic Information

Maria Beger
Kimberly A. Selkoe
Eric Treml
Paul H. Barber
Sophie von der Heyden

See next page for additional authors

Follow this and additional works at: https://digitalcommons.csumb.edu/sns_fac

Recommended Citation
Authors
Maria Beger, Kimberly A. Selkoe, Eric Treml, Paul H. Barber, Sophie von der Heyden, Eric D. Crandall, Robert J. Toonen, and Cynthia Riginos

This article is available at Digital Commons @ CSUMB: https://digitalcommons.csumb.edu/sns_fac/9
Evolving coral reef conservation with genetic information

Maria Beger 1*
Kimberly A Selkoe 2,3
Eric Treml 4
Paul H Barber 5,6
Sophie von der Heyden 6
Eric D Crandall 7
Robert J Toonen 8
Cynthia Riginos 9

ABSTRACT.—Targeted conservation and management programs are crucial for mitigating anthropogenic threats to declining biodiversity. Although evolutionary processes underpin extant patterns of biodiversity, it is uncommon for resource managers to explicitly consider genetic data in conservation prioritization. Genetic information is inherently relevant to management because it describes genetic diversity, population connectedness, and evolutionary history; thereby typifying their behavioral traits, physiological climate tolerance, evolutionary potential, and dispersal ability. Incorporating genetic information into spatial conservation prioritization starts with reconciling the terminology and techniques used in genetics and conservation science. Genetic data vary widely in analyses and their interpretations can be challenging even for experienced geneticists. Therefore, identifying objectives, decision rules, and implementations in decision support tools specifically for management using genetic data is challenging. Here, we outline a framework for eight genetic system characteristics, their measurement, and how they could be incorporated in spatial conservation prioritization for two contrasting objectives: biodiversity preservation vs maintaining ecological function and sustainable use. We illustrate this framework with an example using data from *Tridacna crocea* (Lamarck, 1819) (boring giant clam) in the Coral Triangle. We find that many reefs highlighted as conservation priorities with genetic data based on genetic subregions, genetic diversity, genetic distinctness, and connectivity are not prioritized using standard practices. Moreover, different characteristics calculated from the same samples resulted in different spatial conservation priorities. Our results highlight that omitting genetic information from conservation decisions may fail to adequately represent processes regulating biodiversity, but that conservation objectives related to the choice of genetic system characteristics require careful consideration.
In our world of declining natural resources and biodiversity (Burke et al. 2011, Barnosky et al. 2012), targeted conservation and management programs are crucial for mitigating anthropogenic threats (Hughes et al. 2010, Pandolfi et al. 2011). A large body of literature deals with conservation prioritization methods, the effect of different data types on outcomes, and how to decide on conservation strategies under high uncertainty (Margules and Pressey 2000, Rondinini et al. 2006, Regan et al. 2009). Similarly, much has been written about how to spatially delineate populations and conservation units using genetic data (Moritz 1994, Mace et al. 2003, Waples and Gaggiotti 2006, Palsbøll et al. 2007, Waples et al. 2008, Funk et al. 2012). However, it is relatively uncommon for resource managers to explicitly consider genetic data in conservation prioritization processes (Laikre et al. 2010, von der Heyden et al. 2014). Such genetic data is not overlooked without cause, see Waples et al. (2008) for a discussion of scientific and institutional issues that have hampered this integration. This is surprising, because evolutionary processes underpin extant patterns of biodiversity. Descriptions of genetic diversity, population connectivity, and evolutionary history are inherently spatial, and are therefore reasonable to consider in spatial planning (Faith 1992, Crozier 1997, Bowen 1999). For example, the genetic makeup of organisms largely determines their vulnerability to global change by typifying their behavioral traits, physiological climate tolerance, evolutionary potential, and dispersal ability (Crozier 1997, Carvalho et al. 2010). Furthermore, observed genetic diversity can provide information about a species’ long-term persistence and may also indicate adaptive potential or evolutionary resilience (Stillman 2003, Sgro et al. 2011). Genetic methods can provide this type of information, which is otherwise not available in the conservation decision-making toolbox (Bowen 1999). Here we provide a guiding framework linking information gained from population genetics with spatial conservation planning.

Spatial conservation prioritization aims to identify suites of locations for protection or regulation, considering biodiversity, socio-economic needs, future environmental change, and other sources of uncertainty (Regan et al. 2009, Wilson et al. 2009). These choices are increasingly made through structured, evidence-based decision making, where the explicit formulation of conservation objectives allows the efficient and transparent integration of often conflicting factors, e.g., biological representativeness and cost (Wilson et al. 2009), and connectivity (Beger et al. 2010a). Many sophisticated approaches to address conservation objectives exist, often facilitated by decision support tools, such as the “minimum set” approach aiming to represent a set of conservation targets (e.g., 30% of each habitat) while minimizing the resources spent (Fernandes et al. 2005, Green et al. 2009, Wilson et al. 2009). This decision process relies on the spatial representation of biodiversity and factors that affect biodiversity such as recruitment, competition, environmental niches of species, and local environmental conditions. These factors are often inferred from habitat maps and modeled species distributions across a set of planning units that provide spatial standardization of prioritized sites (Hannah et al. 2007, Beger and Possingham 2008). To date, the use of genetic information in spatial planning systems has been limited to a few examples. Diniz and Telles (2006) proposed an algorithm to develop reserve systems incorporating spatial autocorrelation patterns of intra-species genetic variability. Spatial reserve systems designed with evolutionary significant units were shown to be more effective in protecting rare endemics than designs using species (Vasconcelos et al. 2012). Conservation prioritization with
species richness alone fails to represent phylogenetic diversity, indicating a need to include genetic measures to better represent biodiversity (Karl and Bowen 2001, Pio et al. 2011). Other studies, many terrestrial-based, have highlighted how evolutionary processes may inform conservation prioritization using phylogenetic diversity (Faith and Walker 1992, Forest et al. 2007, Pio et al. 2011, Diniz et al. 2012), evolutionary management units (Moritz 2002, Barber et al. 2011, Toonen et al. 2011), phylogeography (Moritz and Faith 1998, Carpenter et al. 2011), evolutionary innovation (Davis et al. 2007), and surrogate data for genetic diversity (Carvalho et al. 2010), but their implementation in spatial decision making is lacking. Nonetheless, our understanding of what set of objectives is appropriate across the varying spatial and temporal scales of genetic data types is poorly developed (Mace et al. 2003, Carvalho et al. 2011, Diniz et al. 2012).

