The role of molecular genetics in sculpting the future of integrative biogeography
Brett R. Riddle, Michael N. Dawson, Elizabeth A. Hadly, David J. Hafner, Michael J. Hickerson, Stacy J. Mantooth and Anne D. Yoder
Progress in Physical Geography 2008; 32; 173
DOI: 10.1177/0309133308093822

The online version of this article can be found at:
http://ppg.sagepub.com/cgi/content/abstract/32/2/173
The role of molecular genetics in sculpting the future of integrative biogeography

Brett R. Riddle,1* Michael N. Dawson,2 Elizabeth A. Hadly,3 David J. Hafner,4 Michael J. Hickerson,5 Stacy J. Mantooth1 and Anne D. Yoder6

1School of Life Sciences, University of Nevada Las Vegas, 4505 Maryland Parkway, Las Vegas, Nevada 89154-4004, USA
2School of Natural Sciences, University of California at Merced, 5200 North Lake Road, Merced, California 95344, USA
3Department of Biology, Stanford University, Stanford, California 94305-5020, USA
4Biosciences Department, New Mexico Museum of Natural History, 1801 Mountain Road NW, Albuquerque, New Mexico 87104, USA
5Biology Department, Queens College, CUNY, 65030 Kissena Boulevard, Flushing, New York 11367-1597, USA
6Department of Biology, Duke University, Durham, North Carolina, USA 27706

Abstract: We review the expanding role of molecular genetics in the emergence of a vibrant and vital integrative biogeography. The enormous growth over the past several decades in the number and variety of molecular-based phylogenetic and population genetics studies has become the core information used by biogeographers to reconstruct the causal connections between historical evolutionary and ecological attributes of taxa and biotas, and the landscapes and seascapes that contain them. A proliferation of different approaches, sequences, and genomes have provided for the integration of a ‘biogeography of the Late Neogene’ with other Earth and biological sciences under the rubrics of phylogeography, landscape genetics, and phylochronology. Approaches designed explicitly to take advantage of unique properties of molecular genetic information have led to the re-emergence of dispersal as an analytically tractable process that historical biogeographers can now use, along with vicariance, to reconstruct the geographical context of diversification. Concomitant with the expanding amount of information available, molecular data sets often provide for estimates of lineage divergence dates, and analytical tools for doing so continue to improve. The comparability of molecular-based estimates of phylogenetic and population genetic histories across non-related...
taxa has stimulated deployment of new methods to test for spatial and temporal congruence across co-distributed taxa and ecosystems, and thus increased rigour in hypothesis-testing. We illustrate how a molecular genetics framework has provided robust and novel reconstructions of historical biogeographical pattern and process in three different systems, and finish with some thoughts on the role a molecular genetic-based biogeography will play in predicting alternative futures of biodiversity.

**Key words:** biodiversity, congruence, dispersal, historical biogeography, phylochronology, phylogeny, phylogeography, vicariance.

### Introduction – the emergence of a dynamic historical biogeography

Biogeography aims to decipher the geography of speciation, dispersal, and extinction of lineages and clades and, in combination with increasing knowledge of Earth history, to reconstruct the assembly and disassembly of biotas through time. The vast scope of biogeography requires integration across an equally vast array of disciplines – including ecology and evolutionary biology, systematics and phylogenetic biology, population biology and demography, organismal physiology and functional biology, geology, biogeochemistry, climatology, and all of palaeontology. Indeed, an increasingly integrative and revitalized biogeography (Arbogast and Kenagy, 2001; Donoghue and Moore, 2003; Richards *et al.*, 2007; Roy and Goldberg, 2007) is poised to provide a critical service to society. This integrative biogeography is uniquely positioned to make robust predictions about the consequences of human impacts on Earth’s biological diversity, including invasive species, habitat fragmentation and destruction, and global climate change (Whittaker *et al.*, 2005) in ways that cut across more typical species or area-based studies and which promise to reveal insights not otherwise possible.

To a large extent, historical biogeography owes its renewed vitality to the molecular genetics revolution in systematics and population genetics that began with the proposition that neutral molecular markers evolved in a clock-like fashion (Zuckerkandl and Pauling, 1965). The molecular revolution in biogeography received an enormous infusion in the 1990s when phylogenetic systematics and population genetics were conjoined to create phylogeography (Avise *et al.*, 1987), ‘a field of study concerned with the principles and processes governing the geographical distributions of genealogical lineages, especially those within and among closely related species’ (Avise, 2000). A February 2008 topic search of the ISI Web of Science using ‘phylogeograph*’ revealed 4,800 studies to date with phylogeography as a topic. With the advent of phylogeography, we can explore recent and ongoing distributional and demographic changes within contemporary populations, including the processes of range contraction and expansion, hybridization and introgression, and metapopulation structure.

Roughly coincident with the origins of molecular systematics, the emergence of the theory of plate tectonics (Dietz, 1961; Hess, 1962) provided historical biogeography with a testable model of causal association between Earth history and the geographical history of distribution and divergence in plants and animals (Brundin, 1966), which motivated the development of vicariance biogeography (Nelson, 1974; Platnick and Nelson, 1978; Rosen, 1978). Vicariance biogeography is an approach and method highlighting the passive transport of species and biotas to disparate reaches of the Earth on drifting continents and that identifies divergence between lineages as a function of passive separation via Earth events such as mountain-building. A basic premise is that, under a vicariance model, a suite of species or other taxa sharing a common geographical distribution (hereafter, ‘co-distributed species/taxa’) would demonstrate
congruent spatial patterns of isolation and divergence. Vicariance biogeography was conceived as a plausible alternative to dispersalist biogeography, the approach that prevailed prior to the 1970s (Brundin, 1966). The vicariance biogeographers sought to rid historical biogeography of the non-testable ‘just-so stories’ of the dispersalists by grounding the discipline conceptually within a framework that fused Croizat’s (1964) goals of searching for general patterns of distribution with Hennig’s (1966) method that emphasized the critical importance of monophyletic groups in phylogenetic systematics and biogeography (Croizat et al., 1974). Thus, the state of biogeography prior to molecular genetics involved a tension between dispersal and vicariance as hypotheses behind the distribution of diversity, often with the dating of geological events feeding, rather than solving, the debates (for an early example, see Darlington, 1965; Brundin, 1966).

Yet, as we have entered the twenty-first century, a rapidly expanding set of molecular phylogenies is providing historical biogeographers with compelling reasons to believe that dispersal has played a role at least equivalent to vicariance in the history of life on Earth – even across those iconic taxa and biotas inhabiting Gondwanan landmasses that have for several decades been considered exemplars of a vicariance-driven system (Sanmartín and Ronquist, 2004; de Queiroz, 2005; McGlone, 2005; Yoder and Nowak, 2006; Barker et al., 2007). Thus, we have entered yet another stage in the evolution of historical biogeography, fuelled by a new generation of goals and methods motivating progress in the discipline, but with much of the same rancor (eg, Brooks et al., 2004; Siddall, 2005). Unlike the upheaval of the 1970s, however, these conceptual and analytical developments are emerging in parallel along two tracks: on the one hand, within a more traditional area and taxon-based historical biogeography (eg, Ronquist, 2003; Ree et al., 2005; Wojcicki and Brooks, 2005; Ree and Smith, 2008); and, on the other hand, within the increasingly sophisticated realm of single-taxon or comparative phylogeography (eg, Knowles, 2004; Templeton, 2004; Lapointe and Rissler, 2005; Hickerson et al., 2006; 2007; Kidd and Ritchie, 2006; Carstens and Richards, 2007). Both directions are fuelled by increasing focus on dating palaeoenvironmental events and by major advances in acquisition and analysis of molecular genetic data.

