Statistical Phylogeography

L. Lacey Knowles

Department of Ecology and Evolutionary Biology, Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109-1079; email: knowlesl@umich.edu

Annu. Rev. Ecol. Evol. Syst. 2009. 40:593–612

The *Annual Review of Ecology, Evolution, and Systematics* is online at ecolsys.annualreviews.org

This article's doi:
10.1146/annurev.ecolsys.38.091206.095702

1543-592X/09/1201-0593$20.00

Key Words

biogeography, coalescence, demography, genetic parameters, species divergence

Abstract

Phylogeography's objective—to understand the processes underlying the spatial and temporal dimensions of genetic variation—underlies both the prominence and extensive methodological transformations that characterize this nascent field. Here I discuss the insights that come from detailed demographic information and how an understanding of phylogeographic history is crucial to addressing a range of evolutionary and ecological questions, from understanding the source of adaptive divergence to the factors structuring ecological communities. I review recent progress in phylogeography, including its expanding role in evolutionary and ecological study and the molecular and methodological advances that now provide unprecedented details about the factors governing population genetic variation and structure. As a field, phylogeography draws together information across disciplines (e.g., from genetics, ecology, systematics, and paleontology), using a diversity of technical and conceptual approaches. This unified eclectic perspective has been key to phylogeography's success and will be key to phylogeography's enduring future.

INTRODUCTION

Statistical phylogeography: demographic inference that takes into account the stochasticity of genetic processes

Phylogeography endeavors to understand the processes that underlie the distribution of genetic variation within and among closely related species. Although the means by which this goal might be achieved differs considerably from those that spawned the field of phylogeography some thirty years ago, the foundation and conceptual breakthrough made by J.C. Avise are nonetheless the same and are as relevant today as they were two decades ago. Namely, patterns of neutral genetic variation among individuals carry the signature of a species' demographic past (Avise et al. 1987).

This seemingly simple phrase belies the impact of this perspective (see Avise 2000). Phylogeographic study provides detailed species-specific information on how geologic events, environmental influences, and geographic factors interact with aspects of a species ecology and natural history in shaping its evolution. In a comparative context, such causal links can reveal ways in which entire communities or assemblages are structured by shared responses to a common past and unique species-specific historical events. When integrated with other types of information (e.g., phenotypic data), the detailed demographic information extracted from genetic data provides a critical backdrop for an array of evolutionary and ecological questions that could not be addressed without an understanding of phylogeographic history. Moreover, as a field that bridges disciplines, developments in phylogeography have enriched areas tangential to phylogeographic pursuits (e.g., application of coalescent-based approaches in phylogenetics, tree-based thinking to study disease dynamics, and tests of the genetic basis for adaptive evolution). Here I review the advances in phylogeography that have propelled the field forward and examine how its integrative nature has lead to breakthroughs and serves as a primary source of innovation. Reference to some benchmark studies highlights phylogeography's ever increasingly important and broadening role in evolutionary and ecological study.

Some of the examples (and methods) highlighted in this review also show some commonalities with the growing field of landscape genetics. Although the focus of landscape genetics is exclusively on the effects of the contemporary landscape on patterns of genetic variation (see review by Storfer et al. 2007), phylogeographic studies often lie at an intersection where factors impacting both historical and contemporary species distributions are critical to understanding the processes structuring patterns of genetic variation. Consequently, there may be some overlap in methodology between the fields depending on the specifics of the taxa and system, which define the appropriate spatio-temporal scale of study. As a review of statistical phylogeography (that is, studies focused on the processes generating patterns of genetic variation within and among closely related species) the extensive literature on phylogeographic patterns (often detailed in analyses of mtDNA gene trees) is not covered here (see Avise 2000 for a thorough review).

PHYLOGEOGRAPHIC BREAKTHROUGHS

The breakthroughs associated with today's phylogeographic studies extend far beyond the expected increased statistical rigor that comes with greater computational sophistication and molecular advances. The field has witnessed significant conceptual gains, whether the central focus is on the inferred biogeographic and demographic processes or on testing evolutionary and ecological questions that require a historical framework.

The Appeal, Power, and Challenge of Recent Advances

Although rising to prominence on the use of gene trees and mitochondrial DNA (mtDNA) (Avise 2009), the field has moved far beyond this singular approach and the original concept of bridging the traditionally separate disciplines of population genetics and phylogenetics. Today's phylogeographic studies employ a staggering array of different methods and molecular tools (e.g., Cornuet et al. 2008, Dolman & Moritz 2006, Estoup et al. 2004, Excoffier & Heckel 2006, Fagundes et al. 2007, Harrigan et al. 2008, Hey & Nielsen 2004, King & Roalson 2009, Knowles 2001, Lemmon & Lemmon 2008) and celebrate the input from a broad cross section of disciplines, from community ecology, climatology, and geology to molecular and evolutionary genomics (e.g., Bigg et al. 2008, Bonin et al. 2006, Carnaval et al. 2009, Carstens & Richards 2007, Harter et al. 2004, Jha & Dick 2008, Kidd & Ritchie 2006, McCormack et al. 2008, McRae 2006, Rambaut et al. 2008, Waltari et al. 2007).

This diversity has at times created an aura of an uncertain future, as phylogeography has sometimes struggled to overcome the conflicts that emerged from what was initially its most notable contribution—the use of gene trees for demographic inference (Avise et al. 1987). When the rate of coalescence differs among gene lineages, individual gene trees contain information about past demographic events (e.g., a distinct genetic signature of population subdivision is apparent in a gene tree because the coalescence times of gene lineages between populations extend all the way back to when the lineages were present in a common ancestral population, as opposed to when the lineages follow the expected distribution of coalescence times within populations) (see Wakeley 2003, 2007). However, there is an inherent contradiction in using gene trees as a literal interpretation of a species past—gene trees contain historical signal, but also stochasticity (that is, there is a random loss of gene lineages by genetic drift because some individuals leave many offspring and others few). This randomness represents a problem only when this duality is overlooked. Indeed, recent advances (for the most part) eclipse any regretful period marked by overzealous interpretations from a particular method (but see Knowles 2008); phylogeography has undergone both marked methodological and philosophical transformations in how genetic data are used for demographic inference (see **Figure 1**) (reviewed in Arbogast et al. 2002, Hickerson et al. 2009, Knowles 2004, Hey & Machado 2003, Nielsen & Beaumont 2009, Wakeley 2004).

The unwavering popularity of phylogeography is indisputable (**Figure 2**). Much of this core appeal arises from the diverse perspectives the study of phylogeography welcomes, and actively

Coalescence: the process in which genes from the present merge into their common ancestor going back in time

Expected distribution: distribution of values (e.g., Fst or coalescence times) under a specific model that can be compared with observations computed from empirical data

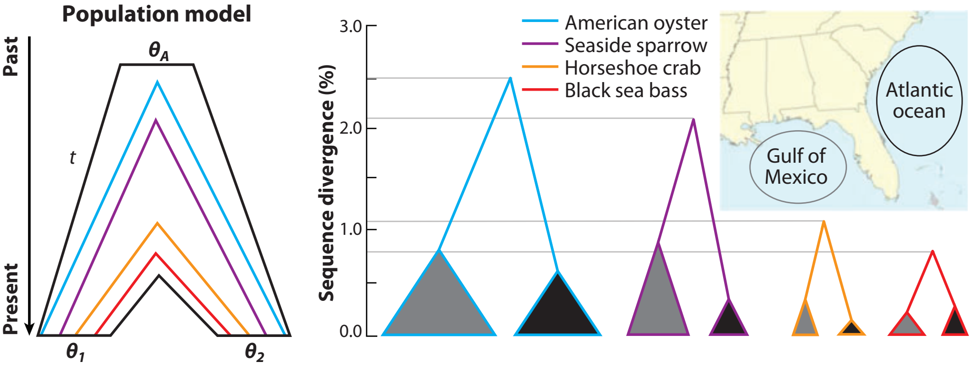

Figure 1

Without a population model, it is not possible to tell whether the differences in gene lineage divergence times between the populations distributed along the Atlantic and Gulf Coast among the different species are (or are not) consistent with a common response to a shared historical factor (such as changes in sea level associated with Pleistocene climate change; Avise 1992). The lack of temporal congruence of gene lineage divergence across species might be within the expected variance given the respective population sizes of the taxa; a model-based perspective is essential to evaluating this hypothesis (for a review on comparative phylogeographic analysis, see Hickerson et al. 2009).