Population genetic studies (including phylogeography) can potentially inform minimum management units (Moritz 1994, Palsbøll et al. 2007, Toonen et al. 2011, Funk et al. 2012), fisheries stock size and boundaries (Hutchinson et al. 2001, Hauser et al. 2002, von der Heyden et al. 2014), and provide benchmarks for the success of management (Uthicke et al. 2004). These and other examples of how population genetic data contributes to conservation and management focus on the scale and intensity of the spatial structuring of neutral genetic diversity, that is, genetic differentiation or genetic structure. Such genetic structure partly reflects the scale and amount of gene flow and therefore can indicate population connectivity of management areas and likely source populations for replenishment following disturbances (Hedgecock et al. 2007, Lowe and Allendorf 2010, Selkoe and Toonen 2011). Of course, genetic structuring may also be affected by differences in effective population size, demographic or colonization history, natural selection, or some combination of these factors, especially for populations that may not have reached migration drift equilibrium. Thus, direct interpretation of population structure in the context of gene flow can sometimes be problematic (Whitlock and McCauley 1999, Hart and Marko 2010, Lowe and Allendorf 2010, Marko and Hart 2011, Karl et al. 2012). However, genetic isolation arising from a lack of gene flow determines the capacity for divergence and independent evolution of populations, and taken to an extreme, can ultimately facilitate speciation (Coyne and Orr 2004) and local adaptation (Sotka 2005, Sanford and Kelly 2011). Genetic tools have also played a major role in discovering cryptic or incipient species and hybrid zones which contribute to spatial patterns of biodiversity, but which often go unnoticed (Knowlton 2000, Bickford et al. 2007, Bird et al. 2011).

Incorporating population genetics into spatial conservation prioritization requires the definition and spatial delineation of genetically distinct populations of randomly-mating individuals within the species of interest (i.e., the population genetic structure). For populations to be genetically distinct, the number of effective migrants exchanged between populations per generation must typically be quite low (generally ≤5 effective migrants for detectable population structure to develop and certainly ≤25) (Mills and Allendorf 1996, Waples and Gaggiotti 2006). This genetic definition of populations contrasts with the broader definition of populations commonly used in conservation science, which often refers to populations as spatially isolated subunits (i.e., biogeographic localities), where genetic distinctness is unknown. In conservation planning, each planning unit could be viewed as a separate "subpopulation," because one of the underlying assumptions of choosing the size of planning units is that the population can persist for some time even if all habitat outside the reserve is
destroyed (Murray et al. 1999, Fernandes et al. 2005). Here, we use the genetic definition of populations as “groups of similar individuals whose genetic make-up is statistically different from other such groups.” Delineating populations can be a tricky task even for terrestrial and freshwater species, which often have smaller population sizes than marine species and thus stronger genetic differences between populations (due to faster genetic drift). For marine species with dispersive larvae that are capable of travelling hundreds of kilometers on ocean currents, the large effective population sizes slow the rate of genetic drift and depress traditional measures of population genetic structure such as $F_{st}$ (Waples 1998, Neigel 2002, Hedrick 2005, Hellberg 2009). Furthermore, when habitat is large and continuous, spatial genetic change may also be continuous, precluding delineation of individuals into discrete groups. However, in many such cases an isolation-by-distance analysis (Rousset 2004) can inform the spatial scale of demographic processes for species with continuous spatial genetic variation, allowing genetically informed choice of planning units (Puebla et al. 2008, Ackiss et al. 2013, Crandall et al. 2014).

Coral reefs are among the most iconic marine habitats and are also one of the most threatened ecosystems in the world (Hughes et al. 2010, Burke et al. 2011), putting these habitats and their associated benefits to mankind in jeopardy (Harley et al. 2006) if no interventions slow or reverse the decline (Mumby and Steneck 2008, Hughes et al. 2010). Common interventions focus on reducing the compounding stresses to the ecosystem, such as closing an area to fishing to enhance population recovery and maintain healthy source populations or reducing non-point source pollution causing nutrient enrichment (McLeod et al. 2009, De’ath et al. 2012). Although most conservation planning efforts in marine environments target ecological timescales, the timescale of evolution is much shorter than is commonly appreciated (Hendry and Kinnison 1999, Palumbi 2001, Stillman 2003, Schoener 2011), and significant anthropogenic effects (e.g., on maturation rate or size) have already been detected on the evolutionary trajectory of many marine species (Devine et al. 2012, Olsen et al. 2004). Even at deeper timescales, evolutionary patterns (e.g., Evolutionarily Significant Units) and processes (e.g., local adaptation) underpin ecological responses, and are thus relevant for conservation decisions (Bowen 1998, Briggs 2005, Rocha et al. 2007, Budd and Pandolfi 2010).

The Indo-Pacific region contains the most biodiverse as well as some of the most threatened coral reef systems worldwide (Veron et al. 2009, Burke et al. 2011, Allen and Erdmann 2012). Within the Indo-Pacific, there are several high-profile conservation initiatives for coral reefs such as the Micronesia Challenge and the Coral Triangle Initiative as well as local spatial prioritization efforts (Green et al. 2009, Baker et al. 2011, Game et al. 2011, Grantham et al. 2013). These on-the-ground programs aim to place large areas of reef under protection, using criteria such as biodiversity, sustainable fisheries, food security, avoiding climate change threats, and integrating land-based influences (West and Salm 2003, CTI Secretariat 2009, McLeod et al. 2009). This region is also notable for many deep genetic divisions over relatively short distances (Barber et al. 2002, 2006, 2011, Carpenter et al. 2011, Martí-Puig et al. 2014). These genetic discontinuities are consistent with regionally isolated populations, such that ecological connectivity among regions is likely to be low for many species.