Our goal in this paper is to explore the current and expanding role of molecular genetic information and molecular approaches to biogeography. We begin with a brief review of its historical underpinnings, follow with an overview of advances in key theoretical and analytical issues, and then present a series of ‘case studies’ from both marine and terrestrial biospheres. We aim to illustrate how a molecular-based biogeography is providing a framework for resolving long-standing and unresolved questions in biogeography by integrating across time and space, by motivating novel questions and approaches, and, perhaps most importantly, by giving insight into probable future changes in biodiversity on our planet.

II The molecular revolution in biogeography

I From albumin to phylogenomics
The molecular revolution in biogeography began rather humbly with the advent of an immunological approach to estimating the underlying genetic variation coding for amino acid differences in a single protein, albumin (Sibley, 1970). Early studies using this method addressed island biogeography in Galapagos iguanas (Higgins, 1977), invasion versus speciation models in western Australian heleioporid frogs (Maxson and Roberts, 1984), Tertiary versus Quaternary speciation models in several Australian anuran species (Roberts and Maxson, 1985),...
and biogeographical histories in a wide range of vertebrates (Gill, 1976; Everaarts et al., 1983; Schill and Dorazio, 1990).

Protein electrophoresis (i.e., allozymes) improved on albumin immunology by providing a cost-effective approach to surveying genetic variation across multiple nuclear markers (Avise, 2004), and resulted in advances in population genetic theory because allozyme data assay Mendelian markers (Avise, 2004). Many allozyme studies have employed phylogenetic and population genetic analyses to infer both intra and interspecific evolutionary and biogeographical histories in animals, plants, and fungi (e.g., Spieth, 1975; Patton and Yang, 1977; Sites and Greenbaum, 1983; Turner, 1983; Hofman et al., 2007; Izawa et al., 2007; Pauly et al., 2007).

The transformation from indirect to direct assay of genetic changes along a DNA sequence came with the discovery of restriction enzymes (Nathans and Smith, 1975) that digest DNA at 4, 5, or 6 base-pair ‘recognition sites’. Hundreds of enzymes are now available to digest DNA, and combinations of them can be used to create DNA fragments that vary in length according to presence or absence of different restriction sites. Multi-enzyme restriction profiles can be generated quickly, producing large amounts of data assaying whole genome base-pair mutation differences between individuals, populations, species, and higher taxa. Following the pioneering study of a species of pocket gopher (Geomys) by Avise et al. (1979), biogeographers were quick to adopt this approach for testing biogeographical hypotheses (e.g., Kessler and Avise, 1984; Bermingham and Avise, 1986; Honeycutt et al., 1987; Riddle and Honeycutt, 1990; Bernatchez and Dodson, 1991; Lovette et al., 1999; Kuchta, 2007).

The advent of the polymerase chain reaction (PCR; Mullis et al., 1986), coupled with the development of chain-terminating DNA sequencing (Sanger et al., 1977), led to the ability to sequence DNA quickly and efficiently, further revolutionizing molecular biogeography. With the availability of ‘universal’ oligonucleotide primers that were useful across a broad range of animal (Kocher et al. 1989) and plant (Taberlet et al., 1991; Demesure et al., 1995) taxa, the mitochondrial cytochrome b gene in animals and chloroplast introns in plants became among the most extensively sequenced molecular markers. The range of substitution rates across mitochondrial genes in animals and chloroplast introns in plants enables their use over a broad range of time-frames (Demesure et al., 1995; Johns and Avise, 1998). A variety of PCR-based techniques now provide ways to assay variation across the nuclear genome, including micro-satellites, amplified fragment length polymorphisms (AFLPs), short interspersed elements (SINEs), and single nucleotide polymorphisms (SNPs) (reviewed in Avise, 2004).

A molecular-based biogeography is now poised to enter a new era with the increasing numbers of whole genomes available for analysis, although few have yet been used for this purpose. The emerging field of phylogenomics is a result of this explosive data boom, representing the study of complete genomes from many closely related species using an evolutionary perspective (Delsuc et al., 2005). For example, Macey (2005) explored the origin of plethodontid salamanders in North America with 27 complete mitochondrial genomes. The potential for a phylogenomics approach to disentangle stubborn phylogenetic and biogeographical relationships can be illustrated with two recent studies that have provided support for a Gondwanan clade of placental mammals (Atlantogenata) distinct from a northern clade (Boreoeutheria) – consistent with the tectonic split of Pangea into southern (Gondwana) and northern (Laurasia) continents about 175 million years ago (Ma). In order to resurrect robust support for Atlantogenata, Hallström et al. (2007) aligned 2,168,859 nucleotides from about 2,840
genes, and Wildman et al. (2007) aligned 1,443,825 nucleotides from 1,698 genes! For many, the ability to explore these data with the current phylogenetic reconstruction methods (e.g., computer algorithms and resources) may be a limiting factor (Delsuc et al., 2005), and the development of innovative analytical approaches will be required in order to take maximum advantage of these enormous data sets (e.g., Edwards et al., 2007; Li et al., 2007).

2 The phylogeographical expansion of biogeography

While classic biogeography rooted in plate tectonics has been reconstructing the overarching assembly and disassembly of continental biotas, phylogeography has been building up the history of species from the ground floor. Indeed, the focus of phylogeography on lineages within species and between closely related species serves to establish a temporal context for phylogeographical studies, which are almost completely constrained within the Neogene, and with the majority focused on the Quaternary (Riddle and Hafner, 2006a). Putatively neutral genetic markers such as mitochondrial DNA (mtDNA) in animals and chloroplast DNA (cpDNA) in plants have frequently been used to estimate the timing of events within lineages and provide an independent assessment of the age of the event relative to geographical and geological history. Although much of phylogeography has been orientated toward the discovery of the geographical distributions of single gene lineages, many studies now are also incorporating data from multiple nuclear loci in attempts to resolve the problem of ‘gene tree versus species tree’ discordance (Edwards and Beerli, 2000; Carstens and Knowles, 2007).

Whereas phylogeography is often considered to be strictly a study of intraspecific pattern and process, one of the important insights from two decades of phylogeographical research has been the high frequency at which ‘cryptic’ temporally deep and geographically structured evolutionary lineages – often suggesting presence of several distinct species – are embedded within a single phenotypically conservative species. Furthermore, discovery of ‘cryptic’ gene tree structure, frequently accompanied by crude molecular clocks to estimate divergence times, allows for postulation of plausible geological and palaeoclimatic events in Neogene Earth history as causal agents of distribution and divergence patterns within species and among closely related species.

From its inception, the strength of phylogeography (Avise, 2000) has been its capacity to provide a connection between conceptually and analytically separate views of history – lying at the junction between the population genetics of microevolutionary processes (e.g., historical population size and structure, migration rates, gene tree coalescence times [the point in time when two allelic lineages diverged from an ancestral lineage]) and the phylogenetics of macroevolution (e.g., the geography of speciation and assembly of biotas). This ‘micro versus macro’ duality is reflected in the incorporation of population genetic, phylogenetic, or some combination of both approaches, which results in a challenge to phylogeographers in choosing from the broad spectrum of methods available (see, for example, Kidd and Ritchie, 2006). However, the expansive nature of phylogeography positions it as a bridge between what have become two essentially separate foci in biogeography that occupy often disparate spatial and temporal regimes.

