Beyond Reproductive Isolation: Demographic Controls on the Speciation Process

Michael G. Harvey,¹ Sonal Singhal,² and Daniel L. Rabosky³

¹Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996, USA; email: mgh272@gmail.com
²Department of Biology, California State University, Dominguez Hills, Carson, California 90747, USA; email: sonal.singhal1@gmail.com
³Museum of Zoology and Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109, USA; email: drabosky@umich.edu

Abstract

Studies of speciation typically investigate the evolution of reproductive isolation between populations, but several other processes can serve as key steps limiting the formation of species. In particular, the probability of successful speciation can be influenced by factors that affect the frequency with which population isolates form as well as their persistence through time. We suggest that population isolation and persistence have an inherently spatial dimension that can be profitably studied using a conceptual framework drawn from metapopulation ecology. We discuss models of speciation that incorporate demographic processes and highlight the need for a broader application of phylogenetic comparative approaches to evaluate the general importance of population isolation, persistence, and reproductive isolation in speciation. We review diverse and nontraditional data sources that can be leveraged to study isolation and persistence in a comparative framework. This incorporation of spatial demographic information facilitates the integration of perspectives on speciation across disciplines and timescales.

Keywords
microevolution, macroevolution, population isolation, population persistence, metapopulation, comparative methods
1. INTRODUCTION

What processes might explain the spectacular diversity of organisms on Earth—from birds and insects to flowering plants and fungi, microscopic bacteria and protozoans? To explore this question, scientists have assumed one of two largely distinct research paradigms. In what might be termed a microevolutionary approach, many researchers focus on the origin of species through the lens of reproductive and geographic isolation between contemporary populations. This research program frequently includes assessments of hybrid sterility and fitness, reciprocal transplant experiments, mate choice trials, and dissection of the genetic architecture underpinning key species-specific traits. Other researchers take a macroevolutionary approach, studying ecological, organismal, and historical factors associated with the dynamics of speciation as measured using phylogenetic or palaeontological data. Unfortunately, these micro- and macroevolutionary research programs rarely meet in dialogue, and this gap is consequential. We might even question whether the two are studying the same phenomena.

A potential bridge between speciation as studied at micro- and macroevolutionary scales involves the origins and ultimate fates of populations that potentially give rise to new species. Since its inception, evolutionary biology has viewed populations as the fuel for new species formation (Darwin 1859, Reznick & Ricklefs 2009). Speciation is contingent on the factors that affect the origin and persistence of populations as well as the advent of reproductive isolation (Allmon 1992). Moreover, the frequency of population origin and the duration of population persistence are potentially critical rate-limiting steps for the speciation process (Levin 1995, Schluter 2016). In nature, these processes have an explicitly spatial dimension—they play out across an arena of local populations. Geographic isolation is widely understood to play an important role in speciation (Mayr 1963), and it is a focus of research in biogeography and phylogeography (e.g., Avise 2000). Population persistence, at least with regard to the maintenance of discontinuity, is a focus of hybrid zone research (e.g., Barton & Hewitt 1985). However, research that incorporates the full range of spatial dynamics of populations explicitly into a speciation framework is lacking.

We argue that a metapopulation perspective offers a natural framework for conceptualizing the spatial dimension of population dynamics associated with species formation. In its broadest sense, metapopulation theory views species as multiple, semi-isolated populations distributed across space. Here, we assume a broad-sense definition of metapopulations (Leibold & Chase 2018), in which local populations exhibit at least partial demographic independence from other such populations but no requirement for populations to represent discrete units embedded within an inhospitable matrix is made. Metapopulation structure is nearly universal by this definition, as almost all species exhibit some degree of spatial population structure resulting from dispersal, landscape heterogeneity, and local adaptation (Harrison & Taylor 1997). We recognize that a stricter view of metapopulations has been key to many theoretical advances in ecology (Levins 1969, Gilpin & Hanski 1991, Gotelli 1991), but the broad-sense definition is a useful heuristic tool for addressing speciation. Viewed through time, the dynamics of metapopulations capture the demographic processes of establishment and persistence. A dynamic metapopulation perspective is compatible with views from the literature on geographical (e.g., allopatric) speciation, allowing us to build on insights into isolation and persistence from biogeography, phylogeography, and hybrid zone research. Moreover, this spatially explicit framework permits investigation of the roles of both reproductive isolation and demographic processes in speciation, an expansion that is essential to bridge key patterns and processes studied at microevolutionary and macroevolutionary scales.

Here, we discuss the integration of demographic processes in speciation research using a spatially explicit framework. We first review potential controls or rate-limiting steps in the speciation
process, highlighting the demographic processes of isolation and persistence. We briefly review conceptual work in speciation, discussing ways in which demography typically enters speciation models. We then summarize recent models and simulation work that explicitly includes population isolation and/or persistence in the speciation process. We review empirical evidence for links between these processes and macroevolutionary dynamics, a key qualification for their importance in the origin of diversity. We explore new sources of spatiotemporal data that can potentially inform speciation studies, drawing from population genetics, phylogeography, ecology, and conservation biology. Finally, we suggest future directions for incorporating demographic processes into the study of speciation.