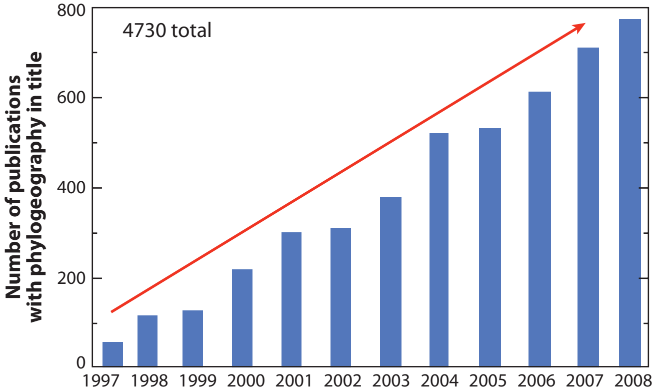

Figure 2

Impressive trajectory showing the increase in the number of phylogeographic studies over the past twenty years; note that the popularity of phylogeography is underestimated considering that only publications in which phylogeography appeared in the title were tallied.

fosters, for understanding the processes underlying patterns of genetic variation within and between closely related species. In the past, the type of statistical analyses a researcher might have used reflected divisions between the fields of population genetics and phylogeography (Hey & Machado 2003, Wakeley 2003), such as deciding to use a gene tree as the basis for inference versus deriving parameter estimates from a huge collection of genealogies (that is, using genealogy samplers; reviewed in Kuhner 2008). However, this traditional divide has now been bridged by statistical phylogeographic approaches (Knowles & Maddison 2002) that consider both coalescent and mutational variance, leaving a variety of approaches that could be used to estimate population genetic parameters and test hypotheses. These include methods that calculate the probability of the data using likelihood-based or Bayesian approaches. These programs (reviewed in Excoffier & Heckel 2006, Kuhner 2008), though computationally intensive, provide detailed estimates of the past, such as population size, growth, subdivision, admixture, patterns of gene flow, and the timing of divergence. These methods have become increasingly popular because of their potential to disentangle complex histories (e.g., distinguishing divergence with gene flow from retention of ancestral polymorphism) (e.g., Carstens & Knowles 2007, Harter et al. 2004, Hey et al. 2004), whereas multiple evolutionary forces have been difficult to distinguish in the past. Alternatively, methods that rely on summary statistics, as opposed to considering the full probability of the data, could be used (reviewed by Hickerson et al. 2009, Knowles 2004). Such approaches offer extreme flexibility, accommodating complex situations involving combinations of processes (e.g., population divergence, admixture, size change, and migration) and any number of populations and samples, while also offering a framework for comparing competing scenarios, estimating parameters, and computing bias and precision measures for any given scenario (e.g., Carnaval et al. 2009, Cornuet et al. 2008, Fagundes et al. 2007, Gray et al. 2008, Hickerson et al. 2009, Muster et al. 2009, Voight et al. 2005).

Population genetic parameters: parameters that characterize populations, such as effective size, gene flow, and time of divergence

Admixture: the founders of a new population comprise a mixture of individuals from two populations with different gene pools

Knowledge of the biological system under study is essential for deciding which approach to use for phylogeographic analyses. What might be an appropriate method for one study might not be for another, depending on the specifics of each species history and the type of data collected (Knowles 2004, Kuhner 2008, Wakeley 2003). Differences between the real population and its simplified representation can affect the results of analyses, limiting the utility of a particular model. In other cases, there may not be sufficient information contained in the data to estimate

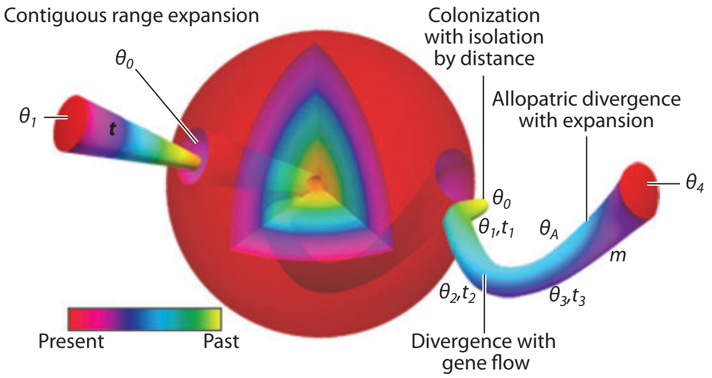

Figure 3

The universe of potential historical scenarios that might be considered in statistical phylogeography is huge. As represented here on the surface of the sphere, there is a continuous array of possible demographic processes and associated parameter values. Depending on the trajectory of a population's history over time (i.e., the path traversed through the sphere), a simple model such as contiguous range expansion (shown on the *left*) might provide an accurate simplified representation, whereas in other cases, a more complex model that incorporates different evolutionary processes at varying time points (shown on the *right*) would be needed for statistical phylogeographic analyses.

or distinguish between complex historical scenarios—there needs to be distinct detectable genetic signatures for competing hypotheses (Knowles & Maddison 2002, Nielsen & Beaumont 2009). Finally, it is the creativity of the individual researcher, not simply their computational prowess, that determines the insights gained from a phylogeographic analysis. The universe of potential historical scenarios is huge. Yet, only a limited subset of these possibilities is considered (**Figure 3**). How the dimensionality of the problem is parsed not only has a direct impact on the quality of a study's findings but is a challenge that also inspires inventive ways for asking phylogeographic questions.

Unprecedented Demographic Details

The shift toward multilocus data and a rigorous model-based perspective in statistical phylogeography provides detailed historical inference that was previously unattainable. This progress has been realized in three different ways: (*a*) refining hypotheses because of greater resolution, (*b*) extending the line of questioning beyond what could be addressed in the past, and (*c*) forging new partnerships in which the integration of additional information with genetic data provides a more powerful basis for inference. Here I highlight a few examples of the different ways in which recent developments have led to unprecedented demographic details, but in no way do these examples cover fully the varied and extensive progress advanced by the statistical phylogeographic perspective.

Increased resolution provides refinement of hypotheses. Understanding when divergence occurred is critical to assessing a number of biological questions, ranging from tests of whether humans rather than climate change caused the extinction of large mammals at the last glacial maximum (e.g., Barnes et al. 2007) to tests that the New World was peopled by a small number of founders (in this case, approximately 1% of the effective size of the estimated ancestral Asian population) during a recent colonization event, as suggested by archeological data of widespread Clovis sites (Hey 2005). Such questions were simply beyond the resolution of the markers and computational techniques used in the past. For example, instead of testing whether species divergence was promoted during the repeated fragmentation of species distributions induced by the glacial cycles (that is, was divergence Pleistocene versus pre-Pleistocene; Hewitt 1999), precise divergence-time estimates are capable of distinguishing glacial versus interglacial divergence, thereby providing insights into the geographic configurations promoting divergence (Carstens & Knowles 2006).

Demographic history: past processes structuring patterns of genetic variation; includes population size, migration rates, divergence times, and variation in these quantities over time

Detailed estimates of other population genetic parameters have also been essential to addressing various biological questions, including the types of demographic changes accompanying species divergence (e.g., Dolman & Moritz 2006, Lee & Edwards 2008, Won & Hey 2005). In a test of whether founder events in silvereyes promoted rapid genetic change, Clegg et al. (2002) assessed the relative impact of drift associated with the founding event itself, as opposed to long-term genetic drift, by comparing the patterns of genetic variation between recently founded island populations and older populations. With the aid of a historical colonization record containing detailed information on the age and timing of natural founding events, the genetic analyses showed that despite the morphological differentiation associated with island colonization (the observation that inspired Mayr's concept of genetic revolution by founder effect speciation), such differentiation could not be attributed to single founder events. Based on (*a*) the genetic diversities of island populations relative to their source populations, (*b*) estimates of the effective size of founder flocks from a Bayesian simulation study, and (*c*) the relatively poor fit of the microsatellite data to a "founder-event" model with a colonization bottleneck (compared to a "gradual-drift" model without a pronounced founder event), the genetic data suggest that differentiation in the silvereyes was promoted by sequential colonization events with fairly large numbers of individuals and/or long-term genetic drift in relatively small isolated populations.