On the one hand, a biogeography of the Late Neogene has emerged to address questions such as locations of and expansion of populations, species, and biotas from refugia following climate change, often while invoking explicit ecological influences (e.g., palaeoclimatically induced range shifts). Biogeography of the latest Holocene now extends into the temporal realm of the past few decades with the emergence of landscape genetics (Manel et al., 2003; Storfer et al., 2007). On the other hand,
the more time-honoured realm of historical biogeography frequently references pre-
Neogene geological processes such as drifting continents, emerging islands, and land bridges. Here, the questions have been more focused on biogeographical phenomena such as the locations, originations, composition, and relationships between areas of endemism, the locations of and expansion of lineages and biotas away from ancestral areas, and the relative importance of dispersal versus vicariance in shaping Earth’s biotas (Riddle and Hafner, 2006a). Much progress is needed to merge these two arenas into a more unified, molecular-based historical biogeography (Riddle and Hafner, 2006a; 2006b).

3 Phylochronology: integration of ancient genetic data with modern genetic variation

Detailed evidence of species evolutionary responses to particular environmental events is lacking and controversial (Vrba, 1993; Alroy et al., 2000; Barnosky, 2001), yet important for reconstructing the biogeographical history of lineages, communities, and biomes. Phylogeographical reconstructions of Neogene events have done much to inform biologists about the age of species yet, surprisingly, have revealed how rarely modern species divergences are tied to profound recent global environmental events such as Pleistocene glaciation (eg, Hewitt, 2000). One limitation of contemporary single-species phylogeography is that modern genetic data alone do not permit an investigation of the individuals that inhabited that location through time, which would provide a dynamic between environmental events and the evolutionary potential of species. Very different historic demographic events, such as dispersal or population growth, can yield identical population genetic signatures. Consequently, we cannot always determine if the pattern of diversity we detect is a result of intra- or interspecific population processes (McCaughley, 1991; Templeton et al., 1995). In addition, we cannot reliably tie the inferred population history to the local environmental history, and thus we have only limited empirical evidence of genetic response of populations to perturbations such as climatic change (Hoffman and Blows, 1993; Bradshaw and Holzapfel, 2008), the significance of which is heightened with ongoing global warming. A means to decipher the relative roles of environmental change, vicariance, and dispersal is to use a comparative multispecies phylogeographical approach (ie, Schneider and Moritz, 1999; Hewitt, 2004; Kirchman and Franklin, 2007) harking back to the fundamental premise of vicariance biogeography. When many species show similar geographical and temporal discontinuities in the genetic diversity of populations, an environmental event, dispersal, or emergence of a barrier can be inferred (Avise et al., 1987; Riddle, 1995; Arbogast and Kenagy, 2001).

An alternative, but powerful, approach is one that uses serial genetic data such as historic or ancient DNA within a single species. By using serial sets of fossil genetic data within a locality it is possible to reconstruct the history of a population through time (eg, Hadly et al., 2004; Shepherd et al., 2005). The advantage of this approach is that we glimpse the dynamics of populations as they evolved. We can measure how population bottlenecks (Paxinos et al., 2003; Shapiro et al., 2004; Chan et al., 2005), population expansions (Barnes et al., 2006), gene flow between populations (Hadly et al., 1998; Belle et al., 2006), and isolation (Hadly et al., 1998; Leonard et al., 2000) influenced the signatures of genetic diversity within species on a landscape. However, ancient genetic data are not easy to collect (Handt et al., 1994), and this type of approach is often limited to non-recombining parts of the genome such as mtDNA. Enzymes that degrade DNA begin working immediately upon death of an organism, cutting up lengths of DNA and decreasing the copy number rapidly (Kelman and Moran, 1996; Hebsgaard et al., 2005). Only under particular conditions (low humidity, low temperature, and low acidity)
will the DNA be preserved (Hofreiter et al., 2001; Smith et al., 2003). Eventually DNA will degrade even under optimal conditions, so that few fossils preserve DNA older than 100,000 years (Lindahl, 2000; Willerslev and Cooper, 2005).

Recent interest in historic genetic data preserved in museum specimens, usually confined to the past century, has facilitated focus on species responses to environmental change (Villablanca, 1994; Bouzat et al., 1998; Austin and Melville, 2006; van Tuinen et al., 2008). When ancient data describing thousands of generations are acquired and analysed using powerful new modelling tools such as the serial coalescent (Rodrigo and Felsenstein, 1999; Drummond et al., 2005), remarkable insights into the history of populations can emerge (van Tuinen et al., 2004). For example, using Bayesian serial coalescent modelling and 10,000 years of genetic data, Chan et al. (2006) reconstructed the most likely timing and magnitude of an extreme bottleneck in tuco-tucos of South America as coincident with a massive Andean volcanic eruption and ashfall (Villarosa et al., 2006), providing a possible cause for this vicariant event and population bottleneck that determined the evolution of this species.

The major limitation for phylochronology is availability of temporal data (Ramakrishnan et al., 2005). Well-dated localities preserving serially stacked fossils are uncommon and their specimens are few. These rare palaeontological sites that preserve the past 100 to 100,000 years are therefore extremely valuable snapshots of evolutionary history for many species.

III Estimating taxon divergence times and inferring biotic responses to Earth history

1 Estimating times of organismal divergence

Historical biogeography will be most effective as a means both for generating and testing hypotheses if methods that allow for accurate calculation of organismal divergence dates can be incorporated within its methodological framework. Precise estimates of clade ages permit ecological and evolutionary investigation on a scale not allowed by relative (ie, hierarchical) age estimates. For example, only absolute dates of divergence among lineages allow investigators to draw firm conclusions about the historical effects of climatological and geological conditions on patterns of speciation and dispersal among organisms. Absolute age estimates can also permit more subtle measures such as the estimation of rates of morphological and molecular evolution and their fit to the predictions of ecological and evolutionary theory. Accordingly, these measures are essential for critically testing models of speciation within a historical biogeographical framework. When an investigator observes that the estimated times of divergence for a given clade (or better yet, multiple clades) correlate with known climatological or geological events, that investigator has reason to postulate causality. Correlation does not mean causation, however. Therefore, correlations between divergence age and geological conditions should be interpreted as tentative hypotheses to be subjected to further external testing with additional data or taxa.

Although methods for estimating time are still in their infancy, tremendous strides have been made within the past decade. Bayesian (Thorne et al., 1998; Kishino et al., 2001; Thorne and Kishino, 2002; Yang and Rannala, 2006) and likelihood (Yang and Yoder, 2003; Yang, 2004) methods have been developed and modified such that they can account for violations of the molecular clock and for uncertainties of the fossil record by incorporating multiple calibration points within a single analysis. Most recently, these methods have been refined to incorporate data partition heterogeneity in genetic parameters (Thorne and Kishino, 2002; Yang and Yoder, 2003), an innovation allowing the investigator to combine different gene loci in a single analysis while accounting for their idiosyncratic patterns of evolutionary change.
In brief, methods for estimating divergence time are evolving and improving at an accelerated rate (Drummond et al., 2006; Rutschmann, 2006; Rutschmann et al., 2007; Cutter, 2008). We predict the coming years will witness an explosion of empirical studies that utilize an explicit temporal framework for investigating biogeographical phenomena, much as the refinement and improved understanding of phylogenetic methods (Hillis et al., 1994) did for evolutionary and other biological studies.

2 Testing for spatial and temporal congruence

Testing for spatial and temporal congruence across different co-distributed taxa is of central importance to biogeography (Lessios, 1998; Wen, 1999; Avise, 2000; Barraclough and Nee, 2001). Under the comparative phylogeographical approach, practitioners aim to use population genetic and phylogenetic data sampled across the ranges of co-distributed species in order to extract the causal processes behind species diversification, distribution, and community assembly (Bermingham and Moritz, 1998; Avise, 2000; Arbogast and Kenagy, 2001). However, the most detailed and rigorous tests of alternative models of speciation rely on estimates of population divergence times across sister-taxon pairs. Most recently, these models are beginning to be integrated with georeferenced ecological and morphological information to test alternative spatial models of climate-driven and ecologically deterministic community diversification and distribution (Hugall et al., 2002; Graham et al., 2004; Carstens et al., 2005; Galindo et al., 2006; Kidd and Ritchie, 2006; Carstens and Richards, 2007; Kozak and Wiens, 2007; Richards et al., 2007; Waltari et al., 2007; Kozak et al., 2008).