Incorporating genetic information into spatial conservation prioritization starts with reconciling the terminology and techniques used in the fields of genetics and
conservation science. Here we build a framework that connects genetic measures, their ecological meaning, and their corresponding use in conservation planning. We illustrate the theoretical framework with an example using data from the boring giant clam, *Tridacna crocea* (Lamarck, 1819), from Philippine and Indonesian coral reefs in the Coral Triangle, to highlight how the reef conservation priorities may change when different representations of evolutionary processes are considered, as compared to the standard practices where evolutionary patterns and processes are not used.

**Genetic Decision Rules**

The first step in conservation prioritization is to develop objectives and decision rules to guide analyses. Such design principles are prolific in the literature and in managers’ guidebooks. In general, they deal with conservation features (e.g., things to protect; Baker et al. 2011, Gilman et al. 2011), thresholds of amounts of features to protect (Fernandes et al. 2005), targets (e.g., what percentage of features to protect; Carwardine et al. 2009, Baker et al. 2011), and replication (Green et al. 2009). Refinements of general design principles may include specific threats such as climate change (Fernandes et al. 2012), or predicted states such as risk level, persistence probabilities or information uncertainty (Game et al. 2008, Foley et al. 2010). These decision rules aim to protect biodiversity representatively, comprehensively, adequately, and cost efficiently by trading off competing priorities represented as “cost layers” (Wilson et al. 2009). Evolutionary data relate to these goals in several ways. Firstly, molecular techniques can identify hidden biodiversity, such as populations with high genetic diversity, cryptic species, or isolated populations that should effectively be treated as separate entities. Conversely, some nominal species may turn out to be genetically indistinguishable ecotypes which, while potentially informative about adaptive variation, can bring into question the justification for independent management strategies (Forsman et al. 2010). Secondly, the evolutionary processes that fundamentally underpin the existence and functioning of biodiversity ideally should be captured in conservation networks to adequately conserve diversity. For example, dispersal between populations may be critical for the replenishment of individual populations (in a meta-population framework) and may also maximize adaptive potential by enhancing genetic diversity in that population.

Genetic data take many forms and the variety and nuances of population genetic analyses and their interpretations can be challenging even for experienced scientists in this field (Whitlock and McCauley 1999, Lowe and Allendorf 2010, Karl et al. 2012). For these reasons, decision rules for using genetic data in conservation are not straightforward and even the objectives might be debatable. For example, should we represent the full range of genotypes equally or maximize genetic diversity by prioritizing sites with high genetic diversity over others? Or instead, prioritize isolated populations because of their unique genetic make-up? Should we protect contact zones where distinct lineages mix to preserve hybrid integration areas as potential adaptive zones (Seehausien 2004, Moritz et al. 2009), or should we focus on protecting the core ranges of these lineages and ignore hybrid zones?

These decisions depend on the overall conservation objective, the basis of all objective-driven decision-making (Gerber et al. 2007, 2011). In coral reef spatial planning, for example, different decision rules would be applied if the objective was to protect
biodiversity (encompassing all levels of diversity such as habitats, species, and genetics), or if the objective was to maintain function (management to ensure food security or coastal protection goals) such as to support a sustainable fishery (Baums 2008). In the first, conservation responses need to incorporate the whole range of diversity types across the range of a population equally, focus on core ranges of genetic lineages, and ensure cryptic and rare species are represented in conservation site networks. To achieve the second, zones with high gene flow have high conservation priority because sources of demographic and evolutionary scale connectivity should be maintained. Following these examples, different system characteristics measured with genetic tools may invoke different conservation responses, depending on the objective (Table 1).

Genetic metrics may be site specific or characterize relationships between two or more sites, but are estimated from the same empirical data. For instance, the types and frequencies of alleles are estimated for populations approximating a specific georeferenced location and each population can be described by its genetic diversity, uniqueness, and distinctiveness (Tables 1, 2), but the aggregation of this allelic and genotypic diversity also informs the relationships among these geo-referenced populations (analogous to concepts of alpha and beta diversity based on species distributions, Petit et al. 2008, Diniz et al. 2012). Both types of information (i.e., diversity at individual locations and relationships between locations) may be incorporated in systematic prioritization but will require modeling or interpolation techniques to map the data (Beger and Possingham 2008, Kininmonth et al. 2010, Carvalho et al. 2011, Pio et al. 2011), with the caveat that spatio-paleoecological modeling of species distributions that best fit deep genetic patterns requires spatial paleoclimate datasets (Espindola et al. 2012) that are rare in marine environments. Descriptive genetic statistics are relatively easily calculated; but where a conservation objective is to preserve ecological function (dispersal connectivity), more quantitative parameters like gene flow may be relevant. Estimates of gene flow between pairs of sites (sometimes asymmetrical depending on the method) are more difficult to obtain as they involve specialized software and may be highly sensitive to model parameters and assumptions (Whitlock and McCauley 1999, Hart and Marko 2010, Lowe and Allendorf 2010, Marko and Hart 2011, Karl et al. 2012). Furthermore, decision rules for connectivity in spatial planning are poorly developed and depend on conservation objectives. Most commonly, the aim is to protect biodiversity by enhancing overall dispersal connectivity in marine reserve networks (Beger et al. 2010b, Treml and Halpin 2012), and this connectivity can be represented by measures of demographic or evolutionary scale connectivity. In cases where maintaining function is more important, or where conservation budgets limit the number of reserves substantially, strong and persistent source populations should be given highest emphasis (Table 1).