2. CONTROLS ON THE SPECIATION PROCESS IN A METAPOPULATION FRAMEWORK

The origin of species involves multiple processes, any one of which might serve as a limiting control on the probability of successful speciation. In this review, we define speciation as the splitting of one evolutionary lineage to produce two or more. This expansive definition allows for consideration of processes important to speciation aside from reproductive isolation (Figure 1). The framework used here is based on the ideas of Mayr (1963) and further developed in the recent paleontological and macroevolutionary literature (Allmon 1992, Dynesius & Jansson 2014, Allmon & Sampson 2016, Rabosky 2016). We recognize four controls: the isolation of populations; the evolution of reproductive barriers between populations; sufficient ecological differentiation for coexistence of differentiated populations; and the persistence of populations for a sufficient time, allowing for completion of reproductive isolation and ecological differentiation (Figure 1). The controls are not chronological, and several might be operating simultaneously in a given instance.

Figure 1

(a) A concept map of controls on the speciation process (based loosely on Allmon 1992). Population isolation, persistence, reproductive isolation (RI), and ecological differentiation are all required for the completion of speciation. There are possible feedbacks between many of these controls (gray arrows). Ecological differentiation between populations, for example, might accelerate the rate at which RI evolves. (b) The bar plot summarizes open questions/future directions posed in reviews of microevolutionary processes impacting speciation (number of papers = 55; time span, 2008–2018; Supplemental Table 1). Each question was categorized to best reflect which stage of the speciation process it reflected. “Other” represents less common question types that are not the focus of the current review (e.g., hybridization, polyploidization). This summary indicates that the current focus of speciation research is largely on RI and speciation genetics.
Speciation control: any part of the speciation process that limits the probability that one species will split into two.

Figure 2
A visual representation of a metapopulation conceptual framework for speciation, with a map of the metapopulation of a single species at present ($t = 0$) on the left and a population history for the metapopulation of that species on the right. In the population history, color shading denotes the acquisition within a population of unique ecological traits that are not shared with other contemporaneous populations, and pattern shading denotes the acquisition of mechanisms of reproductive isolation (RI) with respect to other populations. Populations sharing a color or pattern are not differentiated with respect to one another. An asterisk (*) indicates a recently extirpated population, and red arrows show gene flow between otherwise largely isolated populations (omitted from the right panel for clarity). The scales are intentionally dimensionless, as similar processes might be observed across a range of spatial and temporal resolutions.

of speciation. Moreover, feedbacks across these controls are both possible and expected. For example, ecological differentiation between populations can lead to increased reproductive isolation, as predicted under models of ecological speciation (Schluter 2009). Or, ecological differentiation can facilitate range expansion and sympatry between allopatric populations (Price et al. 2014), improving their chances for persistence.

A metapopulation perspective provides a spatially explicit framework for examining the demographic controls on speciation: population isolation and persistence (Figure 2). Population isolation is a potentially important speciation control because the frequency of isolate formation might determine the number of opportunities for speciation to progress. In a metapopulation framework, population isolation is the formation of new populations through either colonization of a new area or the formation of a discontinuity within a widespread population. A reduction in movement or migration of individuals between populations, relative to movement within populations, is a key feature of isolation in a metapopulation framework. Metapopulation structure could reflect either gaps in population density or reduced movement between groups. Movement across the landscape could both prevent isolation and lead to colonization of new, isolated populations.

Population persistence is relevant to speciation because reproductive isolation and ecological differentiation are not instantaneous. After initial isolation, thousands or millions of years might be
required before populations are regarded as distinct species (Futuyma 1987, Levin 1995, Gavrilets 2003, Dynesius & Jansson 2014). In a metapopulation framework, persistence involves both the survival of populations distributed across space and the maintenance of discontinuity between them. Differentiated populations, therefore, can fail to persist either because all members of the population die out (e.g., local population extinction) or due to increased introgression that leads to a merger between populations.

Reproductive isolation or ecological differentiation might evolve between any populations or sets of populations in a metapopulation framework (Figure 2). Migration across the landscape might impede reproductive isolation and ecological differentiation, or it might lead to new local combinations of traits or alleles that accelerate one or both processes. Even populations that do not persist might contribute important traits or alleles to those populations that ultimately form new species diversity. In a metapopulation framework, speciation is not a simple bifurcating process but could progress simultaneously across a suite of coexisting, variously connected, and potentially interacting populations. In our characterization of speciation as a metapopulation process, we are deliberately agnostic with respect to the spatial scale of the dynamics. Demographic processes might occur over small spatial scales as envisioned in traditional metapopulation studies, or they might play out over entire continents as in phylogeographic and biogeographic research; either of these scales might be relevant for speciation. In sum, a metapopulation perspective injects spatial and demographic processes into thinking about the speciation controls.

3. RELEVANCE OF DEMOGRAPHIC PROCESSES IN THE MODERN PARADIGM OF SPECIATION

Models of speciation have focused primarily on two areas of inquiry (reviewed in Turelli et al. 2001, Kirkpatrick & Ravigné 2002, Gavrilets 2003). First, how does disruptive selection lead to the evolution of prezygotic and postzygotic reproductive isolation between populations? Second, how do negative epistatic genetic interactions (e.g., genetic incompatibilities) between populations evolve? At their core, both lines of inquiry focus on the evolution of reproductive isolation, reflecting the disproportionate attention paid to reproductive isolation in speciation studies (Figure 1). These questions have been extended to explore the likelihood that reproductive isolation evolves under different scenarios, such as varying levels of gene flow or habitat differentiation between populations. For example, numerous models explore how the geographic mode of speciation (e.g., allopatric, parapatric, or sympatric) affects the evolution of reproductive isolation (Kondrashov & Kondrashov 1999, Gavrilets 2000). Other models contrast the likelihood of evolving reproductive isolation when populations are in different environments (e.g., ecological speciation; see Schluter 2009) versus similar environments (e.g., nonecological speciation or mutation-order speciation; Mani & Clarke 1990, Nosil & Flaxman 2010). These models are now being extended to describe the dynamics of genome divergence through speciation (i.e., Wu 2001), but the focus remains on understanding the evolution of reproductive isolation.