Given the importance of divergence time estimates in comparative phylogeographical studies, rigorous methods using modelling approaches that incorporate the coalescent variation among and within lineages are needed to prevent misleading inferences about community history (Arbogast et al., 2002). However, this undertaking is an analytical challenge because biogeographical systems involve a large number of variables and a wide array of hypotheses and models that conform to no generality with regard to model complexity, parameterization, and sampling. Tackling these analytical challenges will require a computationally powerful and flexible interpretive framework that can infer large-scale patterns of spatial and temporal congruence across co-distributed taxa, while incorporating the expected idiosyncratic within-taxon variation in historical demographic parameters (e.g., migration, population sizes, and divergence times) – unfortunately, empirically relevant demographic data are scarce. In special cases, using phylochronological data to validate processes inferred from only modern data may improve simulation techniques, or at least better circumscribe their limitations.

a Temporal congruence: Estimating lineage divergence times with an appropriate statistical model derived from coalescent theory (Tajima, 1983; Knowles and Maddison, 2002) can be accomplished using a variety of well-established approaches (Rannala and Yang, 2003; Yoder and Yang, 2004; Edwards et al., 2007; Hey and Nielsen, 2007). Analysing data from multiple co-occurring species pairs that vary with respect to demographic parameters and pairwise coalescent times is currently less straightforward. One approach would be to conduct an independent analysis on every population-pair and test the hypothesis of temporal congruence based on a resulting set of independent parameter estimates of divergence time. Alternatively, large-scale comparative phylogeographical studies can employ a hierarchical model such as an ‘approximate Bayesian computation algorithm’ (ABC; Hickerson et al., 2006) that can estimate ‘hyperparameters’—representing amalgamations of ‘subparameters’ such as divergence time, migration, and current, ancestral, and founding effective population sizes — that characterize processes across the
co-distributed taxa, such as the variance in divergence times. An advantage to the latter approach is that it can explicitly incorporate uncertainty and variation in the subparameters that independently describe the demographic history of each population pair (Hickerson et al., 2006; 2007). Such hierarchical models can account for both stochastic coalescent and mutational variance, as well as being flexible enough to incorporate and parameterize with uncertainty the biological differences among taxa. Although such a model can involve an enormous number of parameters, a hierarchical model that utilizes an ABC approach can test for temporal congruence without the prohibitive computational burden of explicitly expressing and calculating the likelihood of the model (Beaumont et al., 2002; Chan et al., 2006). Studies that have used hyperparameter estimates to test for temporal congruence include Hickerson et al. (2006), Leaché et al. (2007), and Topp and Winker (2008).

b Spatial congruence: One of the remaining analytical challenges is the interpretation of genetic breaks within species and whether such breaks arise from geographical breaks. Although genetic breaks within species are often found to span putative geographical barriers, computer simulations have demonstrated that such breaks can emerge as stochastic byproducts of the spatial coalescent process (Irwin, 2002; Kuo and Avise, 2005). Such patterns should be interpreted with caution, and causal associations between geographical barriers and genetic breaks should be verified by spatial concordance across loci within species (Kuo and Avise, 2005) and across co-distributed species (Avise et al., 1987). A rapidly expanding approach to testing for spatial congruence across co-distributed taxa involves the integration of population genetic data with georeferenced ecological information collected from both the genetically sampled individuals as well as all georeferenced specimens available for the taxon. While the genetic methods are simply those being developed as described in the phylogeography section above, the novelty of this approach comes from the increasing capabilities to build ecological niche (also called spatial distribution or bioclimatic) models of taxon responses to past (eg, glacial maximum) and predicted future climatic changes. Although the methods for this integrative approach are in their infancy, a growing number of studies have recently emerged (above citations). One of the remaining challenges will be to develop spatially explicit population genetic models (Irwin, 2002; Novembre et al., 2005; Wegmann et al., 2006) so that georeferenced specimen data can be fully integrated with the genetic information.

c Future directions: testing for congruence in 'whole ecosystems': Using genetic markers to resolve deep-seated questions about how climate change drives community assembly and evolution of whole biotas was one of the original objectives of phylogeography generally (Avise et al., 1987) and comparative phylogeography in particular (Bermingham and Moritz, 1998; Avise, 2000; Arbogast and Kenagy, 2001). However, this grand objective has so far been unrealized because comparative phylogeographical studies rarely involve more than a handful of co-distributed species. Although continuing development of methods based on coalescent theory (Beaumont, 2004; Anderson et al., 2005; Beerli, 2006; Carstens and Richards, 2007; Edwards et al., 2007; Hey and Nielsen, 2007; Hickerson et al., 2007; Knowles and Carstens, 2007) and niche-modelling (above citations) are reinvigorating phylogeography, nothing yet in comparative phylogeography has reached the scale of other broad and ambitious initiatives that grew from the 1990s PCR revolution such as assembling the Tree of Life, the Barcode of Life and CIPRES (Cyberinfrastructure for Phylogenetic Research). However, the field of comparative phylogeography is about to explode as collecting DNA sequence data across a
wide diversity of co-distributed taxa scales up to the level of comprehensive ecosystem sampling. Such community-scale comparative phylogeographical, or community genetics, data sets could potentially test classic biogeographical hypotheses (eg, vicariance versus dispersal) at the community level (Carlquist, 1966; Rosen, 1978; Yoder and Nowak, 2006), as well as test controversial and fundamental hypotheses in community ecology such as Hubbell’s Neutral Theory (Hubbell, 2001), Tilma’s stochastic competitive assembly model (Tilma, 2004), and Diamond’s niche assembly rules (Diamond, 1975; Gotelli and McCabe, 2002).

Given this exciting prospect, there is a need for developing and deploying coalescent-based models that are flexible enough to be used across a broad array of idiosyncratic biogeographical contexts. These models should also be capable of estimating hyperparameters of interest such as multi-species patterns in range expansion, colonization, and divergence while also being able to incorporate uncertainty and variance in within-species demographic parameters. Finally, these models should be capable of testing highly parameterized hypotheses without prohibitive computational burden given data sets of 50–500 co-distributed species or populations. Hierarchical ABC models can satisfy these features and thereby potentially allow researchers to address computationally challenging biogeographical questions.

IV Illustrating the value of molecular genetics in biogeography

In this section, we chose three separate systems – a large oceanic island known for its wonderfully unique endemic animals and plants, a continental biota that has diversified in a geologically dynamic theatre, and the oceans of the world – to provide a range of examples illustrating the often novel results and added clarity that have given molecular genetics a pre-eminent place within a modern, integrative historical biogeography. In each case, the molecular perspective is not only revealing previously unknown patterns or demonstrating contested hypotheses in the structure of biotas at a variety of spatial and temporal scales, but is also establishing a robust basis for inferring historical processes. Frequently, the strength of inference from molecular results has suggested new questions for biogeographers and their allies in related disciplines.

