2014

Phylogeography Unplugged: Comparative Surveys in the Genomic Era

Brian W. Bowen
Kartik Shanker
Nina Yasuda
Maria Celia Malay
Sophie Von Der Heyden

See next page for additional authors

Follow this and additional works at: https://digitalcommons.odu.edu/biology_fac_pubs

Part of the Ecology and Evolutionary Biology Commons

Repository Citation
Bowen, Brian W.; Shanker, Kartik; Yasuda, Nina; Malay, Maria Celia; Heyden, Sophie Von Der; Paulay, Gustav; Rocha, Luiz A.; Selkoe, Kimberly A.; Barber, Paul H.; Williams, Suzanne T.; Lessios, Harilaos A.; Crandall, Eric D.; Bernardi, Giacomo; Meyer, Christopher P.; Carpenter, Kent E.; and Toonen, Robert J., "Phylogeography Unplugged: Comparative Surveys in the Genomic Era" (2014). Biological Sciences Faculty Publications. 122
https://digitalcommons.odu.edu/biology_fac_pubs/122

Original Publication Citation
Phylogeography unplugged: comparative surveys in the genomic era

Brian W Bowen 1*
Kartik Shanker 2
Nina Yasuda 3
Maria Celia (Machel) D Malay 4
Sophie von der Heyden 5
Gustav Paulay 6
Luiz A Rocha 7
Kimberly A Selkoe 8
Paul H Barber 9
Suzanne T Williams 10
Harilaos A Lessios 11
Eric D Crandall 12
Giacomo Bernardi 13
Christopher P Meyer 14
Kent E Carpenter 15
Robert J Toonen 1

ABSTRACT.—In March 2012, the authors met at the National Evolutionary Synthesis Center (NESCent) in Durham, North Carolina, USA, to discuss approaches and cooperative ventures in Indo-Pacific phylogeography. The group emerged with a series of findings: (1) Marine population structure is complex, but single locus mtDNA studies continue to provide powerful first assessment of phylogeographic patterns. (2) These patterns gain greater significance/power when resolved in a diversity of taxa. New analytical tools are emerging to address these analyses with multi-taxon approaches. (3) Genome-wide analyses are warranted if selection is indicated by surveys of standard markers. Such indicators can include discordance between genetic loci, or between genetic loci and morphology. Phylogeographic information provides a valuable context for studies of selection and adaptation. (4) Phylogeographic inferences are greatly enhanced by an understanding of the biology and ecology of study organisms. (5) Thorough, range-wide sampling of taxa is the foundation for robust phylogeographic inference. (6) Congruent geographic and taxonomic sampling by the Indo-Pacific community of scientists would facilitate better comparative analyses. The group concluded that at this stage of technology and software development, judicious rather than wholesale application of genomics appears to be the most robust course for marine phylogeographic studies. Therefore, our group intends to affirm the value of traditional (“unplugged”) approaches, such as those based on mtDNA sequencing and microsatellites, along with essential field studies, in an era with increasing emphasis on genomic approaches.
At the NESCent Indo-Pacific catalysis meeting in March 2012, authors met to discuss issues in comparative marine phylogeography. Our goal here is to map out a robust study design for marine phylogeography in this genomic era (see Selkoe et al. 2008, Hellberg 2009, Marko and Hart 2011, Riginos and Liggins 2013 for reviews of the latest analytical approaches). Phylogeographic studies of marine systems differ from terrestrial systems in many respects, including (1) greater diversity of major evolutionary lineages; (2) species with vast ranges, extensive dispersal potential, and large effective population sizes; (3) habitats that can be difficult, expensive, and sometimes dangerous to sample; and (4) poor understanding of factors controlling rates and pathways of gene flow. These conditions apply to all phylogeographic studies, but are nearly universal in marine systems. The corresponding uncertainties may limit phylogeographic inference in marine compared with terrestrial studies of similar scope. In this essay, our goal is to summarize lessons from the past, make recommendations for study design and resolve the path forward for marine phylogeographic surveys, in an era of ever-increasing genomic information.

A Brief History

At the inception of phylogeography, Avise et al. (1987) noted the primary roots of this field in biogeography, phylogenetics, and population genetics. All of those fields in turn have predecessors with foundations in morphology and field observations. As these fields advanced into biochemical and molecular methodologies, a few practitioners lamented the loss of what came before. Although protein gel electrophoresis opened up the possibility of determining gene flow among populations for the first time, it could not replace the knowledge gleaned from hundreds of specimens, or thousands of hours of observations. Field biologists made observations about organisms that could not be revealed by laboratory-based analyses, many which were important for meaningful interpretation of bands on a gel. Sometimes that traditional knowledge was used, sometimes ignored, and sometimes forgotten. Certainly, a little intimacy was lost in the relationship between scientist and subject, as funding priorities for research shifted.

Richard Lewontin, one of the pioneers of protein electrophoresis (Hubby and Lewontin 1966), famously wondered 25 yrs later whether this technology was a boon or bane to organismal biology (Lewontin 1991). He noted the loss of informative studies on morphology and cytogenetics, and the shift in funding priorities to laboratory-bench science. A few years earlier, Wes Brown and colleagues introduced mtDNA for population and evolutionary studies (Brown and Wright 1975, Avise et al. 1979), and soon the tools shifted yet again. The mtDNA sequence information made it possible to estimate the relationships among species based on direct comparisons of nucleotide sequence data, instead of allele frequency data typical of allozyme studies. Tissue preservation requirements were also less stringent, and it became possible to incorporate historical data from museum collections. During the subsequent heyday of single marker mtDNA studies, some practitioners noted that the research lacked equal resolution in the nuclear genome. Researchers in the new field of molecular ecology knew that mtDNA was missing something that predecessors valued: the genetic contribution from both parents (Quinn and White 1987, Karl et al. 1992).

When the chain-termination method of DNA sequencing was published (Sanger and Coulson 1975), the process was impractical for most organismal laboratories
because of the vast resources needed to determine the sequence. The chemical modification and cleavage method (Maxam and Gilbert 1977) allowed direct sequencing of purified DNA, but the process remained technically complex and impractical for population studies. With the exception of a few very well-funded laboratories, attempts to sequence DNA were not routine until computer and laboratory technology advanced to the point where by the early 1990s laboratories were able to sequence thousands of base pairs if they could manage the cost (both in terms of labor and reagents required). Polymerase chain reaction (PCR) assays were used to greatly increase the accessibility of the nuclear genome (Karl and Avise 1993). The Human Genome Project led engineers and scientists around the world to improve the speed and accuracy of sequencing reads, which led to increased availability and a concordant reduction in the overall cost associated with DNA data (Watson 1990).

With the advent of massively parallel “next generation” sequencing it is now possible to generate genomic data and assess phylogeography on scales never before considered (e.g., Puritz et al. 2012a). We can now survey entire nuclear, mitochondrial, and chloroplast genomes, and even the expression of genomes (Carstens et al. 2012). Based on the arc of events described above, is this phase the *end of history as we know it* for single-locus studies in phylogeography? The latest advances into next generation sequencing have opened the gateway to rapid scientific advances, but we currently are limited by the computing power and analytical methods required to take full advantage of the new data and handle the massive increase in numbers of loci and sequence length (Rocha et al. 2013, Willette et al. 2014).

### The Comparative Advantage

The first lesson that stands out from two decades of comparative marine phylogeography is that large-scale patterns in genetic or species diversity are the result of multiple drivers. Even closely related and co-distributed species with similar pelagic larval duration (PLD) can exhibit radically different patterns of genetic structure (Table 1).

Much of large-scale biogeography has been built on top-down approaches, often based on 200 yrs of accumulated distribution data; for example large-scale diversity patterns have led to hypotheses that attribute such patterns to singular evolutionary processes, even though there is every reason to expect that multiple processes are in play (Paulay 1997). The center of origin (or diversification), center of accumulation, and center of overlap hypotheses are classic examples of this plurality in Indo-Pacific biogeography (Barber 2009, Bowen et al. 2013). In contrast, phylogeographic studies are bottom up, built one species at a time: once assembled from many species-level analyses, they provide a powerful framework for evaluating broad distributional patterns. If one wants to understand complex geographic patterns of biodiversity, then species or lineages are the relevant unit of replication: data from a multitude of taxa are needed. Thus far, the diversity of patterns at the species level indicates that global phylogeographic patterns are the result of the combined effects of multiple forces operating over different spatial, temporal, and ecological scales. Greater replication across species is desirable to quantify the prevalence of different processes.