Demographic processes typically enter this speciation paradigm in a restricted fashion. Population isolation in most models is quite simple. Between any two populations, the models assume an initial population isolation event, and the probability of that initial isolation event is not considered. Further, these models do not acknowledge persistence, other than to allow for the possibility of population fusion due to incomplete reproductive isolation. For example, population size is key to population persistence because smaller populations experience more demographic stochasticity (Lande 1993), increasing the likelihood of extinction. Most mathematical and analytical models include population size (as a fixed parameter) because it determines the rate at which most substitutions are fixed and the time to reciprocal monophyly (Gavrilets 2000, Orr & Turelli 2001).
However, the dynamics of population size are largely ignored; population size is primarily considered only as it affects rates of genetic or trait divergence. These simplifications and assumptions are necessary to generate analytically tractable models. But, as described above, capturing the demographic processes of isolation and persistence may be key to further progress in speciation research.

4. SPECIATION MODELS THAT INCORPORATE POPULATION ISOLATION AND PERSISTENCE

A subset of analytical work, simulation studies, and empirical applications in speciation research does address the role of population isolation or persistence in the speciation process. The list below is not exhaustive, but we believe it covers the best developed and most relevant classes of models.

4.1. Multi-Population Models of Reproductive Isolation

Multi-population models have a long history in evolutionary biology and provide an analytical framework for our verbal models of population isolation. For example, multiple, semi-connected populations were a key element of the influential shifting balance theory of Wright (1931, 1932). Since then, a few models of speciation have combined explicit mechanisms of the accumulation of reproductive isolation with multi-population dynamics of isolation and persistence. Analytical and simulation results from these models have found that population subdivision has either a positive effect (Wright 1940), little impact, or an inverse effect on speciation probability (Orr & Orr 1996), depending on the level of migration modeled between populations and the anticipated mechanism of reproductive isolation. Gavrilets et al. (2000) expanded these models significantly to develop a dynamic multi-population model clearly aligned with a metapopulation framework. Here, sizes of subpopulations change through time, expanding or declining to zero. Exploration of this model uncovered a clear positive association between population persistence and the number of species produced. These dynamic, multi-population models are now often explored using individual-based simulations, as discussed in Section 4.4.

4.2. Species Selection on Traits Related to Demography

Although Mayr (1963) discussed the importance of demographic factors in speciation, the topic received little formal attention in speciation models prior to Gavrilets et al. (2000). However, paleobiological work in the 1970s and 1980s gave significant attention to the importance of demographic factors in speciation. Much of this work developed from the general theme of species selection and emphasized the role that heritable and potentially emergent species-level traits might play in shaping differences in rates of species formation and extinction among clades (Jablonski 2008). Differing larval dispersal modes across marine invertebrates were key in developing these ideas. Clades with sessile, nonplanktonic larvae were thought to be predisposed to population isolation (Hansen 1983, Jablonski & Lutz 1983). Because these traits are heritable at the species level, species selection—through faster speciation rates—would be predicted to favor lineages with traits that predisposed such lineages to isolate formation (Arnold & Fristrup 1982). At the same time, others recognized that the same traits favoring the formation of population isolates might also lead to greater rates of population extinction, such that some traits might have correlated effects on population isolation and extirpation (Stanley 1986). These ideas are now commonly explored in empirical data sets using tests of trait-dependent diversification, discussed in detail in Section 5.
4.3. Island Biogeographical Models

Island systems have a clear spatial context analogous to that of metapopulations; thus, it may not be surprising that early explorations of the role of population isolation and persistence in speciation came from modeling island systems. The theory of island biogeography, for example, formalized the species richness of islands in relation to their size and distance from a source of new diversity (MacArthur & Wilson 1967). Although accounting for processes of colonization, isolation, and persistence, the model in its early formulations did not extrapolate to evolutionary timescales and largely ignored speciation. Subsequent efforts noted that observed biogeographic patterns were better described by models that included in situ speciation as an additional source of new species in island systems (Heaney 1986). In general, these studies have found that speciation is more likely to occur on larger islands (Heaney 1986, Coyne & Price 2000, Losos & Schluter 2000) because population isolates—whether due to geographic barriers, different habitats, or changing ranges through time—are more likely to form on bigger islands. If so, we might further predict that species with lower dispersal would be able to speciate on smaller islands than species with higher dispersal, a prediction borne out by comparative analysis across global island systems (Kisel & Barraclough 2010). These results have catalyzed several expansions to the basic theory of island biogeography that allow island species richness to increase through time as a result of colonization, anagenesis, and within-island speciation (e.g., Valente et al. 2015).

Persistence is also an important component of island biogeographical models. MacArthur & Wilson (1967) predicted, in addition to higher rates of persistence on larger islands, greater persistence after colonization in species with larger population sizes and “fast” life histories. Such species can recover more quickly from population declines and are less prone to stochastic demographic events, predictions supported by studies of modern extinction risk in several taxonomic groups (e.g., Cooper et al. 2008).