I Vicariance and dispersal in the biogeographical history of Madagascar

To have an interest in the historical biogeography of Madagascar necessitates a thorough understanding of the origins and gradual sundering of Gondwana. When the supercontinent Pangea began to divide into the southern continental landmasses (Gondwana) and northern continental landmasses (Laurasia), approximately 175 Ma, Madagascar was tucked away, deep within Gondwana (Rakotosolofo et al., 1999; Reeves et al., 2002). Shortly thereafter, Madagascar began the long journey to its current state of remote isolation in the Indian Ocean. Presently, Madagascar is surrounded by a vast oceanic barrier on all sides. It is closest to continental Africa, approximately 400 km to the west, but lies 4000 km from India, 5000 km from Antarctica, and 6000 km from Australia. It is therefore remarkable to consider that each of these landmasses was at one time contiguous with Madagascar.

Initially, Gondwana was a single contiguous supercontinent comprised of what would become Africa, South America, Antarctica, Australia, India, and Madagascar. Shortly after separating from Laurasia, rifting between western Gondwana (Africa plus South America) and eastern Gondwana (Madagascar plus India, Antarctica, and Australia) commenced (Briggs, 2003) as evidenced by the vast outpourings of the Karoo volcanics (182±1 Ma; de Wit, 2003). From that point, Madagascar and the rest of eastern Gondwana began to drift southward relative to Africa, sliding along the strike-slip fault known as the Davie Ridge (Bassias,
There is a general consensus among geologists that this occurred sometime between 165 and 155 Ma (Rabinowitz et al., 1983; Agrawal et al., 1992; Reeves and de Wit, 2000; Scotese, 2000; Reeves et al., 2002; Briggs, 2003). By 140 Ma (Seward et al., 2004), marine conditions are clearly prevalent along the entirety of Madagascar’s west coast. Thus, although separation from Africa began as early as 165 Ma, there was a subsequent period of perhaps 20 million years wherein biotic exchange would have been likely between western and eastern Gondwana. Madagascar reached its current position with respect to Africa by 130–118 Ma (Rabinowitz et al., 1983; Harland et al., 1990; Seward et al., 2004). Although shifts in latitude and relationships to other landmasses remained dynamic for many millions of years subsequent to this positioning, it is certain that Madagascar’s present geographical isolation relative to Africa has been stable for at least 118 million years, and probably considerably longer. Moreover, the Mozambique Channel, approximately 400 km wide and separating Madagascar and Africa, is quite deep (3000 m, at the deepest point) and would not have been notably affected by changing sea levels after the final separation of Madagascar and Africa (Krause et al., 1997).

The dynamic sundering of Gondwana over the past 165 million years or so has been marked by temporal windows wherein biotic exchange would have been facilitated by contact with other landmasses, followed by a long period of progressive geographical isolation. The timing and sequence of fragmentation and isolation may at times have been slowly progressive (eg, Madagascar’s longslide south, relative to Africa), or relatively abrupt (eg, the hypothesized severing of southern connections via Antarctica). The most recent conceivable connection to other landmasses via the Kerguelen Plateau at 80 Ma was approximately contemporaneous with the temporal origins of major extant clades (eg, placental mammals), but considerably more ancient than many of the extant subclades in Madagascar (eg, feliform carnivorans). Thus, the present-day biota of Madagascar must of necessity be composed of groups whose presence is best explained by a mix of ancient vicariant events and transoceanic dispersal. The challenge for biogeographers is to distinguish between these two possibilities.

A recent analysis of dated phylogenies across multiple lineages (plants, vertebrates, invertebrates; Yoder and Nowak, 2006) points to two generalities: (1) that there are numerous endemic clades of Malagasy taxa whose closest sister group relationships are to African taxa; and (2) that there appears to be an overwhelming signal of Cenozoic dispersal. The consideration of outgroup nodes suggests that the majority of lineages ancestral to Malagasy endemics had their origins in Africa. For those groups in which divergence ages have been estimated, the vast majority of both crown and stem ages fall within the Cenozoic. The inescapable inference, although the number of studies relative to the number of endemic taxa is paltry, appears to be that the living biota is predominantly comprised of ‘neo-endemics’ that have evolved from transoceanic dispersers.

This strains credulity for some (eg, Stankiewicz et al., 2006), especially given our knowledge of Madagascar’s extreme and long-standing isolation. However, it is not a novel observation to view the living Malagasy biota as ‘immigrants’ (Simpson, 1952; Krause et al., 1997; 2003). Simpson (1952) disputed hypotheses of landbridge connections between Madagascar and other landmasses, pointing out that the very limited representation of mammalian taxa in Madagascar is definitive indication that the probability of colonization must have been ‘exceedingly low’. Platnick (1981) even considered transoceanic dispersal hypotheses to border on the ‘miraculous’. Yet landbridge hypotheses have continued to be invoked (eg, McCall, 1997), despite no empirical support (eg, Yoder et al., 2003; Poux et al., 2005) because there is
compelling evidence that there may once have been two small subaerial exposures in the Mozambique Channel for a limited geological period during the Miocene that could have broken the trip from Africa into three north to south stages of 295, 210, and 125 km, and in the direction of the prevailing currents (Malod et al., 1991; Bassias, 1992; Krause et al., 1997).

A more useful model of dispersal may be drawn from present-day phenomena. Krause et al. (1997) review contemporary reports of floating 'islands' of vegetation, often with standing trees and mammalian inhabitants, observed in remote oceanic locations, tens and hundreds of kilometres from land (Millot, 1953; King, 1962; Carlquist, 1965). Certainly, there have been numerous empirical reports of transoceanic dispersal in lizards during recent history (Censky et al., 1998; Carranza et al., 2000; Calsbeek and Smith, 2003). Add to this evidence the likelihood that the ancestors of today's Malagasy mammals either were hibernators or had other physiological capacities for reduced metabolic demands (Kappeler, 2000), and the 'miraculous' becomes slightly more routine. Credulity is even less strained by hypotheses of long-distance dispersal in plants and invertebrates given their less stringent life-history attributes (Monaghan et al., 2005; Thiel and Gutow, 2005; Barnes et al., 2006; Cowie and Holland, 2006). Given the data on hand, and the remarkable consistency of the signal across evolutionary lineages, we must consider the following to be the most credible hypothesis for the time being: Madagascar is an island primarily composed of 'neo-endemics' that are the descendants of Cenozoic waif dispersers.

2 Neogene origins of deserts and desert biotas in western North America
The case study of Malagasy biota (above) indicates how a firm foundation of geological studies permitted explicit testing of dispersal versus vicariance hypotheses to account for the origin of that fauna. Current studies of four endemic lineages of Malagasy mammals (lemurs, carnivores, tenrecs, and nesomyine rodents) are now focused on reconstructing their in situ diversification. Intensive field survey and systematic revision have already revealed an enormous amount of cryptic variation: of the 56 currently recognized species of tenrecs and nesomyines, a quarter of these have been described since 1992. Consequently, phylogeographical studies of endemic Malagasy lineages show great promise in revealing geographical patterns important for conservation planning, which is now a process proceeding independently using habitat and species distribution data (Kremen et al., 2008).