While single species cannot elucidate general trends on their own, they provide the building blocks for subsequent comparative phylogeography. Sometimes single taxon studies are the only feasible logistic approach due to funding and geopolitical
Table 1. Comparative phylogeographic studies of marine organisms, listed in chronological order. These cases illustrate the extreme diversity in outcomes that is expected from a group of organisms spread across 500 million years of evolutionary history. However, recent studies that compare many species across the same range have found congruent genetic partitions (Barber et al. 2011, Carpenter et al. 2011, Toonen et al. 2011 for examples). These studies also reveal extensive gaps in comparative studies of plants, prokaryotes, pelagic organisms, oceanic migrants, marine birds, and marine mammals. Abbreviations: n = number of species or taxonomic groups, PLD = pelagic larval duration, COI = cytochrome oxidase subunit 1, ISSR = inter-simple sequence repeat, ITS = internal transcribed spacer, RAG1 = recombination activation gene 1, RFLP = restriction fragment length polymorphism.

<table>
<thead>
<tr>
<th>n</th>
<th>Taxonomic group</th>
<th>Species group</th>
<th>Larval type, PLD, and related information</th>
<th>Markers</th>
<th>Geographic region</th>
<th>Conclusion</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Teleostei</td>
<td>11 reef species, 1 pelagic species</td>
<td>N/A</td>
<td>allozymes</td>
<td>Across East Pacific Barrier: Hawaii and East Pacific</td>
<td>Transpacific taxa have not split into W Pacific and E Pacific species. Long range dispersal highly likely; Overall congruent genetic structuring patterns but totally different pattern for one species. Genetic break between warm-temperate faunas of the Gulf of Mexico and West Atlantic</td>
<td>Rosenblatt and Waples 1986</td>
</tr>
<tr>
<td>10</td>
<td>Actinopterygii</td>
<td>marine fishes</td>
<td>range from no PLD to a few months</td>
<td>allozymes</td>
<td>California and Mexico</td>
<td></td>
<td>Waples and Rosenblatt 1987</td>
</tr>
<tr>
<td>3</td>
<td>Actinopterygii</td>
<td>coastal fishes</td>
<td>various</td>
<td>mtDNA RFLP</td>
<td>West Atlantic</td>
<td></td>
<td>Bowen and Avise 1990</td>
</tr>
<tr>
<td>12</td>
<td>Chordata, Actinopterygii, Perciformes, Blennioidei, marine fishes</td>
<td>PLD up to 2 mo</td>
<td>allozymes</td>
<td>California and Mexico</td>
<td>All species consistently showed little genetic isolation. Species showed genetic divergence between island and mainland populations</td>
<td>Stepien and Rosenblatt 1991</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Perciformes</td>
<td>reef fishes</td>
<td>4 species benthic/18–30 d, 4 species pelagic /15–50 d</td>
<td>mtDNA RFLP</td>
<td>Caribbean</td>
<td>Three of 8 species showed low but significant genetic structuring. Egg type and PLD cannot be a simple indicator of geographic structure.</td>
<td>Shulman and Bermingham 1995</td>
</tr>
<tr>
<td>7</td>
<td>Teleostei</td>
<td>reef fishes</td>
<td>PLD 0–55 d, 4 species have demersal eggs, 2 species (acanthurid and caesionid) are pelagic spawners.</td>
<td>allozymes</td>
<td>Great Barrier Reef</td>
<td>High PLD species (2 species) show no significant structuring across GBR; low PLD species (4 species) show genetic structuring; no PLD species (1 species) shows genetic structuring.</td>
<td>Doherty et al. 1995</td>
</tr>
<tr>
<td>3</td>
<td>Teleostei</td>
<td>reef fishes</td>
<td>24, 39, 60 d PLD</td>
<td>allozymes</td>
<td>New Caledonia</td>
<td>Two of three species showed spatial heterogeneity with break across eastern and western side of New Caledonia lagoon, one sp. had no structure</td>
<td>Planes et al. 1998</td>
</tr>
<tr>
<td>n</td>
<td>Taxonomic group</td>
<td>Species group</td>
<td>Larval type, PLD, and related information</td>
<td>Markers</td>
<td>Geographic region</td>
<td>Conclusion</td>
<td>Citation</td>
</tr>
<tr>
<td>----</td>
<td>----------------------------------</td>
<td>-------------------------------------</td>
<td>------------------------------------------</td>
<td>---------</td>
<td>--------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>9</td>
<td>Cnidaria corals</td>
<td>5 species: brooder; 4 species: broadcast spawner</td>
<td>allozymes</td>
<td></td>
<td>Great Barrier Reef</td>
<td>Considerable genetic subdivision particularly in brooders, but four broadcast and three of brooders show gene flow for entire GBR. Three of four species show structure, not congruent, southward gene flow along Californian/Oregonian provinces.</td>
<td>Ayre and Hughes 2000</td>
</tr>
<tr>
<td>5</td>
<td>Arthropoda, Echinodermata, Mollusca</td>
<td>coastal marine species</td>
<td>pelagic larvae and non-dispersive crawl-away larvae</td>
<td>mtDNA COI</td>
<td>California and Oregon</td>
<td>European older than North American populations, five N. American populations show colonization from Europe. Larval dispersal and Fst closely linked; long PLD showed no structure, shortest PLD and reef living had highest structure.</td>
<td>Wares et al. 2001</td>
</tr>
<tr>
<td>6</td>
<td>Arthropoda, Mollusca, Echinodermata</td>
<td>North Atlantic rocky intertidal</td>
<td>various</td>
<td>mtDNA COI</td>
<td>European and North American</td>
<td>Diverse groups share a generalized pattern. Historical influence of geographical isolation is a dominant factor on the North American coast.</td>
<td>Wares and Cunningham 2001</td>
</tr>
<tr>
<td>3</td>
<td>Teleostei marine fishes</td>
<td>PLD of 18, 24, and 50 d</td>
<td>mtDNA control region</td>
<td></td>
<td>Eastern Pacific</td>
<td>Larval dispersal and Fst closely linked; long PLD showed no structure, shortest PLD and reef living had highest structure.</td>
<td>Riginos and Victor 2001</td>
</tr>
<tr>
<td>19</td>
<td>Arthropoda, Mollusca, Rhodophyta, Teleostei, Annelida, Echinodermata, Anthophyta</td>
<td>Atlantic coastal species</td>
<td>N/A</td>
<td>mtDNA COI</td>
<td>Northwestern Atlantic</td>
<td>Diverse groups share a generalized pattern. Historical influence of geographical isolation is a dominant factor on the North American coast.</td>
<td>Wares 2002</td>
</tr>
<tr>
<td>3</td>
<td>Teleostei surgeonfishes</td>
<td>PLD 45–70 d</td>
<td>mtDNA cytochrome b</td>
<td>Atlantic</td>
<td>Atlantic</td>
<td>Reef specialist has more genetic structure than generalist. None had structure. Different mtDNA diversity between lagoon and outer-slop species.</td>
<td>Rocha et al. 2002</td>
</tr>
<tr>
<td>7</td>
<td>Chaetodontidae reef fishes</td>
<td>3 species with pelagic eggs, 4 species with benthic egg</td>
<td>mtDNA control region</td>
<td>French Polynesia</td>
<td>French Polynesia</td>
<td>Diverse groups share a generalized pattern. Historical influence of geographical isolation is a dominant factor on the North American coast.</td>
<td>Fauvelot et al. 2003</td>
</tr>
<tr>
<td>5</td>
<td>Teleostei marine fishes</td>
<td>N/A</td>
<td>mtDNA control region, allozymes</td>
<td>Mediterranean Sea and northeast Atlantic</td>
<td>Mediterranean in both mtDNA and nDNA, two had genetic break in either mtDNA or nDNA and the remaining one had no differentiation.</td>
<td>Bargelloni et al. 2003</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>Taxonomic group</td>
<td>Species group</td>
<td>Larval type, PLD, and related information</td>
<td>Markers</td>
<td>Geographic region</td>
<td>Conclusion</td>
<td>Citation</td>
</tr>
<tr>
<td>----</td>
<td>-------------------------</td>
<td>---------------</td>
<td>------------------------------------------</td>
<td>------------------------------</td>
<td>----------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>12</td>
<td>Actinopterygii</td>
<td>marine fishes</td>
<td>N/A</td>
<td>mtDNA control region, cytochrome b, allozymes, RFLP, ITS</td>
<td>California</td>
<td>Eight of 12 species showed strong genetic break between Gulf of California and Pacific, and 4 species showed no genetic difference.</td>