Ecological neutral theory (Hubbell 2001) makes perhaps the most explicit connection between speciation and persistence in an island biogeographical framework. Neutral models study community ecological dynamics under an assumed model of speciation (typically, point mutation) and, as such, tell us little about the process of speciation per se (but see Rosindell & Phillimore 2011). However, the stochastic population dynamics that emerge from such models have provided key insights into the demographic processes that facilitate the persistence of new species. In neutral models, extinction is an emergent property of the system that results from demographic stochasticity; it does not require a specific extinction rate as specified by highly phenomenological models of diversification (e.g., the birth-death process; see Nee 2006).

4.4. Simulation-Based Models of Diversification

The emphasis on demographic stochasticity in neutral models, coupled with improved computing power, spurred the development of a new wave of simulation-based, individual- or population-based speciation models that explicitly allow for stochastic fluctuations in population size. These models are increasingly being applied on spatial and temporal scales relevant for speciation. These simulations track the long-term effects of the birth and origination, movement, and death and extinction of individuals and populations across space and time (Gavrilets & Vose 2005, Rosindell & Phillimore 2011, Aguilée et al. 2018, Rangel et al. 2018). They therefore provide explicit links across timescales and across varying environmental conditions, from the fates of individuals and populations to the proliferation of species. The flexibility of these simulations enables them to explicitly model demographic processes and speciation modes.

Across simulation studies, a few patterns emerge, many of which align with predictions from verbal models. Small population sizes lead to extinction (Wang et al. 2013, Aguilée et al. 2018).
If a population is more ecologically similar to the parental species—whether because secondary contact occurs early in divergence or because the rate of divergence is slow, extinction due to competition is likely (McPeek 2008, Aguille et al. 2018). Further, these models showcase the complex role of dispersal in speciation. Increased dispersal can both increase the likelihood of speciation by allowing colonization of new, isolated geographic areas (Rangel et al. 2018) or decrease rates of speciation through the homogenizing effects of gene flow (Gavrilets & Vose 2005, Birand et al. 2012). Initial divergence is often driven by the geography of simulations, so habitat heterogeneity and its interaction with range size can determine speciation rates (Gavrilets & Vose 2005, Aguille et al. 2018). More generally, these simulations often recapitulate broad empirical patterns, such as geographically coarse trends in species richness across South America (Rangel et al. 2018). To the best of our knowledge, only one simulation-based model directly varied the transition rate between population and species (Rangel et al. 2018); variation in this rate (analogous to the rate of reproductive isolation) minimally affected broad-scale diversity patterns. More important was the rate at which populations adapted to changing conditions—low rates of adaptation led to population and species extinction under changing environmental conditions.

4.5. Phenomenological Diversification Models with Demographic Parameters

Phenomenological models allow researchers to study the dynamics of biological processes (e.g., speciation) without explicit modeling of the mechanisms that underlie those dynamics. Birth-death models are among the mostly widely used of such models, and they have been applied to describe variation in speciation rates using both phylogenetic trees and the fossil record (Nee 2006). In their simplest form, birth-death models treat speciation and extinction as instantaneous and random with respect to species’ demography, age, or traits.

Recent work has extended the simple birth-death process to allow the analysis of demographic processes relating to the origin and persistence of incipient species. Weir & Schluter (2007), in their estimation of the relationship between speciation rate and latitude, incorporated a lag time into the simple birth-death model, or an interval between which populations split and are given species designation. This more post hoc approach was later formalized as the protracted speciation model (Rosindell et al. 2010). The protracted speciation model includes separate parameters for the rate at which populations split, the rate at which they convert into species, the rate at which populations collapse, and the rate at which species go extinct. Extensions of this model use phylogenetic data to estimate the length of time required to complete the speciation process (Etienne et al. 2014). The protracted speciation model is a phenomenological model and not a fully realized demographic model because it does not specify a population-level mechanism for extinction. However, the parameters included in these models are consistent with aspects of a demographic model for speciation, in particular, the failure for many populations to persist (Dynesius & Jansson 2014). Thus, this model provides a theoretical framework for describing the empirical disconnect between the rates at which phylogeographic lineages emerge versus those at which new species form (Rosenblum et al. 2012). Further explorations of the protracted speciation model found that varying the transition rate between populations and species has a large effect on macroevolutionary patterns such as species richness (Cutter & Gray 2016). However, phylogenetic diversification models might not be able to distinguish among different demographic scenarios of speciation, which could limit the extent to which inferences about population-level processes can be drawn from phylogenetic data alone (Etienne et al. 2014, Li et al. 2018).

Similar models have been employed outside of the protracted speciation framework. For example, Muir & Hahn (2015) developed a phylogenetic model that explored the role of whole genome duplication in driving broad-scale diversity patterns. This model explicitly accounted for
population processes by accounting for the rate at which population isolates form and then come back into secondary contact. Variance in this transition rate had a larger impact on the speciation rate than the rate at which hybrid incompatibilities evolved, leading the authors to conclude that population dynamic processes may be a more important control on speciation than genetic incompatibilities. Together, the advances made by these phenomenological phylogenetic models have shown the potentially large impact of demographic processes on broad-scale diversity patterns.

5. TESTS OF LINKS BETWEEN DEMOGRAPHIC PROCESSES AND MACROEVOLUTIONARY SPECIATION DYNAMICS

If a given process is an important speciation control, we expect that it should have some impact on macroevolutionary patterns of speciation (Figure 3). Previous authors have advocated for identifying the importance of different sources of reproductive isolation by testing their macroevolutionary impacts (Panhuis et al. 2001, Coyne & Orr 2004). For example, a particular isolating barrier between populations cannot limit speciation if it evolves after complete reproductive isolation or sympatry has already been achieved. Similarly, some of the speciation controls might be primary drivers of speciation, whereas others are secondary. These secondary processes may be important for species differences and therefore merit study even if they are not rate-limiting steps. However, they will not shape the dynamics of speciation and therefore are of limited relevance for understanding the origins of species diversity.