Initially, the foundation of geological knowledge as the basis for phylogeographical studies of the North American aridlands appeared to be equally firm. Four major regional deserts have been recognized since their initial description by Shreve (1942): a colder Great Basin, transitional Mojave, and two warmer southern deserts, Sonoran and Chihuahuan. The distribution and diversification of these regional deserts resulted from a complicated and dynamic geological and climatological history during the Cenozoic (first appreciated and described by Axelrod, 1958; 1983). Their origin followed a major period of mountain-building during the early Eocene, including uplift of the Rocky Mountains and Sierra Madre Oriental and volcanic origin of the Sierra Madre Occidental (Ortega-Gutiérrez and Guerrero-Garcia, 1982; Swanson and McDowell, 1984; Ferrusquía-Villafranca et al., 2005). Putative causal events invoked in the subsequent diversification of aridlands taxa include major climatic shifts, mountain-building, and plate movement. Global cooling of the early Oligocene and middle Miocene marked an abrupt and extreme shift from mesic to drought-tolerant plant species and began the cooling and drying trend that has led to the current glacial episodes (Prothero, 1998; Woodburne, 2004; Hafner et al. 2007; Hafner and Riddle, 2008). Following collision of North America with the East Pacific Rise
in the Miocene, the relatively flat landscape of western North America was reshaped with the block-fault origin of parallel mountain ranges that created massive rainshadows, intensified plate movement that opened the Gulf of California, and volcanic eruptions along the plate boundary that created the Baja California Peninsula (Norris and Webb, 1976; Coney, 1983; Gastil et al., 1983; Lonsdale, 1989; Spencer and Normark, 1989; Stock and Hodges, 1989). The Cape Region at the southern tip of the peninsula was originally part of the continental mainland that was transferred to the Pacific plate about 10 Ma, and was later joined to the California mainland as volcanic eruptions and uplift coalesced to form the peninsula. Sea levels fluctuated and ecological zones shifted (in latitude and elevation) in response to repeated waxing and waning of the late Pleistocene glacial-interglacial cycles, which became markedly longer and more extreme about 700,000 Ka (Webb and Bartlein, 1992).

Comparative phylogeographical studies of the southern regional deserts were based initially on a cadre of arid-adapted rodents (Riddle et al., 2000a; 2000b) and later were expanded to include 22 clades of mammals, birds, reptiles, amphibians, and plants (Riddle et al., 2000c; Riddle and Hafner, 2006b). In contrast to previous biogeographical models that emphasized late Pleistocene events, these molecular phylogeographical studies recovered relatively deep divergence between the Sonoran and Chihuahuan regional deserts (putatively correlated with the uplift of the Sierra Madre and Mexican Plateau) as well as more recent dispersal and vicariance putatively correlated with late Pleistocene glacial cycles. Phylogeographical patterns generally supported the distinction of Sonoran and Chihuahuan regions, revealing idiosyncratic species sensitivity to different filter-barriers and supporting an inclusion of a unique desert plant community (Sinaloan thornscrub) in a southern extension of the Sonoran desert, in agreement with studies based on panbiogeographical methods (Arriaga et al., 1997; Morrone, 2001a; 2001b; 2004; 2005). In even greater contrast to the prevailing view of the Baja California Peninsula as a somewhat depauperate extension of the Sonoran desert, these studies revealed a deep-history divergence of the peninsular and Sonoran-mainland biota, prompting consideration of earlier vicariance, rather than more recent dispersal, as the source of much of the peninsular flora and fauna. Subsequent molecular phylogeographical studies of additional taxa (eg, Sinclair et al., 2004; Lindell et al., 2005; Patton and Álvarez-Castañeda, 2005; Crews and Hedin, 2006; Devitt, 2006; Lindell et al., 2006; Wood et al., 2008) have underscored this deep historical separation, as well as indicating a far more complicated history of vicariance and dispersal on the peninsula itself. Consequently, there is increasing support for recognition of the peninsular desert as the most distinct North American regional desert, as initially proposed by Hafner and Riddle (1997) and supported by panbiogeographical studies of Morrone and his colleagues (see above citations).

At this point, biotic patterns resulting from comparative molecular phylogeographical studies of Baja California species have advanced far ahead of geological evidence: instead of testing phylogeographical hypotheses based on firm geological evidence, geological hypotheses based on phylogeographical patterns await further geological field study (Hafner and Riddle, 2008). It is unknown when terrestrial contact was severed between the peninsula and either the eastern mainland (across the central Gulf of California) or the northern mainland (at the head of the Gulf). While it is generally accepted that a proto-Gulf of California opened as early as 12 Ma, Ledesma-Vázquez (2002) maintains that a mid-peninsular land connection persisted until the proto-Gulf and current Gulf opening were connected as recently as 3 Ma; whereas Murphy and Aguirre-León (2002) indicate complete
separation from the mainland much earlier in the process. Two temporary aquatic barriers (marine inundations and perhaps a chain of freshwater to estuarine lakes) appear to have limited or completely severed terrestrial communication between the peninsula and the mainland at its northern end (Lucchitta, 1979; Boehm 1984; Buizing, 1990; Faulds et al., 2001; Spencer and Perthree, 2001; Murphy and Aguirre-León, 2002; Spencer and Perthree, 2005), and remnants of each barrier or filter-barrier are being revealed in a growing number of terrestrial taxa (Riddle and Hafner, 2006b).

Repeated phylogeographical patterns among diverse vertebrate (terrestrial and marine), invertebrate, and plant taxa indicate past fragmentation of the peninsula itself. As the number and diversity of co-distributed lineages has grown, with each showing a molecular divergence in the vicinity of the central-peninsula Vizcaíno Desert, argument has focused on whether this divergence is due to abrupt ecological and climatic change (eg, Grismer, 2002) or on a past mid-peninsular seaway. Upton and Murphy (1997) initially hypothesized a temporary mid-peninsular seaway at about 1 Ma as the causal vicariant event for a lizard lineage, and subsequent studies of rodent lineages appeared to support this recent date (Riddle et al., 2000a; 2000b; 2000c). Riginos (2005) presented evidence from near-shore reef fishes that strongly supported the past existence of a seaway, but argued that the extensive molecular sequence divergence observed among a variety of terrestrial and marine lineages in the area support a much earlier date (eg, reptiles, Rodriguez-Robles and De Jesus-Escobar, 2000; Lindell et al., 2005; spiders, Crews and Hedin, 2006; near-shore reef fishes, Riginos, 2005). Geological and palaeontological evidence indicate a late Miocene (7 Ma) or early Pliocene submergence of the central peninsula, associated with the opening of the initial proto-Gulf (Smith, 1984; 1991; Ortlieb, 1991; Helenes and Carreño, 1999; Holt et al., 2000; Carreño and Helenes, 2002; Ledesma-Vázquez, 2002). Murphy and Aguirre-León (2002) cite geological evidence (Ochoa-Landin, 1998) that dates the seaway at 3 Ma. Circumstantial evidence in the form of topography, elevation, limited stratigraphic data, and recent volcanic eruptions suggest that limestone in the region may have been deposited by shallow, warm seas sometime during the last million years (Minch and Leslie, 1991). A series of low-elevation salt flats connect the two marine lagoons that penetrate the flat Vizcaíno lowlands, and intense eruptions of the Tres Virgenes volcanic complex that began 1.2 Ma may have triggered uplift and expulsion of marine waters (Hafner and Riddle, 2008).

Lindell et al. (2006: 1329) argued that a single ‘vicariant event affecting all species similarly’ (ie, a single temporary seaway) is the most parsimonious explanation for the large number of co-distributed lineages (nearly 20 terrestrial and 10 marine) in the mid-peninsular region. The wide range of molecular sequence divergence (from extensive to subtle) that has been observed among a variety of terrestrial and marine lineages in the area might argue instead for multiple seaways of disparate age in the central peninsula, as suggested independently by Lindell et al. (2005), Crews and Hedin (2006), Leaché et al. (2007; based on hierarchical Bayesian analysis) and Hafner and Riddle (2008). Until and unless detailed geological or palaeontological evidence is found of multiple seaways in the central peninsula, the hypothesis of abrupt ecological change in concert with Pleistocene climatic shifts cannot be rejected as the source of vicariance of terrestrial lineages, particularly of taxa separated by more subtle molecular divergence.