<td>Bernardi et al. 2003</td>
</tr>
<tr>
<td>6</td>
<td>Crustacea (Amphipoda, Cladocera)</td>
<td>zooplankton</td>
<td>3 species: planktonic; 3 species: benthic</td>
<td>mtDNA COI</td>
<td>Black and Caspian seas</td>
<td>Concordant patterns of genetic break between Black and Caspian sea across all the species. Benthic amphipods have on average four times deeper phylogeographical partition than planktonic cladocerns.</td>
<td>Cristescu et al. 2003</td>
</tr>
<tr>
<td>5</td>
<td>Cnidaria</td>
<td>coral</td>
<td>1 species is broadcast spawner; 3 species are brooder; 1 species has broadcast spawning all species: 7–10 d</td>
<td>allozymes</td>
<td>Great Barrier Reef and Lord Howe Island</td>
<td>No consistent difference in larval type, modest to high gene flow in 5 species. Almost no gene flow between Lord Howe Island and GBR populations for all 5 species. Highly geographically structured. Historical and/or biological factors may be more important than oceanography alone.</td>
<td>Ayre and Hughes 2004</td>
</tr>
<tr>
<td>13?</td>
<td>Gastropoda: Lottidae</td>
<td>limpet</td>
<td>mtDNA COI and 16S rRNA</td>
<td>Indo–West Pacific</td>
<td></td>
<td>Highly geographically structured. Historical and/or biological factors may be more important than oceanography alone. Atlantic colonized from Indo-Pacific. Two lineages (species) introduced to Hawaii.</td>
<td>Kirkendale and Meyer 2004</td>
</tr>
<tr>
<td>3</td>
<td>Scyphozoa</td>
<td>Cassiopea</td>
<td>unknown</td>
<td>mtDNA COI</td>
<td>Global</td>
<td>General lack of phylogeographical congruence between all four species, but some congruence noted between two shallow-water species and two deep-water species.</td>
<td>Holland et al. 2004</td>
</tr>
<tr>
<td>4</td>
<td>Teleostei</td>
<td>seahorses</td>
<td>internal brooders with no PLDs. 2 species shallow-water; 2 species deeper water</td>
<td>mtDNA cytochrome b</td>
<td>Southeast Asia</td>
<td>General lack of phylogeographical congruence between all four species, but some congruence noted between two shallow-water species and two deep-water species.</td>
<td>Lourie et al. 2005</td>
</tr>
<tr>
<td>9</td>
<td>Teleostei</td>
<td>reef fishes</td>
<td>cleaner gobies and sponge-dwelling gobies</td>
<td>mtDNA cytochrome b, nDNA RAG1</td>
<td>Caribbean Sea</td>
<td>Mona Passage is a long term biogeographic barrier within Caribbean for all nine lineages examined.</td>
<td>Taylor and Hellberg 2006</td>
</tr>
<tr>
<td>n</td>
<td>Taxonomic group</td>
<td>Species group</td>
<td>Larval type, PLD, and related information</td>
<td>Markers</td>
<td>Geographic region</td>
<td>Conclusion</td>
<td>Citation</td>
</tr>
<tr>
<td>----</td>
<td>-------------------------</td>
<td>------------------------</td>
<td>-------------------------------------------</td>
<td>------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>5</td>
<td>Mollusca (Gastropoda)</td>
<td>snails</td>
<td>all species: up to 4 wks</td>
<td>mtDNA COI, nDNA 28S rRNA</td>
<td>Tropical Indo–West Pacific</td>
<td>Two (of three) species from continental shelf show congruence. Oceanic species show no breaks at Indo-Pacific barrier.</td>
<td>Reid et al. 2006</td>
</tr>
<tr>
<td>7</td>
<td>Crustacea</td>
<td>brackishwater mysid crustacea</td>
<td>1 species estuaries; 6 species both estuaries and rivers</td>
<td>mtDNA COI</td>
<td>Ponto-Caspian Sea</td>
<td>Four species consistently had genetic break related to major basin, and three did not show consistent geographical differentiation.</td>
<td>Audzijonyte et al. 2006</td>
</tr>
<tr>
<td>3</td>
<td>Stomatopods</td>
<td>mantis shrimps</td>
<td>3 species: 4–6 wks; 1 species: 35 d</td>
<td>mtDNA COI</td>
<td>Coral Triangle</td>
<td>Congruent patterns of phylogeographic breaks and moderate to strong genetic structure between North and South Indonesia.</td>
<td>Barber et al. 2006</td>
</tr>
<tr>
<td>3</td>
<td>Nacellidae</td>
<td>limpet</td>
<td>all species: 3–11 d</td>
<td>mtDNA cytochrome b</td>
<td>Cook Strait in New Zealand</td>
<td>Concordant patterns with moderate to strong structure with discontinuity between North and South Islands.</td>
<td>Goldstien et al. 2006</td>
</tr>
<tr>
<td>18</td>
<td>Teleostei</td>
<td>reef fishes</td>
<td>all are transpacific species</td>
<td>mtDNA ATPase6, ATPase8</td>
<td>Central and eastern Pacific Ocean</td>
<td>Recent separtion of eastern and central Pacific found in 14 species, older separation with high subsequent gene flow in 2 species, old separation, significant differentiation in 2 species.</td>
<td>Lessios and Robertson 2006</td>
</tr>
<tr>
<td>8</td>
<td>Teleostei</td>
<td>reef fishes</td>
<td>benthic spawners; PLD ranges from 0 to approximately 28 d</td>
<td>mtDNA control region, nDNA ISSRs (for a subset of 5 species)</td>
<td>Great Barrier Reef</td>
<td>MdDNA shows structuring in 1 out of 8 species, ISSRS show structuring in 5 out of 5 species. Species show no relation between genetic structuring and PLD.</td>
<td>Bay et al. 2006</td>
</tr>
<tr>
<td>3</td>
<td>Mullusca: Patellogastropoda</td>
<td>limpets</td>
<td>all species: 3–18 d PLD</td>
<td>mtDNA COI</td>
<td>Hawaii</td>
<td>The spatial scales of partitioning differ among the species even though all have similar life history.</td>
<td>Bird et al. 2007</td>
</tr>
<tr>
<td>n</td>
<td>Taxonomic group</td>
<td>Species group</td>
<td>Larval type, PLD, and related information</td>
<td>Markers</td>
<td>Geographic region</td>
<td>Conclusion</td>
<td>Citation</td>
</tr>
<tr>
<td>----</td>
<td>------------------------</td>
<td>---------------------</td>
<td>------------------------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>7</td>
<td>Testudines</td>
<td>sea turtles</td>
<td>pelagic juveniles in most cases</td>
<td>mtDNA control region</td>
<td>Global</td>
<td>Genetic breaks between Atlantic and Indo-Pacific predominate in tropical species. Temperate species are less constrained by thermal barriers.</td>
<td>Bowen and Karl 2007</td>
</tr>
<tr>
<td>4</td>
<td>Asteroidea, Gastropoda, Caridea</td>
<td>starfishes, shrimps, snails</td>
<td>10–34 d PLD</td>
<td>mtDNA COI</td>
<td>Coral Triangle</td>
<td>Discordant patterns across species. Two starfish showed moderate and strong genetic structure, which was mirrored in a specialist parasite, but not in a generalist.</td>
<td>Crandall et al. 2008b</td>
</tr>
<tr>
<td>3</td>
<td>Acanthuridae</td>
<td>reef fishes</td>
<td>all species: similar life history and 90 d</td>
<td>mtDNA control region</td>
<td>Indo-West Pacific</td>
<td>Congruent lack of spatial population structure, likely driven by ecological and environmental factors.</td>
<td>Horne et al. 2008</td>
</tr>
<tr>
<td>6</td>
<td>Crustacea (Cladocera)</td>
<td>zooplankton</td>
<td>N/A</td>
<td>mtDNA COI</td>
<td>Global</td>
<td>Four of six species had semi-allopatric breaks with significant structure, one species had no structure around the globe. The remaining one species had no structure due to small sampling scale.</td>
<td>Durbin et al. 2008</td>
</tr>
<tr>
<td>9</td>
<td>Gastropoda</td>
<td>snails</td>
<td>1 species with 9–15 d, the others unknown possibly 9–15 d</td>
<td>mtDNA COI</td>
<td>Macaronesian Islands, North Africa, Atlantic and Mediterranean</td>
<td>Incongruent patterns due to independent histories and colonization event. But all species showed strong genetic break between islands and continental populations.</td>
<td>Sa-Pinto et al. 2008</td>
</tr>
<tr>
<td>115</td>
<td>Echinoidea, Crustacea, Teleostei, Mollusca</td>
<td>geminate clades of various transisthmian taxa</td>
<td>N/A</td>
<td>various mtDNA and nuclear loci</td>
<td>East Pacific and West Atlantic sides of the Isthmus of Panama</td>
<td>Most geminate species (73 pairs) split before completion of the Isthmus, many more (34 pairs) split at completion of Isthmus, a few (8 pairs) maintained gene flow post-Isthmus. Some adaptation-driven divergences in life histories.</td>
<td>Lessios 2008</td>
</tr>
</tbody>
</table>
Table 1. Continued.