New types of demographic inferences. Being able to distinguish between the different processes that could contribute to patterns of genetic variation is critical to a broad range of questions. For example, a low F_{st} value could reflect gene flow among populations or divergence in isolation among recently subdivided populations. These alternative explanations would carry very different ramifications if a study's goal was to determine whether habitat corridors provide an effective dispersal route (Biek et al. 2007, McRae 2006) or whether selection was strong enough to overcome the homogenizing effect of gene flow during species divergence (Dolman & Moritz 2006, Hey & Nielsen 2004). An incorrect inference in the former might mean a threatened species dependent on migration routes between isolated habitat patches suffers irreparably, or in the latter that a putative case of adaptive divergence is misidentified. Definitive answers to such questions were not possible without the computational advances of statistical phylogeographic approaches that can accommodate complex models involving multiple processes (Kuhner 2008).

Similarly, rather than relying on a priori delimitation of populations for tests of significant genetic differentiation, the number of populations contributing to the genotypes of individuals can be inferred jointly along with the allelic frequencies in each population (Pritchard et al. 2000). Such a framework was used to test a hypothesis of convergent evolution among ecomorphs of *Neospiza* buntings that inhabit similar environments among the Tristan islands (Ryan et al. 2007). Bayesian assignment probabilites of individuals revealed not only that the different ecomorphs correspond to distinct genetic clusters, but also that the similar ecomorphs from the different islands were not founded from a common ancestor. Thus, the close match of the morphological diversity with the available spectrum of seeds on each island had evolved independently.

Similar approaches, when accompanied by information on the proportional contribution of each inferred population to the genotypic makeup of each individual, can also be used to characterize the demographic history accompanying selective divergence, including the potential effects of drift-induced change (Harter et al. 2004; Knowles & Richards 2005). For example, quantification of the rates of genetic drift in domesticated and wild populations of sunflowers, based on comparison of the allele frequency estimates in wild and domesticated strains relative to the ancestral allelic frequencies, showed strong genetic drift during domestication in eastern North America (archeobotanical records had suggested an earlier and possibly independent origin of domestication in Mexico). Comparison with the wild populations, whose genetic composition was nearly identical to that of the wild progenitors of domesticated sunflowers, implies that evolutionary change in the domesticates has occurred at least 50-fold faster than the wild populations since they diverged from their common ancestor (Harter et al. 2004).

Population delimitation (e.g., Knowles & Carstens 2007) and individual assignment tests (e.g., Pritchard et al. 2000) analyses also have particular relevance to conservation, where they can be used to infer admixed populations or detect individuals that are immigrants. These types of demographic inferences have been important in determining whether populations managed as independent units are instead connected through dispersing individuals (as with polar bears; Paetkau et al. 1995), as well as determining migration patterns through identified immigrants (as with wolverines; Guillot et al. 2005).

Integrating genetic data for demographic inference with additional information. Inferences that rely on information from multiple independent sources are inherently more powerful. For example, inclusion of detailed spatial information [as with geographical information systems (GIS); Kidd & Ritchie 2006, Swenson 2008] provides more accurate inferences when estimating the genetic structure of populations, conducting population assignment, or detecting immigrants (Chen et al. 2007). Consideration of not just the genetics of the taxon of interest, but also the pattern of genetic variation of its viral pathogen, revealed details about the recent demographic history of the puma that was only apparent in the fast-evolving virus (Biek et al. 2007). The increased power of combining information is also apparent in comparative phylogeographic analyses for testing how abiotic and ecological processes drive evolution within communities. Such tests of temporal congruence (Hickerson et al. 2006, Leaché et al. 2007, Topp & Winker 2008) and congruence in colonization patterns (Hickerson & Meyer 2008) and refugia (Carnaval et al. 2009) from the simultaneous analysis of multiple species obtain more information from the data than would be obtained from multiple independent single-taxon analyses (reviewed in Hickerson et al. 2009). For example, comparison of alternative hypotheses of assemblage-level responses to late Quaternary climate change for frog species provided an assessment of the explicit demographic scenarios predicted for species distributions in the Brazilian Atlantic forest under paleoclimates and identified how climatically stable and unstable areas contributed to patterns of local endemism (Carnaval et al. 2009).

The goal to provide integrative information can itself serve as a source for methodological innovation and improvement. In addition to identifying the shortest possible path connecting populations, a clever approach that borrows from circuit theory in electrical engineering (McRae 2006) also identifies multiple geographic paths of probable gene flow among populations, as well as barriers to gene flow (**Figure 4**). In other cases, integration of information independent of the genetic data used to make demographic inference can be used to generate plausible hypotheses and select biologically informed population models. Consider that for any set of multiple populations there is an extremely large set of possible histories (**Figure 3**). These could include differing combinations of connections or routes of colonization among contemporary populations, as well as varying numbers or spatial configurations of populations in the past. The pertinent question is how to decide what model would be an appropriate representation? One increasingly common source of information for basing such decisions involves the use of ecological-niche models that

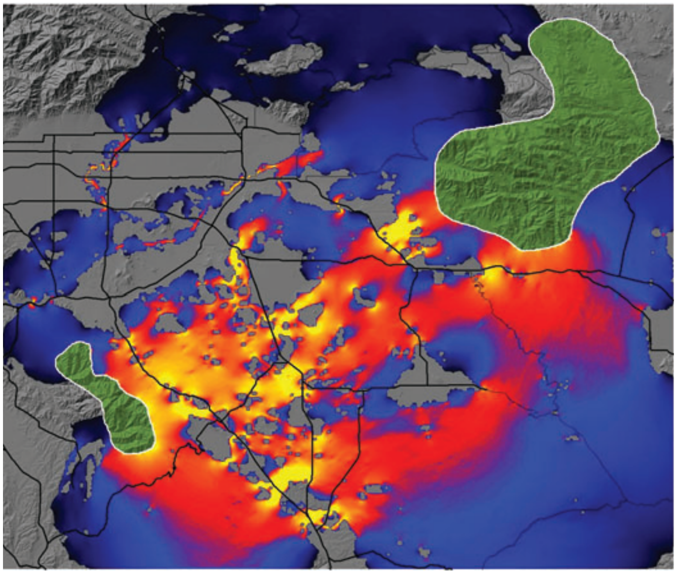

Figure 4

Phylogeographic analyses can be conducted with models derived from information independent of the genetic data, providing biologically informed hypotheses (or models). In this case, suites of landscape features and environmental data, and ideas borrowed from electrical engineering, were used to predict important routes of connectivity between core mountain lion habitat (*green*) in southern California, where areas of low current density (*blue*) are expected to have low densities of dispersing mountain lions, and areas with high current densities (*red*) identify important dispersal routes, with movement bottlenecks (*yellow*) designating areas most vulnerable to habitat destruction. Reproduced with permission from B. McRae, The Nature Conservancy, R. Hopkins, Live Oak Associates, Inc., and B. Dickson, Northern Arizona University.

can provide projections of current distributions, as well as past distributions when the equivalent paleoclimatic variables are used (Bigg et al. 2008, Carstens & Richards 2007, Knowles et al. 2007, Moussalli et al. 2009, Richards et al. 2007, Waltari et al. 2007).

Expanding Role in Evolutionary and Ecological Study

With the recent advances of statistical phylogeography, the estimates of population genetic parameters for any given species' history can be inferred with impressive precision and accuracy. However, the significance of these inferences extend far beyond obtaining an estimate of an effective population size, level of gene flow, timing of divergence, or population expansion or bottlenecking event. The use of genetic data for such demographic inference forms the basis for testing hypotheses and has enriched many areas of evolutionary and ecological study. These include the discovery of unknown phenomena, uncovering previously overlooked processes, and testing long-standing and fundamental questions in evolution and ecology that, until now, have gone answered. The few examples that follow highlight some of these contributions, but again, in no way do they provide exhaustive coverage of the incredibly broad impact of statistical phylogeography.