Comparative tests can be used to evaluate the efficacy of potential controls by assessing their correlation with macroevolutionary patterns of speciation rates or species richness. Studies of single species or species complexes have tremendous potential to show how ecology, demography, and genetic change interact to drive speciation in that lineage, but comparative analyses across multiple lineages that differ in their rate of speciation can provide robust tests of the relationship between population-level processes and macroevolutionary dynamics.

Trait-dependent diversification tests are the primary tool for these types of analyses because they can identify statistically independent effects of a focal trait on the rate of speciation in a phylogenetic framework (Mitter et al. 1988, Maddison et al. 2007). Because comparative data on isolation and persistence are difficult to collect, research has instead tested their importance by using proxy traits. For example, dispersal syndrome has been used in trait-dependent diversification studies as a proxy for population structure or isolation. These studies have found mixed results, with many reporting a negative association between traits related to dispersal and speciation rates (Jablonski 1986, Claramunt et al. 2011, Weeks & Claramunt 2014), others little association (Kisel et al. 2012), and some a positive relationship (Phillimore et al. 2006, Krug et al. 2015). Range size has also been widely studied and may reflect dispersal ability, opportunities for population isolation, or likelihood of population persistence. Fossil data generally find that larger-ranged species have lower speciation rates (Jablonski & Roy 2003), consistent with the ideas that high dispersal hinders speciation and that larger-ranged species have longer species durations (Hansen 1983, Jablonski & Hunt 2006). Geographic range size is heritable across speciation events (Jablonski 1987), suggesting that it can contribute to clade-wide trends in species traits. A variety of other traits have been examined (reviewed in Coyne & Orr 2004, Jablonski 2008, Dynesius & Jansson 2014), many of which might covary with demographic controls. These results are consistent with links between demography and speciation rates.

Variables that serve as rough proxies of speciation controls have their limits. As noted above for range size, any particular variable investigated with respect to diversification rates might impact the speciation process through multiple distinct causal mechanisms. In addition, many traits might
simultaneously elevate speciation rates and reduce extinction, which makes identifying the true mechanism of trait-dependent diversification challenging (Coyne & Orr 2004, Jablonski 2008). Confirming the mechanisms responsible for trait diversity associations would benefit both from more explicit causal modeling and from more direct measures of microevolutionary processes. One way forward is to collect and use traits that more directly capture speciation controls, rather than relying on proxy variables that may have a more tenuous connection to population mechanisms. For example, Rabosky & Matute (2013) used experimental and observational data on intrinsic reproductive isolation in *Drosophila* and birds to test the link between rates of reproductive isolation and macroevolutionary speciation rates. They failed to find support for a link.

Figure 3
A hypothetical example of a comparative and integrative speciation research program based on a metapopulation framework. Demographic processes are explored in multiple species exhibiting disparate traits and/or falling in lineages with different speciation dynamics. Estimates of these processes can then be compared with trait variation and speciation rates. The scatterplots show a stronger relationship between speciation rates and measures of the demographic process itself than that of its proxies. Lizard silhouette by David Orr from PhyloPic (<http://www.phylopic.org>) is licensed under CC BY 1.0.

84 Harvey • Singhal • Rabosky
A few studies have assessed variation in speciation rates with respect to measures that might better capture population isolation and persistence. For example, population genetic metrics of population isolation have been found to exhibit either a positive association with speciation rate (Riginos et al. 2014, Harvey et al. 2017a) or no relationship (Kisel et al. 2012, Singhal et al. 2018). We are unaware of existing studies explicitly examining associations between macroevolutionary speciation rates and measures of population persistence across the landscape. Persistence and its importance have been perennially challenging to study (Dynesius & Jansson 2014). Like species extinction, population loss is difficult to detect without a detailed fossil record (but see Section 6). Population persistence is also challenging to study because, in addition to links to speciation rate, it is likely highly correlated with extinction. In fact, the existence of a gray zone during the speciation process (de Queiroz 2007), in which groups are intermediate between separate populations and separate species, implies that population persistence will likely grade into species persistence. Species with traits predisposing them to population persistence, for example, large range size or phenotypic plasticity, are likely to also experience low extinction. This association might obscure the impact of persistence on speciation rates. The profound disconnect between speciation rates as measured at microevolutionary and macroevolutionary scales (Rosenblum et al. 2012, Rabosky 2016) suggests that persistence is important. An obvious source of this discrepancy involves a general failure of incipient species to persist through time. Low persistence would result in faster estimates of speciation rates at shallow timescales relative to deeper (phylogenetic) timescales. Despite the challenges in studying persistence, it therefore may be the control most likely to reveal fruitful insights into speciation dynamics.

6. SOURCES OF DATA ON DEMOGRAPHIC CONTROLS

Diverse sources of data can potentially provide insight into the spatial demographic processes of population isolation and persistence. For isolation, relevant types of data describe the structure of populations and dispersal and include population genetic differentiation, distributional data on clustering of individuals, measurements of gene flow or individual movement, and data on traits that covary with dispersal. For persistence, direct measures may be possible with long-term survey data, or rarely from genetic or fossil data, but indirect measures based on genetic or phenotypic diversity, experimental responses to environmental changes, short-term population dynamics, or trait data may also be useful. Table 1 provides more details about all of these types of data. More direct and therefore more desirable measures appear toward the top of each subcategory in the table, whereas those that are less direct, such as trait data, appear toward the bottom. Some of these data sources (population genetic differentiation, range size and limits) already appear frequently in speciation research. Others (extinction risk, long-term survey data, animal tracking) appear less frequently. Importantly, for comparative study, units of study and scales of comparison must be comparable in their taxonomic, spatial, and especially temporal context. Below, we suggest data sources that may provide comparable demographic information relevant to speciation.