<table>
<thead>
<tr>
<th>n</th>
<th>Taxonomic group</th>
<th>Species group</th>
<th>Larval type, PLD, and related information</th>
<th>Markers</th>
<th>Geographic region</th>
<th>Conclusion</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Mollusca, Arthropoda, Algae, Vertebrata, Annelida</td>
<td>various temperate North Atlantic taxa</td>
<td>N/A</td>
<td>mtDNA COI, cytochrome b, cox 2–3 spacer, intergenic spacer; plastid DNA</td>
<td>North Atlantic</td>
<td>Multiple glacial refugia identified but patterns of structuring varied between species.</td>
<td>Maggs et al. 2008</td>
</tr>
<tr>
<td>15</td>
<td>Teleostei, Arthropoda, Mollusca, Cnidaria</td>
<td>rocky shore and estuarine variants</td>
<td>various</td>
<td>mtDNA COI, control region</td>
<td>South Africa</td>
<td>Congruent patterns across some phylogeographic breaks, but not congruence for all species.</td>
<td>Von der Heyden 2009</td>
</tr>
<tr>
<td>7</td>
<td>Mollusca, Polychaeta</td>
<td>hydrothermal vent species</td>
<td>1 species: plant octrophic larva, 5 species: lecithotrophic larva</td>
<td>mtDNA COI</td>
<td>Deepsea, the East Pacific Rise</td>
<td>Six of seven species show distinct northern and southern clades. Southern clades for those six species all showed signals of demographic expansions.</td>
<td>Plouviez et al. 2009</td>
</tr>
<tr>
<td>38</td>
<td>Echinodermata, Teleostei, Mollusca, Crustacea</td>
<td>low-dispersal fish and invertebrates</td>
<td>two groups: restricted dispersers and planktonic dispersers</td>
<td>allozymes, mtDNA, and nDNA</td>
<td>Southeast and southwest coasts of the United States</td>
<td>Planktonic larvae may reflect contemporary oceanography, while breaks in restricted dispersers reflect historical processes.</td>
<td>Pelc et al. 2009</td>
</tr>
<tr>
<td>8</td>
<td>Teleostei</td>
<td>triplefin fishes</td>
<td>shallow-water species and deeper water species</td>
<td>mtDNA control region</td>
<td>New Zealand</td>
<td>Three of eight showed strong structure and IBD, three of eight had little/no structure and high gene flow, two of eight had strong structure (no IBD)</td>
<td>Hickey et al. 2009</td>
</tr>
<tr>
<td>23</td>
<td>Cnidaria, Arthropoda, Echinodermata, Mollusca, Chordata</td>
<td>intertidal invertebrates</td>
<td>17 species with planktonic larvae, 3 with direct developing larvae, 3 with brooding larvae.</td>
<td>mtDNA, nuclear DNA, allozymes, microsatellites, minisatellites</td>
<td>Australia</td>
<td>Twelve of 15 planktonic dispersers had no structure, three had strong structure. Seven brooding or direct developing species had structure. Using sheltered habitat predict gene flow across the barrier of SE corner of Australia.</td>
<td>Ayre et al. 2009</td>
</tr>
<tr>
<td>n</td>
<td>Taxonomic group</td>
<td>Species group</td>
<td>Larval type, PLD, and related information</td>
<td>Markers</td>
<td>Geographic region</td>
<td>Conclusion</td>
<td>Citation</td>
</tr>
<tr>
<td>----</td>
<td>-----------------------------</td>
<td>---------------</td>
<td>-------------------------------------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>3</td>
<td>Acanthuridae</td>
<td>reef fish</td>
<td>all similar short PLD but different biogeographic ranges</td>
<td>mtDNA cytochrome b</td>
<td>Hawaii</td>
<td>One wide ranging species has no structure but the two endemic species show stronger structure. Breaks between recently diverged sister-species pairs mostly allopatric and occur peripatrically (13 pairs), between ocean basins (2 pairs), between the E and W Atlantic (2 pairs), between the E Pacific and W Atlantic (1 pair), and along continental margins (1 pair). One case of depth-related sympatric speciation.</td>
<td>Eble et al. 2009</td>
</tr>
<tr>
<td>56</td>
<td>Arthropoda (Decapoda)</td>
<td>hermit crabs</td>
<td>mostly reef-associated</td>
<td>mtDNA COI, 16S; nDNA H3</td>
<td>Indo–West Pacific, east Pacific, west Atlantic, east Atlantic</td>
<td></td>
<td>Malay and Paulay 2010</td>
</tr>
<tr>
<td>3</td>
<td>Chaetodontidae</td>
<td>butterflyfishes</td>
<td>pelagic larvae</td>
<td>mtDNA cytochrome b</td>
<td>Hawaii</td>
<td>No population subdivisions observed in Hawaii, and two species colonized Hawaii from the West Pacific. Despite few genetic breaks, spatial patterns of diversity and differentiation correlate in Southern California but not Baja California.</td>
<td>Craig et al. 2010</td>
</tr>
<tr>
<td>3</td>
<td>Teleostei, Arthropoda,</td>
<td>kelp forest predators</td>
<td>various</td>
<td>microsatellites</td>
<td>Pacific coast of Mexico and California</td>
<td>Despite few genetic breaks, spatial patterns of diversity and differentiation correlate in Southern California but not Baja California. Caule Mendicino shows genetic break. Area north of Oregon shows stronger differentiation than to the south. High-to-mid intertidal species show more structure than low intertidal species.</td>
<td>Selkoe et al. 2010</td>
</tr>
<tr>
<td>50</td>
<td>Annelida, Cnidaria,</td>
<td>various intertidal</td>
<td>various</td>
<td>mtDNA COI</td>
<td>Pacific coast of North America</td>
<td>Cape Mendicino shows genetic break. Area north of Oregon shows stronger differentiation than to the south. High-to-mid intertidal species show more structure than low intertidal species.</td>
<td>Kelly and Palumbi 2010</td>
</tr>
<tr>
<td>27</td>
<td>Teleostei, Arthropoda,</td>
<td>species from Hawaii</td>
<td>N/A</td>
<td>mtDNA COI, cytochrome b</td>
<td>Hawaii</td>
<td>Four concordant barriers to gene flow in Hawaiian archipelago. Single species studies do not show barriers clearly.</td>
<td>Toonen et al. 2011</td>
</tr>
</tbody>
</table>
Table 1. Continued.

