What Are Species Pools and When Are They Important?

Howard V. Cornell and Susan P. Harrison

Department of Environmental Science and Policy, University of California, Davis, California 95616; email: hvcornell@ucdavis.edu, spharrison@ucdavis.edu

Abstract

A regional species pool comprises all species available to colonize a focal site. The roots of the concept are imbedded in island biogeography theory, supply-side ecology, and early propagule addition experiments. The pool concept allows ecologists to examine large-scale effects—including geographic area, evolutionary age, and immigration and diversification—on the diversity, composition, and phylogenetic structure of local communities. Both theory and evidence show that pool influences are greatest when local communities are not strongly and predictably structured by species interactions (e.g., under frequent disturbance or if many species are rare). Practical and conceptual issues to consider when delineating species pools include choosing an appropriate spatial scale, whether to account for environmental filtering, whether to include the species within a fixed geographic area versus those whose geographic ranges overlap with a site, or whether to use databases or geographic data sources. Each issue is discussed in the context of 63 studies using the species pool approach. We conclude that the species pool concept has contributed greatly to our understanding of community dynamics by bridging the gap between large and small spatial scales. Future studies must compare pool characteristics with community structure across multiple regions for a more complete understanding of community assembly.

Keywords

community assembly, environmental filtering, local diversity, phylogeny, regional diversity, species interactions
INTRODUCTION

Ecologists have long recognized that the composition and richness of local communities are influenced by biogeographic and evolutionary processes acting at large spatial and temporal scales (Terborgh & Faaborg 1980, Cornell 1985, Ricklefs 1987). Considering the regional species pool can make the influences of these processes tractable in ecological analyses (Harrison & Cornell 2008). Broadly, a regional species pool is defined as the set of all species available to colonize a focal site (Srivastava 1999). Assessing variation in the size and composition of regional species pools and determining their relationship to the composition of local communities—including their richness, phylogenetic/functional trait structure, and spatial variation (beta diversity)—is a way to include the potential influence of large-scale processes into analyses of community assembly. The goal of this review is to discuss the conceptual and practical issues that should be considered when defining regional species pools and when exploring their influence on community structure.

The species pool concept has been implicitly criticized by some ecologists (e.g., Rosenzweig & Ziv 1999) and evolutionists (e.g., Price et al. 2011) from the viewpoint that regional biotas are under the strong control of locally acting processes, such as competition. Others have argued that local species composition is under such strong regional control that the local community may not even be a meaningful unit of study (e.g., Ricklefs 2008). Metacommunity theory, although not yet well developed empirically, depicts the local and regional scales as dynamically interdependent through frequent dispersal among localities (Leibold et al. 2004). Hence, our first goal is to describe when the regional species pool as opposed to local species interactions best explains variation in local community structure. We then turn to more practical issues of how to define and measure regional species pools.

We note that regional species pools are also an integral part of null model analyses aimed at detecting the influences of local processes—species interactions or environmental filters (environmental mismatches that prevent species from colonizing the site)—on the species, phylogenetic, or trait structure of local communities. In these analyses, regional processes are not the main focus because there is typically only one regional pool from which multiple local communities are drawn (although the pool may be spatially or taxonomically varied as a form of sensitivity analysis). There is already a large literature on such analyses (reviewed by Carstensen et al. 2013), so we do not treat the topic here. The scope of our review treats the subset of null model analyses that consider how variation in regional species pools affects variation within and among local communities (e.g., Kraft et al. 2011; Lessard et al. 2012a,b; Anacker & Harrison 2012).

WHERE DID THE REGIONAL SPECIES POOL CONCEPT ORIGINATE?

Empirical Origins

Appreciation of large-scale influences can be traced back at least to the first decade of the past century when Spalding (1909) showed that seed availability played an important role in structuring desert plant communities. Many subsequent plant ecologists emphasized the role of seed dispersal in addition to the filtering effects of the local environment and competition (e.g., Clements 1916, Gleason 1927, Harper 1977). Research on tropical trees pointed to the importance of recruitment limitation and demographic stochasticity to community assembly and was particularly influential in challenging the locally determined, equilibrial view of community diversity (Hubbell & Foster 1986). As the “species pool hypothesis” (Pärtel et al. 1996) to explain plant community diversity gained momentum, ecologists began to consider the species found in the area surrounding the community, the ecological characteristics of these species, and the likelihood that each species
could disperse to the locality (Mueller-Dombois & Ellenberg 1974, Pärtel et al. 1996). Seed addition experiments frequently resulted in increased species diversity in local plant communities, indicating unfilled niches and the potential for species enrichment from the regional pool (reviewed by Myers & Harms 2009).

In fisheries biology, fluctuations in yields were traditionally attributed to variations in oceanographic conditions affecting population recruitment (Hjort 1926, Ricker 1954). In the mid-1980s, the implications of open recruitment for the structure and dynamics of intertidal, pelagic, and benthic communities came strongly into focus. “Supply-side ecology” proposed that propagule availability was at least as important to community assembly as postrecruitment interactions (Gaines & Roughgarden 1985).