**Which Molecular Techniques Measure Which System Attributes?**

Viable genetically-informed conservation objectives are listed in Table 1. These objectives are purposefully presented as general principles without assuming advanced population genetic knowledge on the part of the reader. Likewise, we are not trying to discuss the nuances of the population genetic data sets or the methods for estimating or interpreting them, but rather focus on a mechanism of how to incorporate genetic data into spatial conservation prioritization. The implementation of these
objectives requires statistical estimates from empirical genetic data. Table 2 defines relevant genetic metrics and lists some of the commonly employed approaches for obtaining these estimators. There are a multitude of approaches for estimating any of these attributes. Here we concentrate on some of the most commonly reported statistics, which could be extracted for spatial planning from most journal publications describing spatial population genetics or phylogeography of a species. Detailed comparisons of methods can be found in recent reviews (Holsinger and Weir 2009, Saenz-Agudelo et al. 2009, Hart and Marko 2010, Lowe and Allendorf 2010, Bird et al. 2011, Ho and Shapiro 2011, Marko and Hart 2011, Karl et al. 2012) and programs for undertaking many of these methods are reviewed by Excoffier and Heckel (2006). Our goal at present is not to advocate the utility of any specific estimator for conservation purposes or to enumerate the nuances of applying such estimators, but simply to illustrate how such metrics could be incorporated in a spatial prioritization methodology. To that end, we have selected proposed genetic conservation objectives (and associated metrics) that typify the most common analyses conducted on single species. The degree to which genetic properties of a single species are representative of other species in the same community is unclear (Kelly and Palumbi 2010, Carpenter et al. 2011, Toonen et al. 2011, Bailey et al. 2012, Whitham et al. 2012). For example, rocky shore fishes and limpets each displayed highly different evolutionary patterns despite being closely related (Bird et al. 2007, von der Heyden et al. 2013); in contrast, broad geographic scale delineations between evolutionarily distinct populations frequently co-occur (Fortuna et al. 2009, Barber et al. 2011, Carpenter et al. 2011, Jackson et al. 2014). Regardless of whether a single species can represent a community, establishing methodologies based on the simple case of a single species is an obvious precursor to more complex scenarios where genetic information from multiple species could be combined.

Changing Conservation Priorities with Genetic Data in the Coral Triangle

Based on a data set for the boring clam *T. crocea* from Indonesia and the Philippines (DeBoer et al. 2014), we evaluate how spatial conservation priorities change when different types of genetically measured system attributes are incorporated in decision making, and compare them to outcomes based solely on optimizing for habitat representation (Fig. 1). While single species approaches are not common in spatial conservation prioritization; here, using a single species exemplar serves to illustrate the approach and to unambiguously evaluate the influence of different genetic system attributes.

Genetic Data.—Inputs for the prioritization analysis consisted of 524 individuals genotyped at 8 microsatellite loci from 27 populations across Indonesia and the Philippines (Online Fig. S1, Online Table S1) with methods described in DeBoer and Barber (2010). We chose to focus on four genetic descriptors that capture different aspects of the data and represent a subset of the genetic system attributes that might be selected based on the conservation objectives (Table 1). The goal of this exercise is to illustrate how such attributes could be used and we have selected representative estimators (Table 2). First, genetic sub-regions were determined by site clustering inferred with Structure (Pritchard et al. 2000). This assignment method
Table 1. Decision rules based on genetic system attributes for two contrasting conservation objectives, to conserve biodiversity (representing habitat, species and genetic diversity) or to maintain function (managing for sustainable use and basic ecological function). For technical details on tools to derive these system attributes, see Table 2.

<table>
<thead>
<tr>
<th>System attribute</th>
<th>Scale</th>
<th>Interpretation</th>
<th>Conservation response</th>
<th>Representation in decision support software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic diversity of sites</td>
<td>Any spatial scale</td>
<td>• Highly diverse sites may have better adaptive potential and more source populations, making them more resilient to environmental change; unusually low genetic diversity may indicate low resilience and, in the extreme, inbreeding depression can lower fitness of individuals.</td>
<td>• Include representative sample of genetic variants (alleles) across sites to preserve the current spatial pattern of diversity.</td>
<td>• Cost layer inversely proportional to genetic diversity.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Sites with rare or “private” alleles (genetic variants found only at that site) harbor potentially unique genetic types. These geographically restricted alleles are likely to coincide with isolation of this population from populations from other sites and may also indicate local adaptation in this population.</td>
<td>• Prioritize alleles from loci that are assumed or confirmed to be selectively neutral, and loci showing diversifying selection according to an “F_{ST} outlier” test can be considered in this context.</td>
<td>• Assign specific targets to genetically unique sites, or ignore such sites.</td>
</tr>
<tr>
<td>Genetic uniqueness of sites</td>
<td>Any spatial scale</td>
<td>• Alleles from loci are assumed or confirmed to be selectively neutral, and loci showing diversifying selection according to an “F_{ST} outlier” test can be considered in this context.</td>
<td>• Ensure sites with unique alleles are protected.</td>
<td>• Assign specific targets to genetically unique sites, or ignore such sites.</td>
</tr>
<tr>
<td>Genetic distinctness of sites</td>
<td>Any spatial scale</td>
<td>• Some sites may contain alleles in unusual proportions/frequencies, indicating they may be isolated and/or have unusual histories; sites that are not genetically distinct are likely to be evolutionarily well connected to other sites.</td>
<td>• Represent present spatial pattern of high and low values of genetic distinctiveness.</td>
<td>• Assign specific targets to different levels of genetic distinctiveness.</td>
</tr>
<tr>
<td>Genetically distinct sets of sites</td>
<td>Global to regional scale</td>
<td>• If genetically distinct clusters of sites are detected, then the exchange of individuals among these distinct clusters has been extremely rare for 10s to 1000s of generations.</td>
<td>• Ensure conservation areas are replicated/distributed across genetically distinct clusters of sites.</td>
<td>• Create spatial representation of genetic subregions and set specific targets for each of them (e.g., stratify).</td>
</tr>
<tr>
<td>System attribute</td>
<td>Scale</td>
<td>Interpretation</td>
<td>Conservation response</td>
<td>Representation in decision support software</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Historical rates of gene flow among sites (direction of flow sometimes available)</td>
<td>Global to regional</td>
<td>• Where gene flow is low, it can be safely assumed that sites are not well connected in the present day. However, inferred high gene flow and/or low differentiation may indicate a historical long term average and are not necessarily indicative of present day connections. Analysis may be able to reveal source strength and identify sink populations.</td>
<td>• If gene flow is low, sites are likely to be genetically distinct. See guidelines above.</td>
<td>• Include pairwise flow rates as connectivity strength between sites. • Represent connection strength as penalty cost where only one site of a pair of connected sites is in a reserve system, e.g. connectivity/boundary cost.</td>
</tr>
<tr>
<td>Recent migration rates among sites (directional)</td>
<td>Regional to local</td>
<td>• Contemporary dispersal events can estimate contemporary migration rates and sometimes identify recent source and destination populations. Analysis may be able to reveal source strength and identify sink populations.</td>
<td>• If gene flow is low, sets of sites are likely to be genetically distinct. See guidelines above.</td>
<td>• Include pairwise flow rates as connectivity strength between sites. • Assign specific targets to particularly strong sources. • Assign specific targets to sinks if they are believed to be genetic repositories.</td>
</tr>
<tr>
<td>Self-recruitment rates of sites</td>
<td>Local</td>
<td>• Highly self-recruiting populations may be resilient to geographically distant events yet vulnerable to local catastrophes (i.e., may be demographically isolated). Levels of self-recruitment may vary greatly year-to-year.</td>
<td>• Preserve a balance of sites with different levels of self-recruitment.</td>
<td>• Ignore or assign higher importance to high self-recruitment sites.</td>
</tr>
<tr>
<td>Hybrid zones</td>
<td>Global to regional</td>
<td>• If hybrid individuals are generally less fit than either parental species, then hybrid zones may be evolutionary sinks. However, recent theoretical and empirical studies have suggested that evolution can proceed quickly in hybrid zones due to the unique genetic combinations found in hybrid individuals.</td>
<td>• Preserve a balance of sites including those with parental species and hybrids.</td>
<td>• Delineate genetic zones that includes hybrid zones and set targets for each zone • Exclude hybrid zones from species’ core ranges when considering their occurrence.</td>
</tr>
</tbody>
</table>
Table 2. Methods of estimating system attributes using genetic data.