<table>
<thead>
<tr>
<th>n</th>
<th>Taxonomic group</th>
<th>Species group</th>
<th>Larval type, PLD and related information</th>
<th>Markers</th>
<th>Geographic region(s)</th>
<th>Conclusion</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Teleostei</td>
<td>marine fishes</td>
<td>benthic and pelagic species</td>
<td>mtDNA, microsatellites, allozymes</td>
<td>Pacific and Atlantic</td>
<td>Reef fishes show genetic signatures of the Torres Strait barrier between Pacific and Atlantic oceans.</td>
<td>Carpentier et al. 2011</td>
</tr>
<tr>
<td>4</td>
<td>Teleostei</td>
<td>marine fishes</td>
<td>various PLD</td>
<td>mtDNA, COI, cytochrome b</td>
<td>Coral Triangle</td>
<td>Direct developers had highest values of F&lt;sup&gt;ST&lt;/sup&gt;.</td>
<td>Riginos et al. 2011</td>
</tr>
<tr>
<td>19+</td>
<td>Arthropoda, Mollusca, Echinodermata, Teleostei</td>
<td>coral reef organisms</td>
<td>NA</td>
<td>mtDNA, allozymes</td>
<td>Various</td>
<td>Reef fishes show genetic signatures of the Torres Strait barrier between Pacific and Atlantic oceans.</td>
<td>Carpenter et al. 2011</td>
</tr>
<tr>
<td>214</td>
<td>Annelida (Polychaeta), Echinodermata</td>
<td>polychaetes and echinoderms</td>
<td>NA</td>
<td>mtDNA, COI, cytochrome b</td>
<td>Various</td>
<td>High connectivity between Arctic and both Atlantic and Pacific. Echinoderms have higher connectivities and geographic ranges than polychaetes.</td>
<td>Carpenter et al. 2011</td>
</tr>
<tr>
<td>4</td>
<td>Teleostei</td>
<td>moray eels</td>
<td>very long PLD</td>
<td>mtDNA, COI, cytochrome b</td>
<td>Indo-Pacific</td>
<td>No genetic differentiation across 22,000 km.</td>
<td>Reece et al. 2011</td>
</tr>
<tr>
<td>3</td>
<td>Teleostei</td>
<td>sea cucumbers</td>
<td>varied</td>
<td>mtDNA, cytochrome b</td>
<td>N/A</td>
<td>No genetic differentiation around Antarctica and with depth.</td>
<td>O’Loughlin et al. 2011</td>
</tr>
<tr>
<td>3</td>
<td>Teleostei</td>
<td>Nototthenid fish</td>
<td>bottom-feeding fish with demersal eggs (2 species) and semi-pelagic species (1 species)</td>
<td>mtDNA, cytochrome b</td>
<td>Southern Ocean</td>
<td>Genetic structuring between High-Antarctic and Peninsular populations observed in all species. Genetic differentiation low to moderate structure within species.</td>
<td>Van de Putte et al. 2012</td>
</tr>
<tr>
<td>3</td>
<td>Teleostei</td>
<td>reef fishes</td>
<td>pelagic larvae</td>
<td>mtDNA, cytochrome b</td>
<td>Hawaii</td>
<td>Lag time after species introduction with rearrangements genetic diversity, with the longest lag time yielding the greatest change in mtDNA diversity.</td>
<td>Drew and Barber 2012</td>
</tr>
<tr>
<td>n</td>
<td>Taxonomic group</td>
<td>Species group</td>
<td>Larval type, PLD, and related information</td>
<td>Markers</td>
<td>Geographic region</td>
<td>Conclusion</td>
<td>Citation</td>
</tr>
<tr>
<td>----</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>-------------------------------------------</td>
<td>----------------------------------------------</td>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>4</td>
<td>Teleostei</td>
<td>wrasses</td>
<td></td>
<td>mtDNA COI and control region</td>
<td>Indo-Pacific Ocean</td>
<td>Outer reef species have less structure than lagoon species.</td>
<td>Ludt et al. 2012</td>
</tr>
<tr>
<td>4</td>
<td>Mollusca (Gastropoda)</td>
<td>amphidromous and marine gastropoda</td>
<td>no rafting, unknown PLD</td>
<td>mtDNA COI</td>
<td>South Pacific</td>
<td>Where stepping stones are available, amphidromous species have connectivity similar to marine species but in the region without stepping-stones, marine species showed higher gene flow. Partitions between Red Sea and Indian Ocean range from shallow population structure to deep evolutionary separations.</td>
<td>Crandall et al. 2012</td>
</tr>
<tr>
<td>7</td>
<td>Teleostei</td>
<td>reef fishes</td>
<td>pelagic larvae</td>
<td>mtDNA cytochrome b and COI</td>
<td>Red Sea, Indian Ocean</td>
<td></td>
<td>DiBattista et al. 2013</td>
</tr>
</tbody>
</table>
constraints. Furthermore, species of greater interest for management and conserva-
tion may be subject to tightly-focused studies. However, where possible, the com-
parisons of multiple co-distributed species are desirable to flesh out biogeographic
history (Carpenter et al. 2011, Toonen et al. 2011). For example, congruent phylo-
geographic breaks among multiple species may indicate a shared history, especially
when found at the boundaries of biogeographic provinces (Avise 1992, Dawson 2001,
Wares et al. 2001, Bernardi et al. 2003, Barber et al. 2006, 2011; see Table 1).

The comparative advantage in phylogeography is that concordant patterns in
multiple species may reveal shared evolutionary events. This concordance can ap-
ply across both geography (biogeographic barriers) and time (such as colonization
of Hawaii; Craig et al. 2009). Many examples are provided in Table 1. The NESCent
working group recognized that one of the central challenges of contemporary phy-
logeography is the dearth of analytical tools for comparing and combining analyses
of separate species. Toonen et al. (2011) used a simple chi-square test for detect-
ing where congruent genetic breaks appear more often than at random along the
Hawaiian archipelago for 27 reef species. Barber et al. (2011) found strong concordant
phylogeographic patterns in 14 species of mantis shrimp in Indonesia. Carpenter et
al. (2011) identified multiple regions of the Coral Triangle that have congruent breaks
in mtDNA data across fish and invertebrates. Kelly and Palumbi (2010) used an as-
sortment of 50 intertidal data sets from the west coast of North American to discern
a trend toward stronger genetic partitioning north of Oregon vs the south. However,
none of these studies were able to make strong biogeographical inferences based on a
comprehensive statistical model (see Pelc et al. 2009). One way to address this limita-
tion is through new statistical methods to examine cross-community trends in spa-
tial genetic differentiation across codistributed species (Hickerson and Meyer 2008,
Mirams et al. 2010, KA Selkoe et al. unpubl data).