Effective population size (N_e): number of breeding individuals, assuming random mating, isolation, and stationarity, that determine the amount of genetic diversity and drift

Discovery of unknown phenomena. Here I refer to three examples that focus on different species from different geographic regions and that are based on analysis of different molecular markers (microsatellites, single-nucleotide polymorphisms, and DNA sequences plus amplified fragment length polymorphisms, respectively). However, in each case, phylogeographic study led to the discovery of unknown phenomena with ecological consequences (as with tropical tree population connectivity), genomic implications (as with sex-specific demographies in humans), and conservation concern (as with shifts in plant mating system, and the survival of Tahitian tree snail lineages previously thought to be extinct).

Example 1. Shade coffee is traditionally cultivated under a diverse canopy of overstory shade trees. The coffee benefits from increased fruit set from the broad spectrum of pollinators, but the finding that shade coffee farms also benefit the remnant tree populations across the tropical landscape was a surprise (Jha & Dick 2008). Interestingly, comparison of the patterns of genetic variation of tropical tree populations growing with and without coffee revealed that this benefit was mediated by the effect of the coffee on conserving resident bird populations, which help maintain gene flow and genetic diversity in the tropical trees surrounding the coffee plants.

Example 2. Selection and recombination are the typical focus for understanding the forces structuring patterns of genetic variation across the genome. However, comparison of genetic variation between the X chromosome and autosomes recently uncovered a signature of sex-specific demographic histories in humans (Bustamante & Ramachandra 2009). This finding suggests that social practices may have given rise to differences between the effective population size of men and women and, hence, also contributed to divergent patterns of evolution across the human genome.

Example 3. Certain life-history and ecological traits determine differential susceptibility to genetic erosion in fragmented habitats. Such characters are of special conservation concern, especially for species having to contend with rapidly changing environments. By means of a meta-analysis on the effects of habitat fragmentation on genetic diversity, Aguilar et al. (2008) discovered an alarming pattern that suggests fragmentation itself is driving a shift toward selfing, thereby further constraining the potential for adaptive change of rare plant species.

Similarly, taxa in certain geographic settings have been disproportionately subject to extinction on a rapid and staggering scale—namely, oceanic island endemics. Tahiti's tree snail populations are one such group with this unfortunate distinction, where introduction of introduced predators drove a precipitous decline of this once diverse fauna. Phylogeographic analysis of museum samples collected before the predator's introduction compared to wild and captive populations revealed that, despite significant decreases in genetic diversity, representatives of the different Tahitian *Partula* mitochondrial clades present in the museum samples have survived (Lee et al. 2007). With this discovery, and targeted conservation measures of montane refuge populations, the genetic diversity that was previously thought to have been lost might be preserved in the remaining remnant populations of the surviving species.

Uncovering previously overlooked processes. The studies featured here emphasize how a fundamental process essential to understanding the respective biological systems remained undetected without detailed phylogeographic study. In the first example, the temporal and geographic scale of the process fell outside a typical ecological study (that is, metapopulation dynamics at a continental and geological scale) and, therefore, was overlooked. In the second example, direct modeling of historical processes, as opposed to relying on diversity patterns to infer putative tropical refugia, reveals areas of conservation concern in the Brazilian Atlantic rainforest that would otherwise have gone undetected.

Example 1. Cold-adapted species show pronounced disjunct distributions with links between Arctic and alpine habitats, raising questions about the underlying population dynamics (e.g., degree of connectivity). Quantification of migration rates and tests under various demographic histories from genetic data suggests a species of wolf spiders from these regions share characteristics of a metapopulation, but the dynamics are occurring at a continental scale over geologic time (Muster et al. 2009). The patterns of habitat connectivity that dated to glacial maxima illustrate how an organism's response to past climate change is influenced by its ecology and life history, which, in this case, can produce population structures and dynamics that parallel the metapopulations typically studied on shorter ecological and local geographic scales.

Example 2. Although biodiversity hotspots have been mapped across the globe, the processes generating such species-rich regions are far from understood. Using ecological-niche models under paleoclimates and genetic analyses to test hypotheses about regional differences in community persistence, Carnaval et al. (2009) discovered a central region in the Brazilan Atlantic forest that served as a biodiversity refuge during climatic extremes. This discovery was made through a contrast of the relative contribution of temporally stable (refugial) and unstable (recently colonized) regions to this biodiversity hotspot using a new Bayesian analysis of multiple species that quantifies the fit of assemblage-level data to these different spatially explicit demographic scenarios. By using an approach that directly models historical processes, conservation priorities can be set that will not only preserve the species themselves, but also the historical processes that led to the diversity (Moritz 2002).

Testing longstanding unanswered questions. These last examples underscore the importance of detailed phylogeographic analysis. These are questions that cannot be addressed without studies of genetic variation. Not surprisingly, at the center of all these studies are questions about the species divergence process. As a historical and typically protracted process, studies on speciation, out of necessity, rely heavily on the inferences provided by phylogeographic study, including the integration of genetic data with other sources of information.

Example 1. The importance of geographic isolation in speciation is noncontroversial. In contrast, whether and how speciation might proceed in the face of gene flow have been the subject of much debate. Comparison of patterns of divergence from multiple loci has not only provided support for divergence with gene flow models (e.g., Hey & Nielsen 2004), but comparison of the gene trees among loci can be used to identify genes that play an active role in the speciation process because the alleles of such genes do not move across species boundaries, faithfully tracking the history of species divergence (e.g., the *Odysseus* locus in drosophila; Ting et al. 2000). Analyses of genetic data integrated with behavioral and morphological data can also provide critical tests of how reproductive isolation can be built up despite the threat of erosion by gene flow, especially in species where the actual genes involved in speciation cannot be identified because of limited genomic resources (in contrast to the example above). For example, detailed behavioral observation coupled with genetic data revealed how behavioral mechanisms—in this case, the imprinting of males and females on their host—promote cohesion of parasitic indigobird populations following a host shift (Sorensen et al. 2003). Analyses of microsatellite data identified significant genetic differentiation between indigobird species consistent with the assortative mating of indigobirds reared by a particular host species.

Example 2. Although species underwent rapid diversification during the dynamic Pleistocene, the effects of the glacial cycles appear to have varied widely among taxa (Weir & Schluter 2004). Detailed analysis of the demography of speciation in a group of montane grasshoppers showed that the repeated distributional shifts caused by the glacial cycles constrained diversification to taxa in which reproductive isolation apparently evolves very quickly (Carstens & Knowles 2006). For example, parameter estimates under a divergence-with-gene-flow model in *Melanoplus* grasshoppers from the sky islands of the Rocky Mountains in western North America indicated that divergence took place around 250,000 years ago, which means that the species experienced repeated opportunities for gene flow during subsequent climate-induced distributional shifts. Yet, very little gene flow occurred between codistributed species. Consequently, reproductive isolation must have evolved quickly to maintain incipient species boundaries in the face of multiple opportunities for gene flow.

Example 3. Detection of concordant suture zones—shared regions of secondary contact between geographically isolated lineages—among the various inhabitants of the tropical Australian rainforest shows how localized ecological adaptation has played a role in generating diversity patterns (Moritz et al. 2009). Spatial and temporal concordance (or lack thereof) of species divergence in patterns of genetic variation likewise show how differences in environmental suitability across a given region or biome, rather than past environmental conditions, underlie the different patterns of diversification observed among taxonomic groups (Byrne et al. 2008).

Stochasticity of genetic processes: the variance caused by genetic drift (random sorting of gene lineages) and mutational process (nucleotide substitution)

TRANSFORMATIVE THINKING STIMULATED BY PHYLOGEOGRAPHY

In addition to its expanding role in evolutionary and ecological study, developments within the realm of statistical phylogeography have also spread to areas outside its purview. As I highlight here (see also Avise 2009, Hickerson et al. 2009), the links to other fields in some cases stem from (*a*) drawing parallels to the benefits and limitations inherent in the way in which genetic data are used for phylogeographic inference, (*b*) applying phylogeographic principles, or (*c*) adopting the integrative tradition championed by the phylogeographic perspective.