Similarly, elements of the Cape Region flora and fauna may include relicts of its Miocene-age Mexican mainland connection (eg, a number of reptilian lineages listed by Murphy and Aguirre-León, 2002), may retain records of temporary Pliocene-age seaways in that region (McCloy, 1984; Riddle et al., 2000c), or may represent recent immigrants.
via Pleistocene glacial-maximum dispersal corridors. A broader reconnaissance of possible Pliocene and Pleistocene marine deposits coupled with a more exact understanding of the history of uplift, subsidence, and fault dynamics in this region will be required to better evaluate patterns of genetic divergence observed in Cape Region lineages.

Both the Malagasy and North American aridlands case studies have demonstrated that molecular phylogeographical studies based on hypotheses derived from geological evidence can eventually lead heuristically to hypotheses that await testing with new geological, palaeontological, or climatological evidence. For Madagascar, diversification of endemic mammalian lineages almost certainly predated the late Quaternary, for which the earliest palaeoclimatic data are available. For the North American deserts, phylogeographical studies have indicated specific areas on the Baja California Peninsula and the surrounding mainland that are deserving of heightened geological and palaeontological survey.

3 Vicariance, dispersal, and connectivity in the seas

Classical examples of marine biogeography include regional faunas separated vicariantly east versus west of the Isthmus of Panama, across the East Pacific Barrier, and among the many basins of the Indo-West Pacific (eg, Briggs, 1974; 1995; Santini and Winterbottom, 2002) among others (eg, Dartnell, 1974; Wilson and Allen, 1987). Concomitantly, the classic phylogeographical studies described reciprocally monophyletic, geographically discrete, mitochondrial clades in co-distributed maritime taxa of southeastern USA (Avise 1992), a pattern that also characterizes intraspecific variation in a number of other regional biotas, including southwestern North America (Burton, 1998; Dawson et al., 2006), southeastern Australia (eg, Waters and Roy, 2003; Waters et al., 2005), and the aforementioned exemplars of vicariance (eg, Benzie, 1998; Lessios et al., 2001; Williams and Reid, 2004; Barber et al., 2006). Thus, the gradualist allopatric evolutionary mechanisms predominant in evolutionary theory for over 50 years (Mayr, 1942; 1954; Palumbi and Lessios, 2005) and explicit in the phylogeographical hypotheses (Avise et al., 1987) have garnered considerable support even in an environment as fluid as the oceans.

In each classical example of process, however, there also is an undercurrent of dispersal. For instance, divergence times of geminate species pairs across the Isthmus of Panama vary depending on, among other things, species’ proclivities to disperse across shallow water (Knowlton and Weigt, 1998; see also Marko, 2002). Dozens of taxa have dispersed across the 5000 km wide East Pacific Barrier (Leis, 1984; Scheltema, 1988; Briggs, 1995; Lessios and Robertson, 2006), the Indo-West Pacific is traversed by numerous species (Williams and Reid, 2004), and the Floridian fauna includes taxa with little or no population subdivision (eg, Avise et al., 1987; Avise, 1992). Such remarkable feats of dispersal have come to epitomize marine biogeography, particularly as a counterpoint to terrestrial systems, for example in the long-standing focus on relationships between pelagic larval duration and gene flow (eg, Crisp, 1978; Doherty et al., 1995; Bohonak, 1999; Lester et al., 2007).

That marine taxa, like terrestrial and freshwater taxa, should show varying degrees of geographical isolation and of dispersal is not unexpected considering the complexity of biological and environmental systems (Rosenblatt, 1963; Dawson and Hamner, 2008). Efforts to understand how physical environmental and organismal traits interact to determine community and population genetic structure have received renewed attention in studies of connectivity (eg, Kinlan and Gaines, 2003; O’Connor et al., 2007; Ramon et al., 2008).

Studies of connectivity originated in the early to mid-1990s, when the term was first used in organizational and prepublication
settings, but rarely in journals, in contexts such as fluxes across the landscape or between ecosystems, or dispersal and gene flow among marine populations (http://aims.gov.au/pages/research/trp/pages/trp2-43.html; J.A.H. Benzie, R.K. Cowen, personal communication; see also DeFreese, 1995; Doherty et al., 1995; Ray, 1996). The term appeared in mainstream marine science at the turn of the century, when it was used to describe dispersal between geographically distinct populations in terms of larval advection, diffusion, and mortality (Cowen et al., 2000). Connectivity is now the key concept in coastal resources management (Mora and Sale, 2002; Steneck, 2006) and an umbrella term for processes ‘pivotal to our understanding of the population dynamics, genetic structure, and biogeography of many coastal species’ (Cowen et al., 2006: 522).

It has attained this degree of success in part by encapsulating in a single word the pluralistic framework in which molecular genetics, satellite oceanography, individual dispersal, and selection may be integrated empirically with mathematical oceanography and computational simulation (eg, Dawson et al., 2005; Cowen et al., 2006; Galindo et al., 2006; Follows et al., 2007), largely separate from long-standing, sometimes polarizing, debates in biogeographical theory. It represents a milieu in which studies at the extremes of biogeography, from global microbial biogeography to island evolution, can co-exist and, moreover, be synthesized. We specifically mention these two areas of research because, in addition to representing the apparent extremes of biogeographical patterns, microbial biogeography and island evolution also mirror historical debates in marine biogeography – the perceived pre-eminence of dispersal (see above) and consequent absence of closed marine systems (Hatcher, 1997) – and will play a large role in advancing our understanding of marine biogeography.

The study of marine microbial phylogeography because the underexplored ‘rare biosphere’ (Sogin et al., 2006) must harbour many microbes difficult to culture ex situ (Kaeberlein et al., 2002), and in situ methods bias the community sampled (Yasumoto-Hirose et al., 2006). Thus, only environmental genomic methods, typically DNA fingerprinting such as automated rRNA intergenic spacer analysis (ARISA) and terminal restriction fragment length polymorphism (TRFLP), can adequately attempt to describe natural microbial diversity. The results are intriguing, ranging from extreme local heterogeneity, through ‘moderate endemism’ or regionalization within broad climatic bands or ecologically relevant depth strata, to global dispersal (Baldwin et al., 2005; Bell et al., 2005; Noguez et al., 2005; Pommier et al., 2005; van der Gast et al., 2005; Bass et al., 2007; Darling et al., 2007; Medlin, 2007; Yutin et al., 2007; Zwirglmaier et al., 2008).

The currently rudimentary knowledge of how microbial ribotype diversity is influenced via phenotype by the environment, however, precludes an important perspective on the major thesis in microbial biogeography: that ‘everything is everywhere, the environment selects’ (eg, see Green and Bohannan, 2006). Experimental environmental manipulations, inclusion of functional genotypes in molecular analyses, and selective in situ sampling of community composition (eg, Battin et al., 2003; Hewson and Fuhrman, 2006; Yasumoto-Hirose et al., 2006) are addressing this issue and undoubtedly will in time explain why, seemingly against the odds, the distributions of some microbes with potentially massive dispersal ability appear to show island biogeographical patterns on local scales (Bell et al., 2005; Noguez et al., 2005; van der Gast et al., 2005).