As the field accumulates spatial genetic data sets at ever increasing rates, there
will be more opportunity to build on these important steps toward extending our
conceptual and empirical understanding of connectivity. These will include more
sophisticated statistical frameworks that combine individual data sets to leverage
statistical power and test predictions about the historical, environmental, taxonom-
ic, and ecological mechanisms underlying shared and divergent patterns. Examples
of promising avenues include application of Hierarchical Approximate Bayesian
Computation (HABC) for phylogeography that can use individual taxa as units of
replication; individual coalescent models for each species can be used to parameter-
ize larger models of community dynamics (Hickerson and Meyer 2008, Carnaval
et al. 2009, Beaumont 2010). HABC methods analyze all the phylogeographic data
sets at once to make across taxon-pair inferences about biogeographic processes
while explicitly allowing for uncertainty in the demographic differences within each
taxon-pair (Hickerson and Meyer 2008). Therefore, HABC approaches are expected
to provide a bridge linking comparative phylogeography with community ecology
(Hickerson and Meyer 2008). On another front, population graph approaches can
build congruence networks based on two or more spatial genetic data sets (Dyer et al.
2010). Population graphs connect the population pairs exchanging migrants in the
network by edges whose length is inversely proportional to the genetic covariance be-
tween the populations (Dyer and Nason 2004). Because conditional genetic distance
(cGD) in a population graph is based upon the differences in genetic covariation as-
associated with both direct and indirect gene flow among populations, cGD should be
more sensitive than conventional pairwise $F_{ST}$ and better suited for landscape genetics (Dyer et al. 2010). Another promising avenue incorporating genetic data sets for conservation planning would be Bayesian Belief Network analysis (BBNs), which can overcome variable sampling coverage across data sets to build a single probability distribution (e.g., Kininmonth et al. 2010a,b). BBNs consist of qualitative and quantitative components. The former is a Directed Acyclic Graph (DAG), which denotes dependencies and independencies between the variables that represent the modeled system. The latter component is Conditional Probability Tables (CPTs), which denote the strengths of the relationship among the variables (Aguilera et al. 2011). The advantage of a BBN is that uncertainties as well as expert opinions can be incorporated, while the strength of the causal relations between the networks variables can be quantified through conditional probabilities (Landuyt et al. 2013). BBNs are now being used for evidence-based conservation management of both aquatic and terrestrial ecosystems (reviewed by Landuyt et al. 2013).

In discussing the greatest gaps in sampling design, a consensus quickly emerged around the need for thorough geographic sampling. The literature contains many examples where species ranges (or evolutionary lineages within species) exceed the boundaries defined by faunal breaks, and sparse sampling across vast geographic regions can miss these important cases. For example, different species show different distributions, levels of divergence, and levels and patterns of introgression at the major faunal break between the Indian and Pacific basins. The level of exchange between Pacific and Indian basins fluctuates during each glacial cycle, causing some lineages to diverge and then come into renewed contact (Benzie 1999a). Are these sister lineages hybridizing? Hobbs et al. (2009) and Montanari et al. (2011) demonstrated that there are focal points of hybridization. However, the Coral Triangle (Indo-Malay-Philippines archipelago) may also be the world’s largest suture zone, where clines between genetic lineages may slosh about such that no single geographic feature is the site of a distinct concordant break for most species in the region (Carpenter et al. 2011). Another important process in this region is range expansions resulting from periodic flooding of the Sunda and Sahul shelves adjacent to the Coral Triangle (Barber et al. 2002, Klanten et al. 2007, Crandall et al. 2008a, 2012a, DiBattista et al. 2012a). In particular, the more diverse and upstream Pacific assemblage seems to be expanding west and south into the Indian Ocean (Williams and Benzie 1997, 1998, Benzie 1999b, Lessios et al. 2001, Craig 2008, Drew and Barber 2009, Briggs and Bowen 2012, but with some exceptions, see Kochzius et al. 2009, Gaither and Rocha 2013). The conclusion is inescapable: dense sampling across species ranges is necessary to understand phylogeographic dynamics. Intense sampling in limited areas may uncover a faunal boundary, but the source of the genetic variation may originate beyond the sampled area.

**Don’t Search for a Gold Mine by Digging Holes**

The great promise of genomic comparisons (Berner et al. 2008, Hohenlohe et al. 2011, Rocha et al. 2013, Willette et al. 2014) and occasional perils (Roesti et al. 2012) continue to unfold. As genomic data become available, DNA segments can be parsed across selected and neutral panels of loci, and will likely reveal the main forces of local adaptation at work at the species and community levels (Bertness and Gaines 1993, Sotka 2005, Gaggiotti et al. 2009, Nielsen et al. 2009, Sanford and Kelly 2011).
However for the immediate future, it is infeasible to conduct genome-scale comparisons on hundreds to thousands of specimens. The technology is widely available, but the analytical software is not. The broad accessibility of genomic tools is creating a rush where researchers are abandoning tried and true tools in favor of novel genomic methods. However, the indiscriminate application of genomics is like randomly digging holes to find a gold mine. It is far more efficient to first find signs of gold, whether through insights from morphology, mtDNA patterns, or discordance between loci, and follow up with more comprehensive genomic approaches when emerging patterns warrant.

In all areas of wildlife genetics, there is a long history of casting off well-established methods for the next technology. The push to incorporate next-generation sequencing methods in phylogeography is leading to the systematic abandonment of previous techniques, a process endorsed by both funding agencies and editorial policies where journals prohibit single marker studies and/or impose requirements for minimal numbers of loci. Funding agencies prioritize genomic research over field studies. The NESCent working group agreed that methods should be chosen based on their cost and ability to address a hypothesis, rather than their novelty. In the medical field it is widely accepted that a thermometer or stethoscope are still invaluable tools, despite the availability of more technologically advanced tools like MRI and CAT Scanners, yet phylogeographers continually cast aside valuable tools in favor of the next big thing. Given limited resources, a data set that samples populations of many species for a single mtDNA locus may be most appropriate for hypotheses focused on comparative and community level patterns. Hypotheses should dictate methods, not vice versa.

Recent years have seen several papers challenging the utility of mtDNA, citing limited examples of selective sweeps, non-uniparental inheritance, and non-clonality (e.g., Bazin et al. 2006, Galtier et al. 2009). However, the ability to find limited exceptions to typical mitochondrial evolution and inheritance among the thousands of published studies should provide more solace than concern regarding the ability of this marker to faithfully record evolutionary history (Wares et al. 2006, Karl et al. 2012). Despite assertions to the contrary, evidence of noise and/or bias in spatial genetic patterns associated with mtDNA is not overwhelming; a review of more than 48 avian studies found phylogenetic estimates from mtDNA vs microsatellite markers were overwhelmingly comparable (Zink and Barrowclough 2008), and a recent review of nuclear-mitochondrial discordance in animals found a low level of incongruence (<18%; Toews and Brelsford 2012). Similar results from marine organisms indicate that mtDNA will often (but not always) be the first locus to show genetic structure due to its smaller effective population size (Timm et al. 2012, DeBoer et al 2014a).

It is important to recognize that one or a few loci may be a poor representation of the entire genome. However, the power to detect shared patterns in single marker studies comes from replication across species rather than multiple loci within species. The consensus among the NESCent working group is that comparative phylogeography should contain an initial survey with a simple set of genetic markers; mtDNA sequences in many cases, but also nuclear introns, microsatellites, or single nucleotide polymorphisms (SNPs) as appropriate for the questions being addressed (Fig. 1). One of the most important insights to be gained from the initial surveys is to pinpoint the regions, zones, taxa, or populations that contain the highest variance.
or extremes of the genetic spectrum for further investigation. Moreover, before detailed genomic studies of selection and local adaptation can be undertaken, it is critical to understand (and be able to model) the underlying population structure and demographic history of the population under investigation (Williamson et al. 2005, Excoffier et al. 2009).

**Moving from Simple Genetic Markers to Genomics**

These “unplugged” surveys (using single or a few genetic markers prior to genomics) across a broad range of species and geography provide signposts for further genetic explorations in select taxa and areas (Fig. 1). How does the phylogeographer then incorporate genomics into comparative geographic surveys? By focusing on the potential “gold mines” identified in initial surveys. Here we consider five general outcomes, and the genomic explorations that may be mandated by these results.