<table>
<thead>
<tr>
<th>System attribute</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic diversity of site</td>
<td>Genetic diversity typically correlates with genetic effective population size ( (N_e) ) (Charlesworth 2009, Hare et al. 2011). The most common estimates of genetic diversity are based on (1) expected heterozygosity ( (H) ), the chance of that any two alleles drawn from the population are different (Nei 1973), or (2) allelic richness ( (AR) ), the number of alleles per locus. Mutation rate also influences diversity so comparisons of diversity across sites must be made with the same marker(s). AR may give better indications of adaptive potential (Allendorf 1986), and should always be first rarefied (subsampled) to standardize for sample size variation (Leberg 2002). Tests for outlier loci showing signs of diversifying selection can improve indications of adaptive potential (Foll and Gaggiotti 2008).</td>
</tr>
<tr>
<td>Genetic uniqueness of sites</td>
<td>Uniqueness can be estimated as the percentage of rare or private alleles ( (PA) ). PA are alleles found in no other sites, and are therefore analogous to range restricted species. In practice, a threshold of rarity might be used, (for example, frequency less than 10%), instead of focusing only on private alleles, which are sensitive to sampling design.</td>
</tr>
<tr>
<td>Genetic distinctness of sites</td>
<td>Distinctiveness indicates how far the genetic composition of each site deviates from the mean genetic composition of all sites, similar in concept to analyses of variance [node position in a population graph: Dyer and Nason (2004), local ( F_{st} ): Foll and Gaggiotti (2006), or concepts of beta diversity; differentiation component of diversity: Petit et al. (2008), turnover: Diniz et al. (2012)].</td>
</tr>
<tr>
<td>Genetically distinct sets of sites (genetic subregions)</td>
<td>Many approaches can estimate which subsets of sites form distinct clusters. Methods draw upon allele frequencies [such as AMOVA: Excoffier et al. (1992), SAMOVA: Dupanloup et al. (2002), BAPS: Corander et al. (2008)] or correlations among loci [such as Structure: Pritchard et al. (2000), GENELAND: Guillot et al. (2005) and Population Graphs: Dyer and Nason (2004)]. These and additional relevant implementations for estimating groupings of sites are listed in Excoffier and Heckel (2006).</td>
</tr>
<tr>
<td>Historical rates of gene flow among sites</td>
<td>Historical rates of gene flow can be estimated using coalescent approaches [Migrate: Beerli and Felsenstein (2001), IMa: Hey and Nielsen (2007)]. Traditionally many investigators have used the inverse of ( F_{st} ) as an estimate of gene flow, this expectation, however, relies on an idealized “island model” whose assumption are unlikely to be met for most populations Whitlock and McCauley (1999).</td>
</tr>
<tr>
<td>Recent migration rates among sites</td>
<td>Assignment testing programs [such as BayesAss+: Wilson and Rannala (2003), Structure: Pritchard et al. (2000), BIMr: Faubet and Gaggiotti (2008)] use the multilocus genotype of each individual to “assign” it back to the population of origin, whereby migrant individuals are those whose genotype is best assigned to a different population from that which they live in. These programs work best when migration rates are low and populations are highly distinct (Berry et al. 2004, Faubet and Gaggiotti 2008).</td>
</tr>
<tr>
<td>Self-recruitment rates of sites</td>
<td>Parentage Analysis based on DNA fingerprinting can provide several demographic metrics such as realized dispersal distance and source-destination relationships, depending on study design, especially if combined with assignment tests (Brazeau et al. 2005, Saenz-Agudelo et al. 2009).</td>
</tr>
<tr>
<td>Hybrid zones</td>
<td>Hybrid populations are typically recognized as harboring elevated genetic diversity, with atypical combinations of alleles found in some individuals. In practice, distinguishing between sympatric cryptic species and hybridization can be challenging (Barton and Hewitt 1985, 1989).</td>
</tr>
</tbody>
</table>

clusters individuals to minimize deviations from Hardy-Weinberg equilibrium allele frequencies (the frequencies expected for a randomly mating, idealized population) within each cluster and thereby delineates emergent genetic groups. The three resultant geographic clusters described considerable partitioning of genotypic diversity \( (F_{CT} = 0.067, P < 0.00001) \) with three groups corresponding to the Indian Ocean,
central Indonesia and Philippines, and the Bay of Cenderawasih (DeBoer et al. 2014). Second, allelic richness was estimated in Fstat version 2.9 (Goudet 1995) with rarefaction based on a sample size of 12. Allelic richness is a measure of the diversity of alleles standardized for sample size, and is a good reflection of genetic population size where greater population size can indicate long term stability (persistence) (Allendorf 1986) (see Table 1). Third, the data were run through the program GESTE (Foll and Gaggiotti 2006) to get estimates of local $F_{ST}$ for each population, using default parameter value settings. Local $F_{ST}$ describes the distinctiveness of a population in how far that population’s genetic composition differs from the mean of other sampled populations. Fourth, we generated estimates of recent, one-way migration rates between the 22 populations within the largest, central cluster identified by Structure using BayesAss+ (Wilson and Rannala 2003) with ten replicate runs and default parameter value settings.