6.1. Population Genetic and Phylogeographic Structure

Genetic data have been a key source of information on demography and population structure since the development of F-statistics and models of isolation by distance (IBD) (Wright 1943). Spatial patterns of turnover in genetic diversity provide information about the metapopulation structure of species. These genetically differentiated populations are often assumed to reflect potential incipient species, which has motivated whole areas of research (e.g., phylogeography; see Avise 2000). Landscape genetics, although typically conducted at smaller scales, also provides insights
<table>
<thead>
<tr>
<th>Measure</th>
<th>Data</th>
<th>Details</th>
<th>Example Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Isolation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population structure</td>
<td>Genetic clusters (preferably with age information)</td>
<td>Genetic breaks reveal isolation; age information permits estimation of the isolation rate through time</td>
<td>Harvey et al. 2017b</td>
</tr>
<tr>
<td></td>
<td>Phylogeographic lineages</td>
<td>Distinct lineages in a tree of populations (e.g., from mitochondrial DNA) and their ages</td>
<td>Harvey et al. 2017a</td>
</tr>
<tr>
<td></td>
<td>Isolation by distance curves</td>
<td>The spatial decay in genetic similarity, preferably scaled by the time over which it has accumulated</td>
<td>Riginos et al. 2014</td>
</tr>
<tr>
<td></td>
<td>Distributional data/ecological niche models (ENMs)</td>
<td>Distributional gaps or areas of reduced density can reveal isolation; longitudinal data can reveal age of isolation events (e.g., in introduced species)</td>
<td>Carnaval et al. 2009</td>
</tr>
<tr>
<td><strong>Dispersal</strong></td>
<td>Gene flow estimates</td>
<td>Estimates of gene flow from genetic data estimated using a multi-population demographic model or summary statistics</td>
<td>Harvey et al. 2017b</td>
</tr>
<tr>
<td></td>
<td>Dispersal kernel, genetic</td>
<td>Use of individual-based genetic approaches to infer the distribution of dispersal events (e.g., genetic parentage analysis)</td>
<td>D’Aloia et al. 2015</td>
</tr>
<tr>
<td></td>
<td>Dispersal kernel, tracking</td>
<td>Individual-based data (e.g., mark–recapture, radio and satellite tracking) on movement within or between populations</td>
<td>Kays et al. 2015</td>
</tr>
<tr>
<td></td>
<td>Dispersal mode (trait data)</td>
<td>Use of intrinsic species proxy traits (seed size, flying versus terrestrial, animal versus wind pollinated, pelagic larval duration/larval mode) to broadly describe dispersal extent</td>
<td>Claramunt et al. 2011</td>
</tr>
<tr>
<td><strong>Persistence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual persistence</td>
<td>Long-term population survey data</td>
<td>Direct measures or estimates of changes in population size through time from longitudinal survey data</td>
<td>Sanderson et al. 2006</td>
</tr>
<tr>
<td></td>
<td>Long-term range shift data</td>
<td>Longitudinal monitoring efforts or observation databases stratified by time</td>
<td>Tingley et al. 2009</td>
</tr>
<tr>
<td></td>
<td>Genetic diversity</td>
<td>Signatures of merged populations may be identifiable in species (e.g., when multiple sets of alleles are in linkage disequilibrium)</td>
<td>Villanea &amp; Schraiber 2019</td>
</tr>
<tr>
<td></td>
<td>Fossil data</td>
<td>For species with a good fossil record, the range and population size can be tracked through time</td>
<td>Smits 2015</td>
</tr>
<tr>
<td></td>
<td>Retrodiction/prediction of species distributions using ENMs</td>
<td>Using historical and projected future climate data to determine how species distributions change through time</td>
<td>Carnaval et al. 2009</td>
</tr>
<tr>
<td>Genetic load</td>
<td>Fitness impact of deleterious alleles</td>
<td>Measurement of homozygous-derived mutations and missense and loss-of-function variants; functional predictions of mutation impact</td>
<td>van der Valk et al. 2019</td>
</tr>
<tr>
<td>Plasticity</td>
<td>Reaction norms</td>
<td>Response to adverse environmental conditions analogous to those that might be encountered over evolutionary time</td>
<td>Jiguet et al. 2006</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Dataa</th>
<th>Details</th>
<th>Example Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptive potential</td>
<td>Genetic diversity</td>
<td>In particular, diversity at potentially important coding loci (e.g., major histocompatibility complex genes)</td>
<td>Saccheri et al. 1998</td>
</tr>
<tr>
<td></td>
<td>Phenotypic diversity</td>
<td>Diversity in functionally important traits that might facilitate adaptation to new environments</td>
<td>Bolnick et al. 2007</td>
</tr>
<tr>
<td>Population size and spread</td>
<td>Survey data on population size</td>
<td>Counts or estimates of total population size</td>
<td>Fink et al. 2018</td>
</tr>
<tr>
<td></td>
<td>Genetic estimates of population size, size fluctuation</td>
<td>Allele frequency distributions or coalescent events can track population size history</td>
<td>Lorenzen et al. 2011</td>
</tr>
<tr>
<td></td>
<td>Distributional data</td>
<td>Might include range size information and data on spatial dispersion of individuals</td>
<td>Carraro et al. 2018</td>
</tr>
<tr>
<td></td>
<td>Stochastic population models</td>
<td>Using species-specific parameters of life history to inform statistical models of population growth and extinction risk (i.e., population viability analysis)</td>
<td>Keith et al. 2008</td>
</tr>
<tr>
<td></td>
<td>Life history traits (fecundity, growth rate, recruitment)</td>
<td>These are traits with anticipated impacts on population viability and growth</td>
<td>Cooper et al. 2008</td>
</tr>
</tbody>
</table>