Lessons from Phylogeography Usher in a New Phylogenetic Paradigm

The potential (and the actual probabilities) for discord between the gene tree of any single locus and the actual species history of divergence has long been recognized (e.g., Maddison 1997, Takahata 1989). This has led to a dramatic rise in the use of multilocus data in both phylogeographic and phylogenetic study. However, how these data are typically used differs substantially between these two disciplines.

The tradition of equating an estimated gene tree with the history of species divergence has predominated phylogenetics. As a result, sequences from different loci are typically combined (that is, concatenated) and analyzed in very much the same format as data from a single gene (albeit, with a partitioned model that accommodates locus-specific models of nucleotide evolution). Hence, the problem of differing gene trees across independent loci is purportedly eliminated by concatenation because only a single tree is estimated (but see Degnan & Rosenberg 2009).

Phylogeographic studies, in contrast, have adopted methodological approaches that explicitly take into account genealogical discord rather than forcing loci to conform to a single genealogical history (which would be an obvious violation of the patterns of genetic descent) (Hey & Machado 2003, Kuhner 2008, Wakeley 2007). In fact, the variance among loci provides information relevant to understanding evolutionary history (e.g., Nielsen & Wakeley 2001).

The lessons learned in phylogeography about the inherent benefits of explicitly considering the stochasticity of genetic processes when making historical inference, or conversely, the limitations of equating a gene tree with the history of species divergence are now being heeded in phylogenetics. Recent studies show that discordant gene trees retain significant phylogenetic signal and that it is just a matter of extracting this information (Edwards et al. 2007, Knowles & Chan 2008, McCormack et al. 2009). As with phylogeography, coalescent-based approaches have been developed (e.g., Kubatko et al. 2009, Liu & Pearl 2007, Maddison & Knowles 2006) such that data from multiple loci can be used to obtain a direct estimate of the history of divergence—the species tree—as opposed to focusing on the idiosyncrasies of individual gene trees (**Figure 5**) (e.g., Belfiore et al. 2008, Brumfield et al. 2008, Carstens & Knowles 2007, Kubatko & Gibbs 2009, Liu et al. 2008, Oneal et al. 2009).

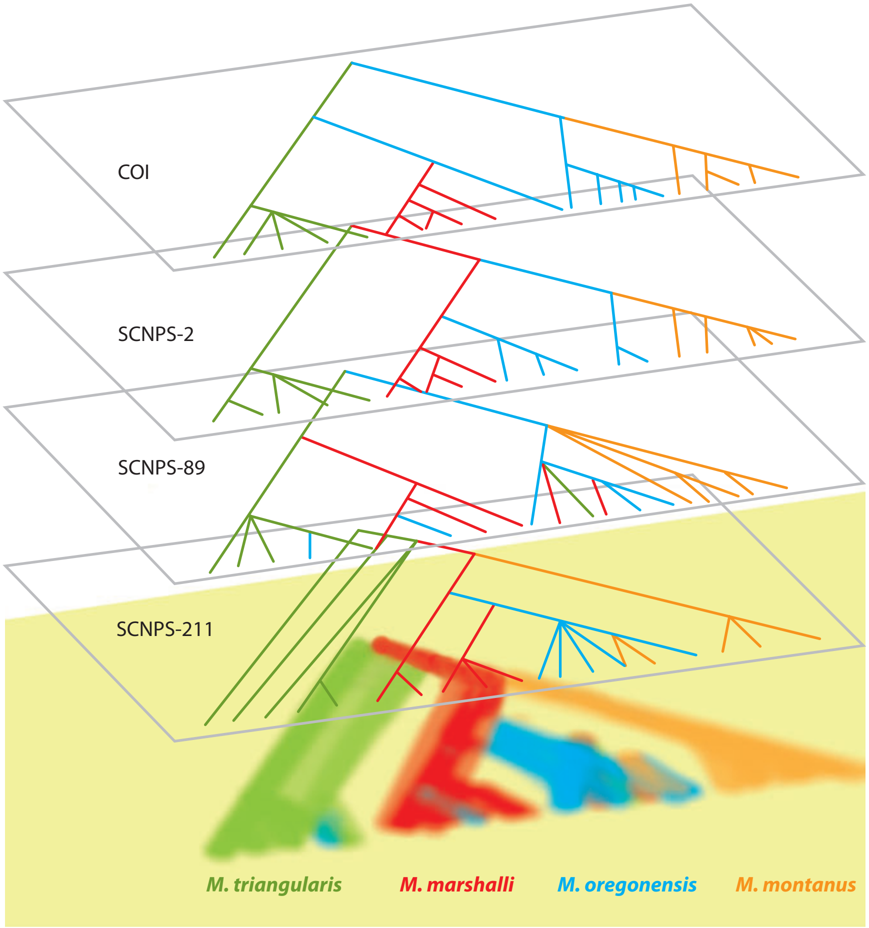

Figure 5

Rather than focusing on the idiosyncrasies of the tree topologies of individual loci, where coalescent and mutational variance can give rise to gene tree discord, as shown here for a sample of independent loci sequenced in montane grasshoppers (from Carstens & Knowles 2007), new coalescent-based approaches provide direct estimates of the underlying history of divergence—the actual species tree. Extracting the common historical signal of species divergence in the montane grasshoppers from the gene trees of the independent loci, despite widespread incomplete lineage sorting, highlights how this advance provides resolution of the traditionally unresolvable, estimating the pattern of species splitting during recent evolutionary radiations.

Tree-Based Thinking for Understanding the Drivers of Infectious Disease

The shape of a genealogical tree, in particular its topology and branch lengths, can be very revealing about the demographic past (Wakeley 2003, 2007). Such tree-based thinking has permeated the study of the drivers of infectious disease (e.g., Adams et al. 2006, Koelle et al. 2006, Lindesmith et al. 2008, Rambaut et al. 2008), where the cause of interannual variability in disease outbreaks is typically approached through the detailed study of the time series of infected individuals (Cobey & Koelle 2008). However, such ecologically based models (with separate susceptible, infected, and recovered classes in the host population, as well as specified transition rates between

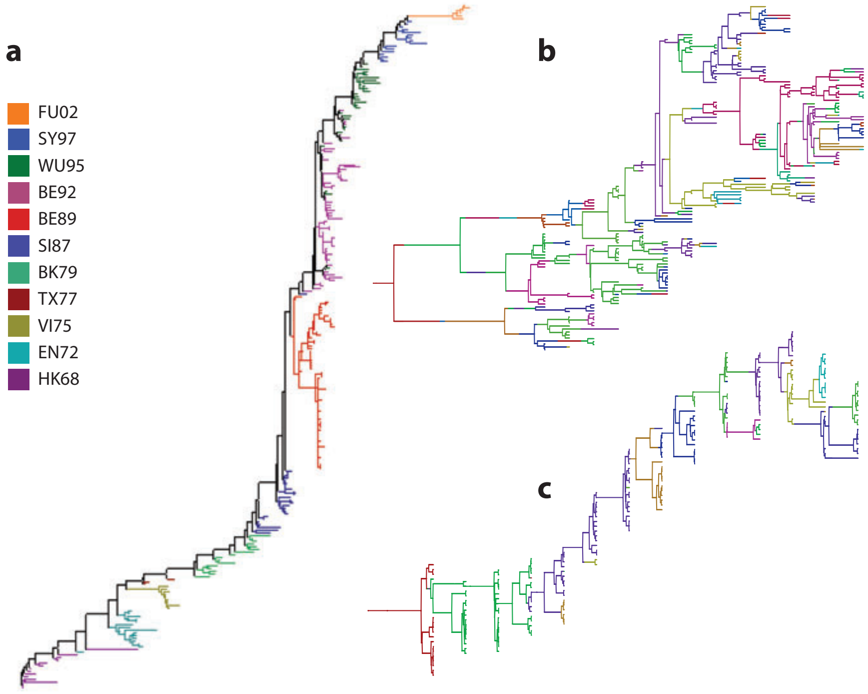

Figure 6

Genetic profile of influenza, or specifically, (*a*) a gene tree of the virus's hemagglutinin glycoprotein (with known antibody binding sites corresponding to different antigenic properties labeled) that shows limited variation within clusters (which are color coded) and episodic contraction of genetic diversity during cluster transitions. This tree shape contrasts with a characteristic pattern under neutral expectations (*b*), unless unrealistically low population sizes are assumed (*c*) in order to obtain the skewed or ladder-like topology observed in the empirical flu data (*a*). Modified from Koelle et al. (2006); epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans (*Science* 314:1898–903). Reprinted with permission from AAAS.

these classes) fall short of predicting the dynamics of rapidly evolving pathogens (Koelle et al. 2006).