This method detects recent migration among populations by capitalizing upon multilocus genotypes to infer directional migration rates from inbreeding coefficients without assuming Hardy-Weinberg equilibrium or other restrictive assumptions. As the genetic subregions were delineated under the assumption of limited recent genetic exchange, the calculation of recent migration rates only applied within each sub-region. The number of populations in the Indian Ocean and those in Cenderawasih Bay was not sufficient to conduct this analysis.

**Conservation Planning.**—We developed spatial conservation prioritization scenarios that include baseline conservation features of five reef types from an
unsupervised classification of satellite imagery (Kakuta et al. 2010). For the reason that the aim of prioritization was to establish no-take conservation areas on coral reefs, we used an estimate of the amount of artisanal fishing (Halpern et al. 2008, http://www.nceas.ucsb.edu/globalmarine/impacts), to represent the cost of reserves as lost opportunity for fishers. Our analyses applied the conservation planning software Marxan, a widely used and freely available program (http://www.biology.uq.edu.au/marxan). Marxan finds solutions to an objective function aiming to minimize the cost of an overall reserve system while meeting the targets set for conservation features. We set the target to represent 30% of each baseline conservation feature equally (see Table 1 for features). These parameters for baseline conservation features and cost remained the same across all scenarios, and in each scenario we added parameters representing genetic system characteristics in the context of the biodiversity conservation objective (compared in Table 1). As many different spatial configurations of solutions may fulfill the objective function, each scenario included 100 repeat runs to ensure that system variability is captured. We established five scenarios to evaluate the differences between using different objectives, and different system characteristics (Table 3).

Spatial conservation decision systems require continuous spatially explicit data surfaces, and point data associated with genetic sampling need to be interpolated throughout the entire planning region. This requirement constitutes a major obstacle to using genetic point data in spatial conservation prioritization, as it is unclear how genetic and evolutionary system characteristics could be modeled using environmental parameters in the marine environment, but see Carvalho et al. (2010) for a terrestrial example.

The entire seascape was discretized into squares (30 × 30 km) and those containing reef habitat were used as planning units in the spatial prioritization (Online Fig. S1). Values for genetic system attributes were assigned to these planning units based on location within a subregion (structure) and the values assigned by interpolations (allelic richness, Local $F_{st}$). We used a resampling procedure in ArcGIS (Zoraster 2003) to interpolate values for genetic diversity (allelic richness) and genetic distinctness of sites (Local $F_{st}$) that represents a simplified representation of the patterns (Online Fig. S2). From interpolated values, five classes were defined and targets were set for each class (Table 3). Recent migration rates required a more complicated procedure, because flows are not site specific, and need to be assigned between pairs of planning units in Marxan (Beger et al. 2010b). To represent this directional connection strength in Marxan, we made the simplifying assumption that each sampling site is representative of the surrounding seascape and therefore the pairwise migration rates can be applied to the proximate neighborhood. We defined the local seascape neighborhoods represented by each sampling site based on Thiessen polygons and applied the pairwise migration data to the planning units within these neighborhoods (Online Fig. S3). This method effectively extrapolated the genetic-based connectivity estimates (27 sites) to all 1449 planning units used in the analysis. This planning unit connectivity matrix was used as the connectivity strength matrix in Marxan (Beger et al. 2010b).

**Resulting Changes in Conservation Priorities.**—The baseline scenario using only habitat representation as an optimization criterion captured 30% of all habitat types across the region by selecting reserve networks where most sites are in
the Java Sea around southern Sulawesi and on the periphery of Palawan, Philippines (Online Fig. S4). This pattern was driven mostly by the cost data. When omitting the cost data, planning units were selected at equal frequencies of approximately 30%. Reef habitats and species assemblages varied widely across the region (Edinger et al. 2000, Meyer et al. 2005, Veron et al. 2009), but the classification of habitats used here did not capture this variability. Nevertheless, this caveat remained the same across all scenarios. When adding genetic data, conservation priorities changed for all genetic system attributes, but differently. For example, applying a 30% target for each of three subregions to represent structure (Indian Ocean, Central Indo-Pacific and Cenderawasih Bay) shifted some conservation focus to Cenderawasih Bay and the Indian Ocean sites that were poorly represented in the baseline scenario (Fig. 2A). Genetic diversity (allelic richness) and genetic distinctness (Local $F_{ST}$) changed conservation priorities in almost identical ways (Fig. 2B,C), spreading conservation priority sites more evenly around the planning region, and assigning higher priorities to sites in the Central Philippines, the Sangihe Talaud Archipelago, and Indian Ocean sites on Sumatra. In contrast, the recent migration rates scenario changed conservation priorities predominantly by dotting single priority sites evenly across the seascape (Fig. 2D).