1. **Concordant Patterns Among Species.**—Concordant patterns among species indicate shared biogeographical barriers or common history (Avise 2000, Rocha et al. 2002, Hickerson et al. 2010). One of the most exciting frontiers in this respect is the
shift from single-species conservation programs to ecosystem-level management of marine resources. Comparative phylogeography can be valuable in this initiative, by defining the boundaries between marine assemblages that are shared across taxonomic and trophic categories. For example, Toonen et al. (2011) documented genetic breaks in the Hawaiian Archipelago that are shared between mollusks, corals, echinoderms, fishes, and marine mammals. The spatial scale of genetic structuring tends to be smaller in the 600-km Main Hawaiian Islands than the 1500-km Northwest Hawaiian Islands. Thus, a follow-up study (possibly using SNPs, microsatellites, or other more variable loci) would focus on heavier sampling in the main islands and sparser sampling in the remote northwest to further capture spatial patterns of structuring. Given the challenges in cost, technology, and bioinformatics still facing genomic analyses, efficiency in the number and locations of samples is crucial. With this more informed sampling design, getting appropriate coverage with genomic scale markers is more feasible. Note that concordant patterns and processes are not always obvious from a glance at a haplotype network or a set of pie charts. Crandall et al. (2012b) found that a biophysical model of larval dispersal among South Pacific archipelagos provided the best explanation for mitochondrial data in four gastropod species, despite qualitative differences in the structure and distribution of their haplotype networks.

2. Non-concordant Patterns Among Species or Loci.—Non-concordant patterns among species or loci may be the stepping off point for genome-wide studies, as well as ecological or morphological surveys. Variation in mating systems, fertilization biology, larval biology, population size, ecology, etc. all influence distributions and connectivity. The mode and scale of dispersal are known to have a broad impact on genetic structuring, speciation, and resulting diversity pattern (Paulay and Meyer 2006, Malay and Paulay 2009). Egg size, mode of development, and pelagic period can evolve rapidly and often differ among closely related species (Schulze et al. 2000, Collin 2004, Marko and Moran 2002, Hart et al. 2011, Puritz et al. 2012b). Aside from these ecological effects, the large effective population sizes of most marine species may obscure events such as recent speciation and range expansion, due to non-equilibrium allele frequencies, incomplete lineage sorting, and non-concordant patterns among loci. Under such non-equilibrium processes, the use of equilibrium models can inflate estimates of gene flow between populations or closely related species (Marko and Hart 2011). Therefore using only a few neutral loci may fail to catch the “true” patterns of species boundaries and/or population structure. Gamete recognition loci in free spawners and mating systems in internal fertilizers test and influence emerging species boundaries, impact the tempo and scale of speciation (Hellberg et al. 2000, Landry et al. 2003), and can yield different evolutionary dynamics among related taxa (Rocha-Olivares et al. 2001, Palumbi and Lessios 2005). While past studies have tested for evidence of selection at specific loci involved in fertilization (Lessios 2011), sensory perception (Seehausen et al. 2008), and bioluminescence (Feder and Velez 2009), phylogenomic approaches are allowing broad evaluation of which loci are under selection, thus permitting the evaluation of the importance of reproductive and ecological factors in speciation (Excoffier et al. 2009, Gagnaire et al. 2011, Bird et al. 2012, Deagle et al. 2012, Pespeni et al. 2012).

3. Discordance Between Molecular and Morphological Differentiation.—Discordance between molecular and morphological differentiation can provide
another appropriate starting point for genomics. Comparative studies can distinguish adaptive and non-adaptive radiation (i.e., ecological opportunity). While evolutionary radiations can be inferred from temporal patterns of diversification (exemplified by “early bursts”), adaptive radiations are associated with ecological opportunity (Losos 2010), which is often revealed through selection on morphological traits. For example, color is a very important taxonomic character in many groups, and can indicate underlying genetic differentiation (e.g., Meyer et al. 2005, Drew et al. 2008, Timm et al. 2008, Malay and Paulay 2009). However, coloration is not always a reliable taxonomic character; the brightly-colored pygmy angelfishes (Centropyge spp.) have taxonomic divisions based on color patterns that do not align with genetic partitions (DiBattista et al. 2012b). Conversely, a phylogeographic survey of bonefishes (genus Albula) using the same mtDNA marker revealed deeply divergent genetic lineages that were almost indistinguishable with external morphology (Bowen et al. 2007). A combination of these patterns is also found in the hamlet (genus Hypoplectrus) species flock in the Caribbean (Puebla et al. 2007). These cases raised evolutionary questions that can now be addressed using genome-wide techniques, but those questions might never have surfaced without the initial single marker studies.

4. Geographic Contexts of Speciation.—Comparative studies can explore the relative importance of allopatric, parapatric, and sympatric speciation, providing insights into the causes of diversification (Rocha and Bowen 2008, Bird et al. 2011, Dawson et al. 2012). For example, younger species with non-overlapping distributions on either side of a biogeographic barrier and range overlap by older species are signatures expected from allopatric speciation followed by secondary contact. Multiple species with coincident ranges bounded by barriers are also the consequence of divergence in geographic isolation (Jordan 1904, Mayr 1954). Conversely, sympatric speciation can arise anywhere within a species range and thus produce a pattern in which geographic overlap bears no relation to potential barriers. Members of the gastropod genus Lunella include sympatric species along coastlines, whereas species distributed in two-dimensional island arrays are allopatric (Williams et al. 2011). In this case, sympatric species showed greater phenotypic diversity than allopatric species, indicating that cryptic species are more likely to occur in taxa with allopatric distributions. However these expectations can be confounded by unequal rates of accumulation of reproductive isolation (Palumbi and Lessios 2005) or by rapid dispersal that erases traces of the actual point in space where new species arose; nevertheless, they are useful guides for the causes of speciation events, particularly if they are duplicated in multiple unrelated taxa. Allopatric species complexes are common in the Indo-Pacific region, and future taxonomic and biogeographic studies could benefit from focusing on low dispersal species in this region. Here again, the resolution of markers available today can provide critical insights, but genome-wide studies can provide resolution, especially in species complexes with fuzzy boundaries (Forsman et al. 2010, Nosil and Feder 2012, Radwan and Babik 2012, Stat et al. 2012, Rocha et al. 2013).

5. Origins of Marine Biodiversity.—Patterns of origination and dispersal have begun to appear through phylogeographic studies in recent years (Palumbi 1997, Eble et al. 2011, Bowen et al. 2013, Cowman and Bellwood 2013, Gaither and Rocha 2013, Crandall et al. 2014). The Center of Origin hypothesis (Ekman 1953) to explain biodiversity hotspots was criticized for lacking a plausible process for speciation within
the Coral Triangle (Jokiel and Martinelli 1992). Recent comparative phylogeography studies in this region highlight the role of Pleistocene sea level fluctuations and the Halmahera Eddy in driving lineage diversification (Barber et al. 2006, Crandall et al. 2008a, Barber et al. 2011). While these patterns are clear in some species (e.g., Barber et al. 2002, DeBoer et al. 2008, 2014b), they can be less clear in others. For example, Vogler et al. (2008) demonstrated that the crown of thorns sea star, *Acanthaster planci* (Linnaeus, 1758), is likely a species complex and multiple "species" and haplogroups are distributed across the Coral Triangle. However, it is unclear whether overlapping distributions in the Coral Triangle result from secondary contact (with potential hybridization) following allopatric speciation, or whether speciation resulted from natural selection in sympathy. Similarly surveys of coral-feeding nudibranchs across the Pacific (Faucci et al. 2007) and gastropods in the Coral Triangle (Simmonds et al. unpubl data) have revealed that genetic lineages are associated with feeding preferences, rather than geography, indicating sympatric speciation. In the above cases, mtDNA provided critical insights, but genome wide studies are essential to fully understand evolutionary histories by distinguishing between neutral and selective divergence and testing for introgression.

**Ecologically-grounded Hypotheses**

The five general patterns described above (discordance between loci or data sets, geographic distributions and origins, species dispersal) provide not only signposts toward further exploration of the genome, they also can send the perceptive researcher back into the water for additional sampling and study. To understand the complex process of speciation in the sea, it is important to distinguish between populations that are differentiating as a result of selection (potentially despite exchanging migrants) from those that differentiate primarily through genetic drift in geographic isolation. Marine phylogeography has a traditional emphasis on geography, but increasingly the study design is focused on ecological partitions (Munday et al. 2004, Rocha et al. 2005, Choat 2006, Selkoe et al. 2010). When there are evolutionary issues such as ecological divergence between closely related species, there is good reason to explore further into the genome as well as the ocean.