Studying the phylodynamics of a disease (**Figure 6**) provides a means for disentangling the ecological and immunological processes driving disease dynamics and evolution, which is critical for managing infectious disease and mitigating the toll they take on human populations every year (Cobey & Koelle 2008, Ferguson et al. 2003, Grenfell et al. 2004). For example, characterization of the genetic patterns of influenza (based on sequence analysis of the virus's hemagglutinin glycoprotein, which is involved in antibody binding) has revealed an unusual tree shape (**Figure 6**). Despite high mutation rates, influenza surprisingly exhibits limited standing variation at any single point in time, with episodic contraction of the diversity within influenza clusters during abrupt transitions between antigenic strains (**Figure *6a***). A coupled epidemiological transmission model and genotype-phenotype model implemented within a neutral network reproduce influenza's characteristic genetic and antigenic patterns of boom and bust viral evolution (Koelle et al. 2006). In this case, the phylodynamics suggest an evolutionary history in which sequence changes in the hemagglutinin correspond to essentially neutral or almost neutral phenotypic space (that is, no change in antigenic properties, or phenotypic stasis) and abrupt changes between flu clusters occur as diffusion across a neutral network provides access to adjacent genotypic space that enables dramatic phenotypic change (that is, stepwise emergence of new clusters that replace old clusters with differing antigenic properties) (**Figure 6*a***). Such studies highlight that, for rapidly evolving pathogens, the drivers of disease cannot be understood without reference to how the pathogens evolve and how their evolutionary and epidemiological dynamics interact.

Integrative Phylogeographic Tradition Applied to the Analysis of Non-Neutral Variation

As a field that routinely seeks out information independent of the genetic data upon which a demographic inference is based (not to mention, incorporating an explicit spatial context), phylogeographic analysis of neutral genetic variation stands in stark contrast to the typically sterile pursuit of detecting selection in molecular data (e.g., the tabulation of dN/dS ratios; McDonald & Kreitman 1991). Nonetheless, the spread of phylogeography's integrative tradition to analyses of non-neutral variation is becoming increasingly evident.

With broad sampling of individuals across different environments, data on the degree of differentiation across large suites of loci (that is, interlocus contrasts; Beaumont & Balding 2004, Beaumont & Nichols 1996) can be used to identify loci whose divergence is correlated with environmental or ecological transitions. For nonmodel organisms, such approaches provide a means for searching out the loci underlying adaptive divergence (e.g., intertidal divergence in snails, Wilding et al. 2001; floral divergence in hawk-moth pollinated plants, Herrera & Bazaga 2008; adaptation to altitude in frogs, Bonin et al. 2006).

An integrative perspective might also provide insights into how adaptive divergence might proceed. For example, a virtually untested, yet critically important determinant of whether environmental change will out-pace the potential for adaptive evolution is whether selection acts on new mutation versus standing genetic variation. Tests that rely solely on genetic data for distinguishing between these alternative genetic sources of adaptation are difficult and suffer from limited power (Barrett & Schluter 2008). However, in a broader context, where not only the spatial distribution of phenotypes and selected loci but also the distributional patterns of neutral variation are compared, adaptive divergence arising from pre-existing variation or via new mutation might be distinguished. Such pioneering work on the source of adaptive changes associated with the transition from a marine to freshwater habitat in the three-spine stickleback (Colosimo et al. 2005) reinforces the virtues of the integrative tradition upon which phylogeographic study derives much of its power. Sequencing of individuals from different populations (as shown on the map, **Figure 7**) indicated a common origin for the allele conferring the adaptive phenotype of low-plate armor in stickleback fish from freshwater habitats; the molecular basis of the adaptive phenotype was traced to the gene ectodysplasin (*Eda*) using a quantitative trait loci mapping procedure (**Figure 7*b***). Geographic sampling also revealed that members of this clade of low-plate alleles are present at low frequencies in marine fish, suggesting that standing genetic variation (that is, pre-existing ancestral variation) may have been the source for rapid, parallel adaptive phenotypic change in these freshwater sticklebacks, as opposed to new mutations. Comparison of the *Eda* gene tree (**Figure 7*c***) from marine and freshwater fish to the topology of a tree for random nuclear loci (**Figure 7*d***) rules out the possibility that all present-day low-plated populations are derived from a single population sharing alleles at most genes, thereby confirming that parallel evolution has occurred by repeated selection of pre-existing *Eda* alleles (Colosimo et al. 2005).

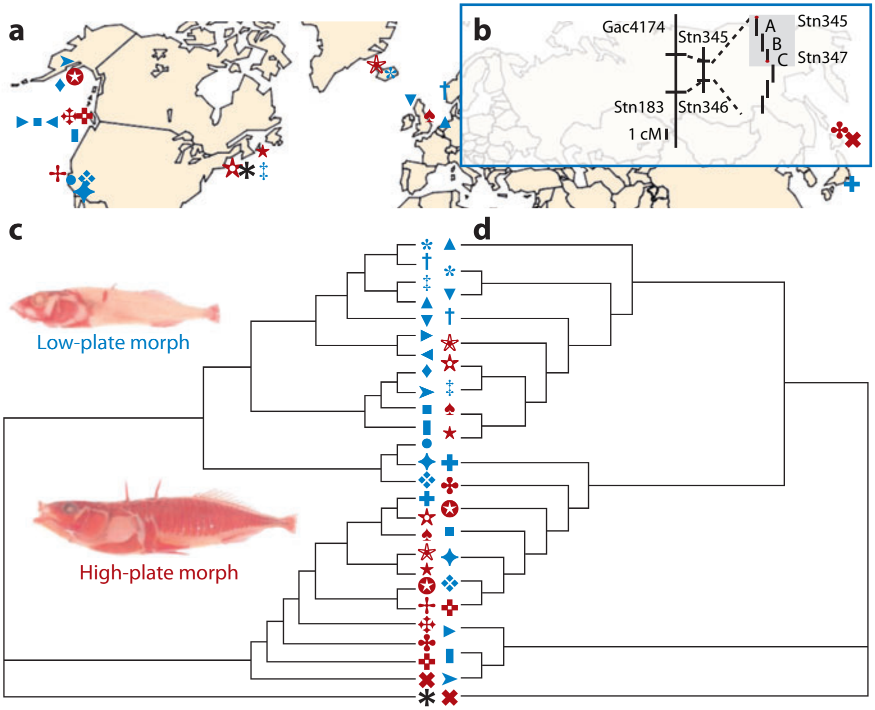

Figure 7

Geographic distribution of haplotypes across marine and freshwater stickleback populations (*a*) that exhibit an adaptive shift toward low-plate armor in the derived freshwater populations; the molecular basis for which was traced to the gene ectodysplasin (*Eda*) using a quantitative trait loci mapping procedure (*b*). Comparison of the tree topology for *Eda* (*c*) to one based on random nuclear loci (*d*) shows that the source of parallel adaptive divergence arose from selection acting on pre-existing ancestral variation in the fish, not new mutation. Modified from Colosimo et al. (2005); widespread parallel evolution in sticklebacks by repeated fixation of *Eda* alleles (*Science* 307:1928–33). Reprinted with permission from AAAS.