aWithin each category, ordered from most promising (direct measures with readily available data) to least promising.
bThese citations are not meant to be exhaustive but rather to provide examples relevant for comparative research.

into the origins and maintenance of population structure and especially the role of landscape heterogeneity in impeding or facilitating gene flow (Manel et al. 2003).

Genetic differentiation ranges from discrete breaks between clusters to continuous differentiation across space (IBD) (Figure 4). Discrete breaks have clear implications for speciation—in geographic speciation models, such breaks are a necessary part of the speciation process (Mayr 1963). Continuous differentiation across landscapes—whether due to environmental gradients or neutral or demographic processes (e.g., Seeholzer & Brumfield 2018)—can also lead to speciation, as documented in cases of speciation with gene flow (Schluter 2009). Future studies might evaluate whether IBD or the presence of discrete genetic breaks better predicts speciation dynamics within a lineage.

Population genetic structure has often been inferred using single-locus or multi-locus phylogeographic studies. Substitution rates or fossil or geological calibrations can be used to identify the timescale over which isolation has occurred. High-throughput sequencing and new methods for analyzing genomic data (e.g., Petkova et al. 2016) have increased our ability to resolve genetic breaks across species and to determine patterns of migration across these breaks. Range-wide sampling is important for detecting structure, however, even with genomic data. High-throughput workflows will ultimately also facilitate large comparative studies of population genetic structure from many species.

How population genetic structure forms within a lineage can impact how it should be modeled in studies of speciation. Is population genetic structure at equilibrium within a species, or are patterns of structure within species actively changing as a result of ongoing processes of population expansion and differentiation? The former might be the case if structure is inherited from an ancestral species, whereas the latter might be a better model if a species began as a small or
Panmictic population and then expanded and differentiated through time. These alternative perspectives may be responsible for different results in studies comparing population structure with speciation rates (Harvey et al. 2017a, Singhal et al. 2018). A related question is whether variation in structure reliably reflects the traits of a species, such as dispersal, selection, and spatial variation in abundance, or whether it primarily reflects idiosyncratic historical factors, such as landscape change through time. If the latter, then population genetic structure may provide limited insight into the long-term demographic attributes of a species.

### 6.2. Movement Ecology (Migration and Dispersal)

Data on the movement of organisms can inform our understanding of dispersal across metapopulations. Reduced dispersal is indicative of population isolation in a metapopulation framework. Elevated dispersal, however, might reflect traits that predispose a species to colonize new areas or persist through time. Studying dispersal on scales relevant to speciation has been challenging. However, new tools are facilitating the estimation of dispersal from more individuals and at unprecedented spatial scales.

Animal tracking has seen dramatic improvements in recent decades. In particular, the tracking of small animals is now possible due to the development of small tags and a more powerful, space-based, animal-tracking infrastructure (Kays et al. 2015). These improvements permit the tracking of animals over large distances, and battery improvements make it possible to obtain tracking...
information throughout the entire lifetimes of organisms. Other technologies can be used to obtain large data sets for particular groups, for example, Doppler radar for movements of birds and other aerial organisms (Horton et al. 2016). These tools might permit the estimation of realized patterns of dispersal across the distributions of many species.

Experimental studies may also provide species-level estimates of dispersal abilities that are more precise or relevant for speciation than proxies based on morphological or natural history traits. For example, flight trials in birds have been used to evaluate both physical capacity and psychological propensity to cross habitat barriers (Moore et al. 2008). Similar studies could be used to generate broadly comparable data sets across taxa that vary substantially in their rate of speciation to maximize statistical power for trait-dependent diversification analyses. Experiments and observational studies in marine organisms are leading to more and better estimates of pelagic larval duration (Shanks 2009), which is thought to correlate with oceanic dispersal. In cases for which detailed dispersal patterns are not available from tracking data, improved dispersal trait information may be useful.

It remains unclear which measurements of dispersal are most relevant for speciation. First, dispersal does not always lead to gene flow. Second, rare, long-distance dispersal events are likely to have the largest impact on population differentiation but are also the most challenging to measure. Further technological developments, however, may facilitate tracking of enough organisms for long enough intervals to identify overall dispersal trends and capture rare events.

6.3. Range Limit Research and Niche Modeling

Range size and shape have critical impacts on both population isolation and probability of persistence (Mayr 1963, Gaston 2003). Contemporary data reflect the metapopulation structure of species, whereas longitudinal data can reveal metapopulation dynamics through time. Niche modeling is a primary tool for mapping species distributions based on current, past, or future abiotic and biotic environments across space (Carnaval et al. 2009). Correlative niche models are built using environmental data from current distributions, and they generally assume niche conservatism. Thus, they cannot be used to estimate occurrence in environments without contemporary analogs, nor does potential for plastic or evolutionary responses exist (Pacifici et al. 2015). Mechanistic niche models permit the incorporation of information on niche evolution or experimental data on physiological responses (e.g., Kearney & Porter 2009, Cunningham et al. 2016) but require detailed data lacking from many species. By dissecting how dispersal and fitness limit a species’ range, experimental studies of range limits allow even more mechanistic perspectives of what determines population persistence at range margins (Sheth & Angert 2018). Robust comparative data sets on isolation and persistence may be obtainable using a combination of correlative and mechanistic niche modeling approaches.