Genomic scale studies will enable us to ask questions about the feedbacks between ecological and evolutionary phenomena. For instance, the Marquesas are especially interesting for genetic studies, as they have both oceanographic isolation and ecological differences from the adjacent Society Islands, including upwelling, variable temperatures, and little coral cover (Randall 2001, Hickerson and Meyer 2008). The Marquesas are also genetically distinct from the Society Islands in six of eight species surveyed to date (Planes and Fauvelot 2002, Gaither et al. 2010, Crandall et al. 2012, Szabo et al. 2014) and harbor the third highest endemism rate in the Pacific Ocean (Briggs and Bowen 2012). Is geographical concordance between population genetic and biogeographic evolutionary patterns the norm? In other words, do areas with high levels of genetic diversity and structure tend to have high rates of endemism? Are the same ecological drivers behind both phenomena? These are exciting frontiers in marine phylogeography. The availability of genomic sequencing combined with newly available environmental databases for marine ecosystems (Sbrocco and Barber 2013) makes it possible to simultaneously examine the variability of neutral and selected regions of the genome across geographic and ecological parameters.
Combining Evolution and Ecology to Understand Species Richness Patterns

Traditionally, studies either have explored temporal patterns of diversification from a biogeographic perspective, or spatial patterns in relation to parameters such as environment, area, and range. Wiens and Graham (2005) proposed a niche conservatism model based on phylogenetic history to explain contemporary patterns of species richness, thereby making a case for combining ecology and evolution. Rahbek et al. (2007) developed models combining range and environment in predictions of richness patterns at continental scales, using a bird data set from South America. Gotelli et al. (2009) presented a theoretical framework for modeling richness, combining evolution and ecology. Such models require fine scale understanding of species boundaries and range, as well as diversification rates, which can only be inferred from phylogeographic and phylogenetic data.

Conclusion: Not the End of History

The historian Francis Fukuyama famously remarked that the end of the Cold War could be the end of history, based on a perceived stabilization of the world order (Fukuyama 1992). Karl Marx made a similar claim in 1859 based on the presumed supremacy of communism. It did not happen then, and it will not happen in phylogeography now. However, the field may be reaching a more stable state after two decades of swiftly changing approaches. With entire genomes in hand, we will soon have the capacity to evaluate the relative power and utility of different markers, and determine which loci are the key drivers of marine diversification.

The optimal study design we endorse for comparative phylogeography requires the surveys of hundreds of specimens, range-wide coverage where possible, across 10, 20, or more species. This should include a field effort commensurate with the issues, an initial survey with simple genetic markers such as mtDNA, microsatellites, or SNPs, followed by judicious applications of genomics when necessary. Before gigabases of population genomic data are collected, it is essential that the investigator consider the hypotheses being tested, as well as the data requirements and limitations of the model that will be used for inference. Population (and management) level issues for individual species are amenable to applications of SNPs and microsatellites, evolutionary issues can be pursued at the appropriate level of nDNA analysis. If the question is phylogenetic distinctiveness, fewer loci are needed than in the quest for loci under selection.

An obvious and valid criticism of using mtDNA for an initial sweep is that it would miss some evolutionary patterns present in the nuclear genome (Eytan and Hellberg 2010, DiBattista et al. 2012b). As noted earlier, while the caveat of mitochondrial-nuclear discord is routinely raised as a criticism of single-locus studies, such cases are rare (Karl et al. 2012, Toews and Brelsford 2012). While introgression and paralogy can impact mtDNA analyses as that of any single marker analysis, error rates in most groups are low (<2% in large mollusk data sets; Meyer and Paulay 2005), with perhaps the exception of taxa prone to nuclear pseudogenes of mitochondrial origin (NUMTs; Williams and Knowlton 2001). Furthermore, the pitfalls of mtDNA studies are fairly well known and can usually be recognized and tested, while the drawbacks of nuclear or genomic approaches are still being discovered. Therefore,
mtDNA still provides the most versatile tool for taxonomically and geographically broad-scale studies because of its ease of amplification across taxa, level of resolution, and low cost. The more broadly used nuclear markers often do not distinguish closely related species (Andrews et al. 2013), while microsatellites lack phylogenetic breadth in application. A major promise of genomic approaches is developing access to an abundance of nuclear markers (Etter et al. 2011, Faircloth et al. 2012, Lemmon and Lemmon 2012, Rocha et al. 2013), perhaps with the eventual goal of a core set of nuclear markers applicable across a wide range of species.

Single locus studies offer another advantage in providing standardized DNA sequence data sets that can be shared between laboratories, and augmented with new data. The sea turtle research community has exploited this advantage to allow comparisons between new and existing data sets based on sequences of the mtDNA control region. When new haplotypes are reported, the sequences are archived in a database maintained by the Archie Carr Center for Sea Turtle Research (http://accstr.ufl.edu/resources/mtdna-sequences/). Researchers can use this archive to link haplotypes observed on nesting beaches to haplotypes observed in distant feeding aggregates, greatly accelerating the scientific foundations for protection in international waters (Bowen and Karl 2007, Bjorndal and Bolten 2008). For additional examples of pre-existing data sets, see Keyes et al. (2014). Comparable DNA data sets also allow the barcoding that has emerged as a powerful tool, using the same marker in many taxa permits matching of unknowns against a library of accumulated sequences (Hebert et al 2003). Emerging metagenomic approaches are now taking advantage of such barcode libraries to evaluate community diversity (e.g., Leray et al 2013).

Another major reason for keeping single locus markers in the phylogeographer’s tool kit is to advance marine phylogeography regardless of economic conditions (Barrowclough and Zink 2009). The vast majority of the world’s biodiversity hotspots are centered in developing countries (Roberts et al. 2002) that have limited funding for basic research. While next generation sequencing methods will undoubtedly decrease greatly in cost, it may still be cost-prohibitive in many science communities. Simple mtDNA studies are still effective in addressing key evolutionary and conservation questions (DeBoer et al. 2014a).

Finally, harking back to our roots in morphology and field observations, the phylogeographer should not forget the value of knowing the organisms intimately. This has been borne out countless times in our field, as the astute marine phylogeographer sees a known organism in a new location for the first time, and says “that looks different”. That “different” organism is supposed to be the same species as elsewhere but has a new shape or color, a distinctive behavior, or a different host (Rocha 2004, Faucci et al. 2007, DiBattista et al. 2013). Phylogeographic studies based on such observations are fueling a revolution in the discovery of cryptic marine diversity, and providing foundations of further evolutionary studies (Gittenberger and Gittenberger 2011, O’Loughlin et al 2011). Good naturalists also notice differences in the diversity, distributional, or abundance trends across environments and taxa, which can guide selection of species targeted to test particular hypotheses. For example the rarity and low diversity of *Calcinar* hermit crabs or certain land crabs in the Coral Triangle indicates that ecological restriction to oceanic regions is an important selective force in some taxa that can lead to inverse patterns of diversity (Malay and Paulay 2010, Paulay and Starmer 2011). Indeed it is fascinating when these field observations lead to the descriptions of new species or novel evolutionary lineages. However, it can be
equally interesting when the differences observed in the field do not yet translate into genetic partitions at neutral loci (DiBattista et al. 2012b, Baums et al. 2013). Regardless of the outcome, a primary challenge of our day is to fit the plethora of genomic data into phylogeographic conclusions. The emphasis on genomics is apparent in granting trends, as funding has shifted yet again. Our emphasis here is on the less appreciated truth that a strong field component, and adequate sampling, is the universal foundation for robust marine phylogeography.

Acknowledgments

We thank the National Evolutionary Synthesis Center at Duke University for hosting the meeting “Molecular Ecology and Evolution of the Indo-Pacific” that produced this paper. That meeting was funded by a National Science Foundation grant to ED Crandall and C Riginos. We thank special editor C Riginos, U Ramakrishnan, and an anonymous reviewer for comments that improved the manuscript.

Literature Cited


Forsman ZH, Concepcion GT, Haverkort RD, Shaw RW, Maragos JE, Toonen RJ. 2010. Ecomorph or endangered coral? DNA and microstructure reveal Hawaiian species