SUMMARY POINTS

1. The unprecedented demographic details now provided by statistical phylogeographic study highlight the impact that the shift toward a model-based multilocus perspective has had on historical inference.
2. The population genetic parameters and geographic patterns of genetic variation inferred from phylogeographic analyses are not just relevant to testing hypotheses about the demographic past. They are also the basis for addressing a broad range of questions in evolutionary and ecological study, and have enriched other disciplines as various aspects of the phylogeographic perspective are now applied to areas outside of the field of phylogeography.
3. Integration of the genetic data used in phylogeographic analyses with ancillary information, whether it is ecological, morphological, or functional in nature, is key to maximizing evolutionary and ecological insights. Analyses of genetic data without this broader context will necessarily be limited.

FUTURE ISSUES

1. How can a large universe of potential histories be explored, such that their relative probabilities are considered jointly (e.g., a posterior probability distribution of phylogeographic histories), as opposed to choosing specific models that might be compared or analyzed sequentially?
2. Genomic data, which include not only bulk analysis of lots of markers, but also consider other genomic attributes, such as the chromosomal position or corresponding phenotypes associated with specific markers, need to be integrated into a phylogeographic framework such that the interplay between demographic processes and selection can be tested.
3. A framework is needed in which the potential gains of model complexity (considering a broad array of demographic processes that might vary over time and among populations) can be evaluated relative to the cost of applying complex models. For example, what level of complexity can be tolerated for a given amount of data, such that the reliability of phylogeographic inference is not compromised by the high dimensionality of a complicated model?
4. Synthesis of comparative information across taxa will be invaluable for evaluating different historical and ecological theories about diversity patterns and community assembly from different regions and ecosystems.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Many ideas in this review have benefited from discussions with colleagues, especially Wayne Maddison, Mike Hickerson, Bryan Carstens, and Craig Moritz. I would also like to thank Bob Payne and Josepha Kurdziel for helpful comments, as well as Qixin He, Lucy Tran, Kelsey Gibbons, and Trevor Bedford for their contributions to the figures. This work was supported by National Science Foundation grants (DEB-04–47224 and DEB-07–15487) to L.L.K.

LITERATURE CITED

Adams B, Holmes EC, Zhang C, Mammen MP, Nimmannitya S, et al. 2006. Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok. *Proc. Natl. Acad. Sci. USA* 103:14234–39

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J. 2008. Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol. Ecol.* 17:5177–88

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annu. Rev. Ecol. Syst.* 33:707–40

Avise JC. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos* 63:62–76

Avise JC. 2009. Phylogeography: retrospect and prospect. *J. Biogeogr.* 36:3–15

Avise JC. 2000. *Phylogeography: The History and Formation of Species*. Boston: Harvard Univ. Press

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, et al. 1987. Intraspecific phylogeography—the mitochondrial-DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18:489–522

Barnes I, Shapiro B, Kuznetsova T, Sher A, Guthrie D, et al. 2007. Genetic structure and extinction of the woolly mammoth *Mammuthus primigenius*. *Current Biol.* 17:1072–75

Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23:38–44

Beaumont MA, Balding DJ. 2004. Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* 13:969–80

Beaumont MA, Nichols RA. 1996. Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. London Ser. B* 263:1619–26

Belfiore NM, Liu L, Moritz C. 2008. Multilocus phylogenetics of a rapid radiation in the genus Thomomys (Rodentia: Geomyidae). *Syst. Biol.* 57:294–310

Biek R, Drummond AJ, Poss M. 2007. A virus reveals population structure and recent demographic history of its carnivore host. *Science* 311:538–41

Bigg GR, Cunningham CW, Ottersen G, Pogson GH, Wadley MR, Williamson P. 2008. Ice-age survival of Atlantic cod: agreement between palaeoecology models and genetics. *Proc. R. Soc.London Ser. B* 275:163–72

Bonin A, Taberlet P, Miaud C, Pompano F. 2006. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Mol. Biol. Evol.* 23:773–83

Brumfield R, Liu L, Lum D, Edwards SV. 2008. Comparison of species tree methods for reconstructing the phylogeny of bearded manakins (Aves: Pipridae, Manacus) from multilocus sequence data. *Syst. Biol.* 57:719–31

Bustamante CD, Ramachandran S. 2009. Evaluating signatures of sex-specific processes in the human genome. *Nat. Genet.* 41:8–10

Byrne M, Yeastes DK, Joseph L, Kearney M, Bowler J, et al. 2008. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Mol. Ecol.* 17:4398–417

Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C. 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science* 323:785–89

Carstens BC, Knowles LL. 2006. Shifting distributions and speciation: genomic resolution of species divergence during rapid climate change. *Mol. Ecol.* 16:619–27

Carstens BC, Knowles LL. 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. *Syst. Biol.* 56:1–12

Carstens BC, Richards CL. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution* 61:1439–54

Chen CA, Durand E, Forbes F, Francois O. 2007. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol. Ecol.* 7:747–56

Clegg SM, Degnan SM, Kikkawa J, Moritz C, Estoup A, Owens IPF. 2002. Genetic consequences of sequential founder events by an island-colonizing bird. *Proc. Natl. Acad. Sci. USA* 99:8127–32

Cobey S, Koelle K. 2008. Capturing escape in infectious disease dynamics. *Trends Ecol. Evol.* 23:572–77

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307:1928–33

Cornuet J-M, Santos F, Beaumont MA, Robert CP, Marin J-M, et al. 2008. Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24:2713–19

Dolman G, Moritz C. 2006. A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evolution* 60:573–82

Degnan J, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference, and the multispecies coalescent. *Trends. Ecol. Evol.* 24:332–40

Edwards SV, Liu L, Pearl DK. 2007. High-resolution species trees without concatenation. *Proc. Natl. Acad. Sci. USA* 104:5936–41

Estoup A, Beaumont M, Sennedot F, Moritz C, Cornuet J-M. 2004. Genetic analysis of complex demographic scenarios: spatially expanding populations of the cane toad, *Bufo marinus*. *Evolution* 58:2021–36

Excoffier L, Heckel G. 2006. Computer programs for population genetics data analysis: a survival guide. *Nat. Rev. Genet.* 7:745–58

Fagundes NJR, Ray N, Beaumont M, Neuenschwander S, Salzano FM, et al. 2007. Statistical evaluation of alternative models of human evolution. *Proc. Natl. Acad. Sci. USA* 45:17614–19

Ferguson NM, Galvani AP, Bush RM. 2003. Ecological and immunological determinants of influenza evolution. *Nature* 422:428–33

Gray DA, Huang H, Knowles LL. 2008. Molecular evidence of a peripatric origin for two sympatric species of field crickets (*Gryllus rubens* and *G. texensis*) revealed from coalescent simulations and population genetic tests. *Mol. Ecol.* 17:3826–55

Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, et al. 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303:327–32

Guillot G, Estoup A, Mortier F, Cosson JF. 2005. A spatial statistical model for landscape genetics. *Genetics* 170:1261–80

Harrigan RJ, Mazza ME, Sorenson MD. 2008. Computation vs cloning: evaluation of two methods for haplotype determination. *Mol. Ecol.* 8:1239–48

Harter AV, Gardner KA, Falush D, Lentz DL, Bye RA, Rieseberg LH. 2004. Origin of extant domesticated sunflowers in eastern North America. *Nature* 430:2001–5

Herrera CM, Bazaga P. 2008. Population-genomic approach reveals adaptive floral divergence in discrete populations of a hawk moth-pollinated violet. *Mol. Ecol.* 17:5378–90

Hewitt GM. 1999. Post-glacial recolonization of European biota. *Biol. J. Linn. Soc.* 68:87–112

Hey J. 2005. The number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS Biol.* 3:965–75

Hey J, Machado CA. 2003. The study of structured populations—new hope for a difficult and divided science. *Nat. Rev. Genet.* 4:535–43

Hey J, Nielsen R. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167:747–60

Hey J, Won Y-J, Sivasundar A, Nielsen R, Markert JA. 2004. Using nuclear haplotypes with microsatellites to study gene flow between recently separated cichlid species. *Mol. Ecol.* 13:909–19

Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, et al. 2009. 20 years after Avise et al. 1987: comparative phylogeography fulfilling original promise by integrating with emerging fields. *Mol. Phylog. Evol.* In press