6.4. Experimental and Observational Studies of Persistence

Although population persistence is challenging to measure, experimental and observational studies are increasingly conducted on spatial and temporal scales that might capture this phenomenon. When collected from enough species, these estimates may be useful to address the persistence control on speciation in a comparative framework. Long-term observational studies can take advantage of recurring field surveys to quantify persistence (e.g., Ehrlich & Murphy 1987, Tingley et al. 2009, Laurance et al. 2011). Long-term data, of course, are challenging to obtain, and many empirical studies do not directly estimate persistence (Burgess et al. 2014). A limited number of studies have directly measured population persistence in a metapopulation framework.
(e.g., Cain & Cook 1989, Harrison 1991), but their connection to broader patterns of speciation remains largely unexplored (Allmon & Sampson 2016). The recruitment of citizen scientists (e.g., Fink et al. 2018) and new remote sensing tools for measuring populations (e.g., Lynch et al. 2012), however, are streamlining the collection of long-term survey data. Estimates of population persistence from many species will be critical for linking persistence to both trait variation and speciation probability. Importantly, many of these longitudinal studies were designed to measure the effects of anthropogenic change on species distributions. Before using these data, we should evaluate if and how they are different from species responses to the types of environmental changes encountered over evolutionary timescales.

6.5. Conservation Risk

Traits are used as predictors of extinction risk in conservation biology and may provide a rough proxy of variation in population persistence for use in comparative studies. Traits such as home range size, dispersal ability, and growth rates have been found to effectively predict population recovery or resilience through time (Cooper et al. 2008, Montero-Serra et al. 2018). Biogeographic setting may be an important extrinsic trait to consider, with species in less stable environments potentially exhibiting lower population persistence. This information has been used to determine exposure scores for species risk in the face of climate change (e.g., Gardali et al. 2012). Integrative or holistic metrics of extinction risk might also be a useful proxy. The IUCN red list categories, for example, incorporate a combination of survey results, trait information, and the results of quantitative modeling such as population viability analysis (IUCN 2017).

The links between traits related to extinction risk and population persistence may not always be straightforward. Rarity, for example, is often associated with a high risk of extinction, but chronically rare species may also have low mutation loads (Robinson et al. 2018) or other traits that make them resistant to extinction (Rabinowitz 1981). In addition, persistence can be achieved through a variety of avenues, from niche tracking to plasticity or adaptation, each of which might be useful to organisms with different life histories. Because persistence can be achieved through multiple pathways, measuring persistence by any one trait may be difficult across geographic regions or taxonomic groups (Hatfield et al. 2018). Finally, persistence may vary among populations within a species (Harrison 1991), and the shape of this distribution may be more relevant than the trait-based approximation of the species-wide average. Trait information may be useful primarily in cases for which long-term survey data or niche modeling data pertaining to persistence are lacking or incomplete.

7. CONCLUSION AND PROSPECTUS

In summary, we view a spatial perspective based on metapopulation ideas as a natural conceptual framework that incorporates the controls of population isolation and persistence into speciation research. The advancement of this framework will require additional work in the three primary areas covered above:

1. Further development and adoption of theoretical speciation models containing isolation and persistence (Section 4).

2. Use of comparative data sets to test potential links between demography and macroevolutionary speciation dynamics (Section 5).

3. Gathering and incorporating new sources of data on demography into both single-species and comparative speciation studies (Section 6).
We emphasize that this article represents a current snapshot across a set of rapidly advancing research areas. New studies continually present diverse data from additional species and regions that might be relevant for understanding the role of demographic processes in speciation and for linking micro- and macroevolution. We look forward to what is surely a coming integration of these data sources and ideas.

DISCLOSURE STATEMENT
The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS
For discussion, we thank Robb Brumfield, Brian Tilston Smith, Glenn Seeholzer, the Rabosky laboratory, Rachael Herman, Doug Futuyma, Dan Simberloff, Liz Derryberry, Sara Lipshutz, Niki Patel, Michael Whitlock, Jay McEntee, and Molly Schumer. Funding was provided by National Science Foundation grants DBI-1523893 (to M.G.H.), DBI-1519732 (to S.S.), and DEB-1754398 (to D.L.R.); a California State University–Dominguez Hills Research, Scholarly and Creative Activity grant (to S.S.); and a fellowship from the David and Lucile Packard Foundation (to D.L.R.).

LITERATURE CITED
Bolnick DI, Svanbäck R, Araújo MS, Persson L. 2007. Comparative support for the niche variation hypothesis that more generalized populations are also more heterogeneous. PNAS 104:10075–79


Harvey MG, Seeholzer GF, Smith BT, Rabosky DL, Cuervo AM, Brumfield RT. 2017b. Positive association between population genetic differentiation and speciation rates in New World birds. *PNAS* 114:6328–33


Kisel Y, Barraclough TG. 2010. Speciation has a spatial scale that depends on levels of gene flow. *Am. Nat.* 175:316–34


Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16:97