**Discussion**

In the present study, we present a framework for how to incorporate genetic design principles for conservation prioritization that bridges the fields of population genetics, conservation genetics, and spatial conservation decision science. Our approach catalogued genetic system attributes at different spatial and temporal scales and includes site-specific and between-site parameters (Tables 1, 2). Given the multifaceted nature of genetic data, these system attributes require different approaches depending on the conservation objectives (e.g., protect biodiversity or ensure functionality) (Table 1).
In our Coral Triangle case study, conservation priorities changed when adding genetic data for the boring giant clam *T. crocea*. We use this case study to illustrate that omitting genetic data from basic spatial marine prioritization is unlikely to fully represent ecological and evolutionary processes, particularly for highly structured organisms such as *T. crocea* (DeBoer et al. 2014). Whether *T. crocea* is representative of genetic diversity patterns among codistributed taxa is unknown; this is an issue underlying any recommendation based on a single taxonomic group. Nonetheless, using genetic data for one species highlights potentially significant omissions in conventional habitat-based spatial planning as this genetic diversity represents potential for these organisms to evolve and adapt to changing environmental conditions. Importantly, this example highlights how the conservation priorities changed for different types of genetic data used, indicating the need for careful evaluation of conservation objectives relating to genetic information.

Of the system characteristics tested, the simple measure of assigning subregions based on genetic structure, for example, led to higher priorities in areas where genetic structure can be represented that were not included in the reference reserve system (Fig. 2A). We also found close overlap in conservation outcomes for allelic richness and Local $F_{ST}$ measures, which suggests that some parameters, although having somewhat different interpretations in evolutionary context, may yield similar information for conservation decisions. For this giant clam data set, allelic richness and local $F_{ST}$ were inversely correlated ($R^2 = 0.04, P < 0.001$); insofar as this general result holds true, then (in the absence of admixture) simple statistics such as allelic

![Figure 2. Difference maps of conservation priorities relative to the baseline scenario that did not include genetic data for (A) genetic sub-regions (Structure clusters), (B) adaptive potential (allelic richness), (C) genetic distinctness of sites (local $F_{ST}$), and (D) recent rates of gene flow among sites (connectivity). These were calculated from selection frequencies of 100 Marxan runs. Cells are colored to represent those that were never selected (transparent), those whose high conservation priority remained unchanged between baseline and genetic scenario (most cells had 100% selection frequency), and changes of different intensity with cells having higher or lower priority between baseline and genetic scenario as indicated by the color scale.](image-url)
richness might sufficiently represent the information that population genetics can contribute to spatial planning.

Our results show that important genetic characteristics and biological processes are probably not well represented when using habitats or species richness to develop reserve systems. Furthermore, degradation of reefs makes it increasingly inappropriate to focus conservation solely on biodiversity goals, as pristine and biodiverse reefs are becoming increasingly rare (Hughes et al. 2003, Halpern et al. 2008, Sandin et al. 2008). For instance, maintaining the function and productivity of reefs is highlighted as a primary management goal in countries like the Philippines and Indonesia, which rely on products from reefs and adjacent habitats to a high degree. Our genetic design principles illustrate the different use of genetic information for contrasting objectives, for example, hybridization zones might be viewed as valuable elements of biodiversity, but in the context of functionality conservation effort should focus on the core ranges of hybridizing species (Table 1). Increasing human pressures and limited conservation budgets may often push this trade-off towards maintaining function over biodiversity.

Although many population genetic studies claim a conservation benefit arising from their genetic data, only rarely is such data used in spatial site prioritization for conservation. Why is that? An obvious reason is that population genetic data are often difficult to interpret (Waples et al. 2008, Karl et al. 2012), and conservation approaches that map a parameter to inform conservation decisions cannot easily accommodate nuances and uncertainty. For example, in the present study, we focused on genetically distinct populations of randomly-mating individuals within a species of interest because this concept aligns nicely with the idea, in conservation planning, that planning units or clusters represent a meaningful spatial unit. Focusing on populations may be somewhat problematic, where genetic system characteristics do not tell the whole story. An extreme example is that of the spinner dolphins in which there is no population genetic structure in either nuclear or mitochondrial loci assayed to date among recognized sub-species that are distinct only at Y-chromosomal markers (Andrews et al. 2013). For some marine animals, despite low levels of population genetic structure (and thus previously assumed high levels of gene flow) stepping stone isolation-by-distance dispersal processes probably drive local-scale differentiation (Puebla et al. 2008, Pinsky et al. 2010, Crandall et al. 2012), and shared genetic breaks in suites of coral reef species emerge when considering many species in a multispecies comparison (Toonen et al. 2011). Such patterns would not be captured when using structure as a genetic system parameter in conservation prioritization.

Another challenge for incorporating population genetic data in spatial marine prioritization is the fact that existing spatial planning software requires connectivity (migration) and conservation feature data for all planning units in the analysis. Because genetic sampling will necessarily occur in discrete (and often few) locations, a major challenge remains in extrapolating these data to all unsampled locations throughout the area of interest to represent the “true” patterns of the genetic system attribute in question. While using more quantitative species distribution modeling techniques would be an ideal methodology to determine probable values of unsurveyed sites (Guisan et al. 2006, Beger and Possingham 2008), a low number of samples (27 sites) and high predictor variability across the sampling area (e.g., in reef types) typically results in inconclusive models and very low predictive power. Statistical
and biophysical models of underlying processes (such as larval dispersal) aiming to predict evolutionary patterns are under development for coral reef habitats (Kool et al. 2011, Treml et al. 2012, Sbrocco, NESCent, pers comm). However, given that evolutionary patterns form over long timescales, and it is often unknown what drives a pattern (e.g., geography vs species’ attributes), the spatial predictors of genetic system attributes are often just as sparsely sampled as the genetic attribute itself (Espindola et al. 2012). In the example used here, we therefore interpolated a data surface using a minimum curvature spline technique that considers the islands of the archipelago as barriers (Zoraster 2003). Interpolation techniques using geographic distance bounded by islands may be a technique that adequately estimates patterns for T. crocea, as giant clams are known to be highly differentiated (DeBoer et al. 2008, Nuryanto and Kochzius 2009). Thus, spatial patterns of population characteristics may indeed be driven by geographical patterns. Most interpolation methods use geographical distance to extrapolate point data across space, but are limited because they fail to consider the environmental and ecological drivers of the observed ecological and evolutionary patterns (but see geostatistical kriging approaches, Rossi et al. 1992).