Hickerson MJ, Meyer C. 2008. Testing comparative phylogeographic models of marine vicariance and dispersal using a hierarchical Bayesian approach. *BMC Evol. Biol.* 8:1–18

Hickerson MJ, Stahl E, Lessios HA. 2006. Test for simultaneous divergence using approximate Bayesian computation. *Evolution* 60:2435–53

Jha S, Dick CW. 2008. Shade coffee farms promote diversity of native trees. *Curr. Biol.* 18:R1126–28

Kidd DM, Ritchie MG. 2006. Phylogeographic information systems: putting the geography into phylogeography. *J. Biogeogr.* 33:1851–65

King MG, Roalson EH. 2009. Discordance between phylogenetics and coalescent-based divergence modeling: exploring phylogeographic pattern of speciation in the *Carex macrocephala* species complex. *Mol. Ecol.* 18:468–82

Knowles LL. 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Mol. Ecol.* 10:691–701

Knowles LL. 2004. The burgeoning field of statistical phylogeography. *J. Evol. Biol.* 17:1–10

Knowles LL. 2008. Why does a method that fails continued to be used? *Evolution* 62:2713–17

Knowles LL, Carstens BC. 2007. Delimiting species without monophyletic gene trees. *Syst. Biol.* 56:887–95

Knowles LL, Carstens BC, Keat ML. 2007. Coupled genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Curr. Biol.* 17:1–7

Knowles LL, Chan Y-H. 2008. Resolving species phylogenies of recent evolutionary radiations. *Ann. Mo. Bot. Gard.* 95:224–31

Knowles LL, Maddison WP. 2002. Statistical phylogeography. *Mol. Ecol.* 11:2623–35

Knowles LL, Richards CL. 2005. Genetic drifts importance during Pleistocene divergence as revealed by analysis of genomic variation. *Mol. Ecol.* 14:4023–32

Koelle K, Cobey S, Grenfell B, Pascual M. 2006. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science* 314:1898–903

Kubatko L, Carstens BC, Knowles LL. 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* 25:971–73

Kubatko LS, Gibbs HL. 2009. Phylogenetic relationships among *Sistrurus* rattlesnakes inferred from multiple loci. *Syst. Biol.* In press

Kuhner MK. 2008. Coalescent genealogy samplers: windows into population history. *Trends Ecol. Evol.* 24:86–93

Leaché A, Crews SA, Hickerson MJ. 2007. Did an ancient seaway across Baja California cause community isolation? *Biol. Lett.* 3:646–50

Lee TW, Burch JB, Jung YH, Coote T, Pearce-Kelly P, ÓFoighil D. 2007. Tahitian tree snail mitochondrial clades survived recent mass extirpation. *Curr. Biol.* 17:R502–3

Lindesmith LC, Donaldson EF, Lobue AD, Cannon JL, Zheng DP, et al. 2008. Mechanisms of GII.4 norovirus persistence in human populations. *PLoS Med.* 5:e31

Lee JY, Edwards SV. 2008. Divergence across Australia's Carpentarian barrier: statistical phylogeography of the red-backed fairy wren (*Malurus melanocephalus*). *Evolution* 62:3117–34

Lemmon AR, Lemmon EM. 2008. A likelihood framework for estimating phylogeographic history on a continuous landscape. *Syst. Biol.* 57:544–61

Liu L, Pearl DK. 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Syst. Biol.* 56:504–14

Liu L, Pearl DK, Brumfield, Edwards SV. 2008. Estimating species trees using multiple-allele sequence data. *Evolution* 62:2080–91

Maddison WP. 1997. Gene trees in species trees. *Syst. Biol.* 46:523–36

Maddison WP, Knowles LL. 2006. Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol.* 55:21–30

McCormack JE, Bowen BS, Smith TB. 2008. Integrating paleoecology and genetics of bird populations in two sky island archipelagos. *BMC Biol.* 6:1–12

McCormack JE, Huang H, Knowles LL. 2009. Breaking them down and building them up: evaluating the accuracy of maximum-likelihood estimates of species trees. *Syst. Biol.* In press

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351:652–54

McRae BH. 2006. Isolation by resistance. *Evolution* 60:1551–61

Moritz C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst. Biol.* 51:238–54

Moritz C, Hoskin CJ, MacKenzie JB, Phillips BL, Tonione M, et al. 2009. Identification and dynamics of a cryptic suture zone in tropical rainforest. *Proc. R. Soc. London Ser. B* 276:1235–44

Moussalli A, Moritz C, Williams SE, Carnaval AC. 2009. Variable responses of skinks to a common history of rainforest fluctuation: concordance between phylogeography and palaeo-distribution models. *Mol. Ecol.* 18:483–99

Muster C, Maddison WP, Uhlmann S, Berendonk TU, Vogler AP. 2009. Artic-alpine distributions—metapopulations on a continental scale? *Am. Nat.* 173:313–26

Nielsen R, Beaumont MA. 2009. Statistical inference in phylogeography. *Mol. Ecol.* 18:1034–47

Nielsen R, Wakeley J. 2001. Distinguishing migration from isolation: a Markov-chain Monte Carlo approach. *Genetics* 158:885–96

Oneal E, Otte D, Knowles LL. 2009. Testing for biogeographic mechanisms promoting divergence in Caribbean crickets (genus *Amphiacusta*). *J. Biogeogr.* In press

Paetkau D, Calvert W, Stirling I, Strobeck C. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* 4:347–54

Pritchard JK, Stephens M, Donnelly P. 2000. Inferences of population structure using multilocus genotype data. *Genetics* 155:945–59

Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. 2008. The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453:615–19

Richards CL, Carstens BC, Knowles LL. 2007. Distribution modeling and statistical phylogeography: an integrative framework for testing biogeographic hypotheses. *J. Biogeogr.* 34:1833–45

Ryan PG, Bloomer P, Moloney CL, Grant TJ, Delport W. 2007. Ecological speciation in South Atlantic island finches. *Science* 315:1420–23

Sorenson MD, Sefc KM, Payne RB. 2003. Speciation by host switch in brood parasitic indigobirds. *Nature* 424:928–31

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, et al. 2007. Putting the 'landscape' in landscape genetics. *Heredity* 98:128–42

Swenson NG. 2008. The past and future influence of geographic information systems on hybrid zone, phylogeographic and speciation research. *J. Evol. Biol.* 21:421–34

Takahata N. 1989. Gene genealogy in 3 related populations—consistency probability between gene and population trees. *Genetics* 122:957–66

Ting C-T, Tsaur S-C, Wu C-I. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proc. Natl. Acad. Sci. USA* 97:5313–16

Topp CM, Winker K. 2008. Genetic patterns of differentiation among five landbird species on the Queen Charlotte Islands, British Columbia. *Auk* 125:461–72

Voight BF, Adams AM, Frisse LA, Qian Y, Hudson RR, Di Rienzo A. 2005. Interrogating multiple aspects of variation in a full resequencing data set to infer human population size changes. *Proc. Natl. Acad. Sci. USA* 102:18508–13

Wakeley J. 2003. Inferences about the structure and history of populations: coalescents and intraspecific phylogeography. In *The Evolution of Population Biology*, ed. R Singh, M Uyenoyama, pp. 193–215. Cambridge: Cambridge Univ. Press

Wakeley J. 2004. Recent trends in population genetics: more data! More math! Simple models? *J. Heredity* 95:397–405

Wakeley J. 2007. *Coalescent Theory: An Introduction*. New York: Roberts and Company Publishers

Waltari E, Hijmans RJ, Peterson AT, Nyári AS, Perkins SL, Guralnick RP. 2007. Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS One* 2:e563

Weir JT, Schluter D. 2004. Ice sheets promote speciation in boreal birds. *Proc. R. Soc. London Ser. B* 271:1881–87

Wilding CS, Butlin RK, Grahame J. 2001. Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *J. Evol. Biol.* 14:611–19

Won Y-J, Hey J. 2005. Divergence population genetics of chimpanzees. *Mol. Biol. Evol.* 22:297–307

Annual Review of Ecology, Evolution, and Systematics

Volume 40, 2009

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Ecology, Evolution, and Systematics* articles may be found at http://ecolsys.annualreviews.org/errata.shtml