High β-diversity (a measure of taxonomic turnover) across short spatial scales in marine taxa with potential for long-distance dispersal has typically been explained as the result of environmental heterogeneity, or ‘patchiness’ (eg, Smith and Witman, 1999;
Although a flourish of studies in the 1970s found some evidence in marine systems of island biogeographical patterns (e.g., Schoener, 1974a; 1974b; Dauer and Simon, 1976; Molles, 1978) the implications, such as limits on migration, were not widely adopted – note the absence of island theory from texts by Nybakken (1997) and Castro and Huber (2003), in contrast to its representation in Ricklefs and Miller (1999), Krebs (2001), Krohne (2001), and Whittaker and Fernández-Palacios (2007). In part, this was due to experimental design confounded by scale (Schoener, 1974a; 1974b), a growing emphasis on dispersal (Crisp, 1978), and evidence that endemism was lower in marine taxa than in terrestrial taxa on the same oceanic islands (e.g., Sterrer, 1998; Randall, 1998; Wagner and Funk, 1995; Drew and Roderick, 2005; see also Paulay and Meyer, 2002). However, the commonplace application of molecular tools to marine taxa has demonstrated that many supposed widespread ecological generalists are complexes of cryptic species (e.g., Knowlton, 1993; 2000), often occurring as archipelagic endemics (e.g., Paulay and Meyer, 2002; Williams and Reid, 2004), or in marine lake ‘islands’ (Dawson and Hamner, 2005). It is not surprising, therefore (when one looks) to also find evidence of other aspects of island evolution in marine biotas (e.g., McClain et al., 2006; see also Knowlton, 2001; Robertson, 2001).

Connectivity is a new term, but its gestalt bears the synthetic hallmarks of evolutionary and biogeographical classics (e.g., Mayr, 1942; 1963; Carlquist, 1965), island biogeography theory (MacArthur and Wilson, 1967), phylogeography (Avise, 2000), and population genetics (e.g., Conner and Hartl, 2004). Its purview includes synthesis of colonization-extirpation dynamics; mutation, migration, and selection; and micro- and macro-evolution. Its novelty, particularly in terms of associated technological advances – including molecular genetics – and a readjustment of research focus, also offers renewed potential for long-overdue integration of the biogeography of marine and terrestrial systems suggested a half-century ago (Mayr, 1954; Dawson and Hamner, 2008), and the more recently compared micro- and macrobiota (e.g., Horner-Devine et al., 2007).

**V Predicting the future of biodiversity using genetic approaches and data**

Use of molecular genetics will figure prominently in discourse about the effect of near-term environmental change on local, regional, continental, and global biodiversity. Undeniably, extinction will play an important role in changing the structure of the future biota of the earth. Thus, to best preserve Earth’s biota, we must anticipate and construct an early warning system that identifies species vulnerable to extinction. Further, this warning system must be made more robust by knowledge of how species handle environmental change in natural settings over centuries to millennia to millions of years – knowledge that does not presently exist. While we would expect a number of species to go extinct under the combined influences of global climate change, habitat modification and destruction, and invasive species, we need to understand further the ways in which species have, alternatively, shifted ranges – possibly in combination with rapid morphological evolution (Hellberg et al., 2001) – or shifted phenotypes in situ (Smith et al., 1995; Chiba, 1998; Hadly et al., 1998) under previous episodes of climate change. Integrative biogeography can provide us with a global perspective on the unique set of events that structure species, communities, and biomes yielding insights relevant to conservation not otherwise possible (Barnosky et al., 2001). This new biogeography, aided by advancements in population genetics, genomics, and phylogeography, as well as more sophisticated Earth history reconstructions, provides a powerful toolkit because early warning signs such as population subdivision, local extinction (extirpation), and loss of...
connectivity leave indelible genetic marks, often over the variable timescales most relevant for future global change.

Not only do genomes of species harbour evidence of events past (excluding direct measures of extinction events, which remains the purview of palaeontology), but genomes are species’ futures as well (Lacy, 1987). All the tools available to populations, species, and communities are stored in their genetic codes. In fact, the evolutionary potential of species and populations has been questioned in light of human-accelerated climate change (Myers and Knoll, 2001; Thomas et al., 2006) and habitat degradation (Novacek and Cleland, 2001). The genetic diversity of any given population through time is a result of the addition of novelty through recombination, mutation or gene flow, and the relative change in gene frequency (including loss of unique alleles) through selection or drift. Because mutation is a relatively slow process, its contribution to the addition of genetic novelty and shaping genetic diversity is likely to be less significant over the temporal scales of current environmental change (100 to 1000 years), predicted by the IPCC (2007) to increase in the future, than are genetic drift, gene flow, and selection. Thus the population processes most important for shaping genetic diversity are genetic drift, gene flow, and, potentially, selection (Hoffman and Blows, 1993). Over time, the degree of isolation of a population will determine whether genetic diversity is primarily the product of within-population processes or the result of interactions between populations. Ongoing and future climatic change and habitat fragmentation is already decreasing the potential for interactions between populations, resulting in indirect effects on the genetic diversity of species (Bradshaw and Holzapfel, 2008). Relevant empirical data are scarce, however, leaving uncertain the effects that physical perturbations may have on genetic diversity and connectivity of populations. Human activities also are likely to further influence population sizes and, hence, genetic diversity of many species (eg, Kohlmann et al., 1988; Debinski and Holt, 2000; Root et al., 2003). Thus, understanding the contribution of past environmental change, vicariance and dispersal on the connectivity and genetic diversity of populations and species is paramount to making predictions about the future of species persistence, and ultimately their evolution (Waits et al., 1998; Allendorf and Luikart, 2006). Genetic diversity and its distribution, in fact, may be a valuable harbinger of things to come.

In March, 2007, a group of 25 biogeographers, ecologists, and palaeontologists convened a workshop to consider the future of a new Integrative Biogeography. One of the themes that emerged from that exercise was the need to develop a ‘Map of Life’ project (Kidd, personal communication) – that will be equivalent and complementary to the ongoing Tree of Life (aTOL: http://atol.sdsc.edu/) and Barcode of Life (CBOL: http://barcoding.si.edu/) projects. While such an endeavour will require synthesis across a broad range of disciplines and approaches, the rapid growth in our understanding of the diversification and distributional histories of populations, species, higher taxa, and biotas is directly related to the acceleration in approaches to estimating molecular genetic similarities and differences across taxa. The need for such an initiative is both timely, due to its direct connections to the aTOL and CBOL projects, and critical, in that the diversity that we wish to map is rapidly disappearing due to human agency.

Acknowledgements

We thank the National Science Foundation for their support of a March 2007 workshop (DEB-0639022) in which several of the themes reviewed in this paper emerged.

References


Arriaga, L., Aguilara, C., Espinosa-Organista, D. and Jiménez, R. 1997: Regionalización ecológica y biogeográfica de México. Taller de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (Conabio), México, D.F.


Drummond, A.J., Rambaut, A., Shapiro, B. and Pybus, O.G. 2005: Bayesian coalescent inference of
past population dynamics from molecular sequences. Molecular Biology and Evolution 22, 1185–92.


Hallström, B. M., Kullberg, M., Nilsson, M. A. and Janke, A. 2007: Phylogenomic data analyses provide evidence that Xenarthra and Afrotheria are sister groups. Molecular Biology and Evolution 24, 2059–68.


Lindell, J., Méndez-de la Curz, F.R. and Murphy, R.W. 2005: Deep genealogical history without
population differentiation: discordance between mtDNA and allozyme divergence in the zebra-tailed lizard (*Callisaurus draconoides*). *Molecular Phylogenetics and Evolution* 36, 682–94.


Riddle, B.R. 1995: Molecular biogeography in the pocket mice (*Perognathus* and *Chaetodipus*) and grasshopper mice (*Onychomys*) – the Late Cenozoic development of a North American aridlands rodent guild. *Journal of Mammalogy* 76, 283–301.


Rigos, C. 2005: Cryptic vicariance in Gulf of California fishes parallels vicariance patterns found in Baja California mammals and reptiles. *Evolution* 59, 2678–90.


Scotese, C.R. 2000: *Earth History (paleogeographic maps).* PALEOMAP Project. Department of Geology, University of Texas.


Topp, C.M. and Winker, K. 2008: Genetic patterns of differentiation among five landbird species on the Queen Charlotte Islands, British Columbia. Auk, in press.