Although the 27 sampling sites in our case study is a limited representation of the hugely diverse reefs of Indonesia and the Philippines, this clam study represents one of the densest sampling efforts in the Indo-Pacific to date (Keyse et al. 2014). Clearly, these samples need to be expanded to cover the entire study region (e.g., southwest Indonesia has not been sampled), or an adequate process-based model should be used to interpolate and extrapolate the data. Otherwise, the unsampled regions will not be included in the conservation prioritization, highlighting the practical limitations of applying genetic data to spatial planning. From a perspective of extrapolating point data, a higher density and broader coverage of sampling sites is advisable to get the best representation of genetic parameters across the planning region. Similarly, to use more statistically-sound methods of distribution modeling, future genetic sampling designs would require covering different sites that may represent the variability in the drivers (i.e., variables used to predict genetic parameters at unsampled sites) most likely to shape genetic system characteristics.

Genetic sampling extent can also affect the interpretation of genetic system parameters, especially at the edges of sampled regions. As an illustration of edge effects, Local $F_{ST}$ (genetic distinctiveness) is defined as “the probability that two genes chosen randomly from the population share a common ancestor within that population without immigration or colonization” (Balding 2003, Gaggiotti et al. 2009). Thus, our higher genetic distinctness values in the Indian Ocean and Cenderawasih Bay may relate to these sites being at the margins of our study region, and could thus either be true sites of high Local $F_{ST}$ or it could be high because of uncapture genetic characteristics adjacent to but outside our target region. Measuring rare or unique alleles (Table 2) also requires exhaustive sampling across the entire project range, because the method would be highly skewed if “rare” alleles were quite common at unsampled sites. These edge effects may translate into higher conservation priorities being given to regions marginal to a study area that are not truly genetically distinct, or priority sites with unique alleles might not be captured in a reserve system. To avoid such edge effects, the entire species range must be sampled.

While there are clear technical and methodological hurdles to incorporating genetic data into marine spatial planning, another major barrier is institutional. Cultural differences between scientific communities play a large role in constricting
the pipeline between genetic research and management outcomes (Waples et al. 2008). For example, clear distinctions have been made between evolutionarily significant connectivity (connectivity among populations on evolutionary time scales) and demographic connectivity (connectivity that influences local abundance and dynamics of populations on ecologically relevant time scales) (Hedgecock et al. 2007). While the distinction between demographic and evolutionary connectivity is an important one, this is only one of many attributes to be considered in prioritization. As evolutionary biologists should not overstate the importance of genetic connectivity estimates for informing patterns of demographic exchange, conservation practitioners must also realize that genetic data (e.g., genetic diversity and evolutionary potential) can be important variables to consider in marine spatial planning. For example, at a recent Coral Triangle Initiative workshop (2012), regions in Sumatra and Sulawesi (Makassar Strait) were prioritized for conservation based on information about regional genetic diversity, the former for its genetic uniqueness in the Indonesian Archipelago and the latter for its role in maintaining gene flow across the region (PH Barber pers obs). This example emphasizes an increasing receptivity of marine managers to genetic data once put into a conservation appropriate context.

A second example of an institutional barrier is the high cost of population genetic research that many managers perceive. While genetics laboratory equipment has high capital costs, the recent “next-generation” revolution in genetic technology has brought about an exponential decrease in the cost of genetic analysis per sample. Large multilocus datasets containing tens of thousands of single nucleotide polymorphism (SNP) loci can now be generated in a single step (Genotyping by sequencing, GBS; see Willette et al. 2014). Thus, it has become not only possible, but often cheaper, to generate datasets with hundreds to many thousands of SNPs per individual than it cost to run traditional sequencing of a handful of loci only a few years ago. The availability of so many markers greatly enhances the power of genetic analyses, allowing researchers to, for example, determine the ancestry of Europeans to within a few hundred kilometers of their historical origin (Novembre et al. 2008), or uniquely “tag” (DNA-fingerprint) all offspring of an individual from a small tissue biopsy (Jones et al. 2005, Anderson and Garza 2006). As costs continue to fall and genetic methods such as parentage-based tagging and coalescent modeling continue to improve, spatial coverage and accuracy of genetic system characteristics will become increasingly more suited to spatial planning. Yet, major fixed expenditure is still required to collect samples across a wide range of sites; this field sampling cost is unlikely to decrease.

In the present study we have presented a framework of genetic decision rules using a single species case study from a single species with exceptional sampling across the Coral Triangle region. Such an example highlights the differences in conservation priorities when using genetic data, but it is clear that different life histories across taxonomic groups and species result in drastically different spatial distributions of genetic characteristics in populations (see Treml et al. 2012 for implications of life histories on broad-scale population connectivity). A next step is to evaluate how conservation priorities change when including genetic data for multiple taxa with contrasting patterns in genetic attributes. Similarly, the genetic data in this case study are constrained by including only microsatellite diversity from a handful of loci; more genomically comprehensive examinations of within species data would enhance estimates of functional diversity which is more relevant for evaluating adaptive potential.
(Funk et al. 2012). In mega-diverse systems such as coral reefs, full consideration of genetic diversity among a wide cross spectrum of species is important to represent major parts of overall biodiversity and functions, and our study represents a first step towards routinely incorporating such data into future conservation efforts.

Acknowledgments

We thank the National Evolutionary Synthesis Centre (NEscent) for hosting the Molecular Ecology and Evolution of the Indo-Pacific Catalysis meeting, which brought together a diverse group of scientists working on the Indo-Pacific coral reef genetic ecology and conservation. MB was funded by an Australian Research Council Discovery Early Career Research Award. Data from T. crocea was supported by an NSF CAREER grant (OCE-0349177) to PHB and an NSF PIRE grant (OISE-0730256). RJT and KAS were supported in part by NOAA National Marine Sanctuaries Program (MOA grant No. 2005-008/66882) and NSF OCE-1260169 grants. We are grateful for thoughtful comments and discussion during the workshop and analysis from H Fox, G Ahmadia, N Yasuda, C Fauvelot, and L van Herwerden.

Literature Cited


Allen GR, Erdmann MV. 2012. Reef fishes of the East Indies. Tropical reef research, distributed by University of Hawai‘i Press, Honolulu [Hawaii].


Beger et al.: Evolving coral reef conservation with genetic information