**Theoretical Origins**

An early challenge to the pre-eminence of local processes in community assembly was island biogeography theory, which proposed that island species diversity depended on immigration from a mainland species pool. In tests of this theory (reviewed by Schoener 2010), it was shown that island isolation had negative effects on immigration rates and thus local diversity, providing considerable support for the general idea of species pool effects on islands and fragmented habitats. Such effects were reinforced by experimental studies showing that community structure was often more idiosyncratic than predicted by classic niche assembly (e.g., Sutherland 1974), suggesting that community membership was contingent on dispersal from the species pool. Neutral theory (Hubbell 2001), a direct descendant of island biogeography theory, depicts local communities as stochastic samples from a regional pool in which the probability of dispersal to the community is related to the distance and abundance of a species in the pool rather than its identity. The pool is in turn governed by the vagaries of biogeographic history, expressed as random speciation and extinction events. In general, however, neutral theory focuses more on the absence of niche differences than on the detailed understanding of regional influences on local communities.

Metacommunity theory (Leibold et al. 2004) comprises a broad class of models (including neutral theory) in which multiple local communities are linked by dispersal. In contrast to island biogeography theory, neutral theory, and the species pool hypothesis, many metacommunity models assume that species characteristics matter and that regional processes (i.e., dispersal among localities) operate on a similar timescale to that of local ones (i.e., resource availability, habitat selection, species interactions); thus, all these processes interact to determine which species can persist regionally (i.e., in the set of patches) as well as locally (i.e., within each patch). Support for these models comes mainly from microcosms and mesocosms in which frequent dispersal can be created experimentally (e.g., Holyoak & Lawler 1996, Cadotte 2006). In more natural systems, the assumption of similar timescales for local and regional processes has been difficult to test.

**The Local-Regional Diversity Relationship**

Direct tests of the relationship between local richness and the richness of the regional species pool based on observational data were first proposed by Terborgh & Faaborg (1980), conducted by Terborgh & Faaborg (1980) and Cornell (1985), advocated and further conducted by Ricklefs (1987, 2000), and reviewed by Lawton (1999) and Harrison & Cornell (2008). In the classic test, a linear local-regional diversity relationship was interpreted to mean that local communities were open to regional processes, whereas a curvilinear relationship meant diversity was determined locally and approached saturation. The statistics and interpretation of this test have been questioned, specifically because curvilinear relationships tend to arise if local richness is underestimated and the regional pool is overestimated and because a curvilinear relationship is not sufficient to infer...
Community structure is governed by a continuum of processes including species pool genesis, dispersal to and from the pool, environmental filtering, and niche occupancy (Leibold et al. 2004, Harrison & Cornell 2008, Myers & Harms 2009, Chase & Myers 2011; Figure 1). A brief survey of current ecological theory can indicate when the regional species pool is likely to contribute to community structure.

**Neutral Theory**

Communities are open to regional influences as long as niche space is not completely filled. As demonstrated by neutral theory, the extreme case of community openness is when all species have identical niches and identical competitive abilities such that new immigrants are just as likely to recruit successfully into a community as individuals from resident populations. In this paradigm of
pure dispersal assembly, within-community (alpha) and among-community (beta) diversity arise only from chance local extirpations and dispersal events (Hubbell 2001). However, even in the more realistic case of niche assembly, where community variation arises at least in part from niche differences and environmental heterogeneity (e.g., Kraft et al. 2011, Shipley et al. 2012), communities may still be open to regional influences. For example, local sites varying in their climate, soil fertility, or disturbance regimes may be colonized from a regional pool containing different numbers of species well adapted to each set of site conditions. Under this scenario, sites with high soil fertility may have higher diversity not because they have more niche space but simply because more species in the pool are adapted to fertile than nonfertile soils. Such an interaction between local environmental heterogeneity and the composition of the regional species pool may give rise to variation in local richness along environmental gradients (e.g., Pärtel et al. 1996, Myers & Harms 2009, Belmaker & Jetz 2012).

Unsaturated Versus Saturated Niche Space

Disturbance contributes importantly to the openness of communities by making unused resources available (e.g., Heino et al. 2003, Tilman 2004). The prediction that species pool effects are more important in frequently disturbed localities has been supported by many studies (e.g., Collins et al. 2002; Belote et al. 2009; Myers & Harms 2009, 2011). Pool effects should also be stronger in communities dominated by rare or transient species, which tend to compete less consistently than common and widespread species (Hubbell & Foster 1986). This prediction has been upheld in studies of coral reef fishes (Belmaker 2009) and North American breeding birds (White & Hurlbert 2010). Finally, pool effects are likely to be more important if some species have limited dispersal abilities or there are dispersal barriers among localities within the region (e.g., Roslin 2001). Such limited dispersal can result in unused resources so that species that manage to reach a locality are likely to colonize successfully.

In contrast, when resources are consistently and fully used by resident species, the community is saturated, colonizers are unlikely to succeed, and the regional species pool has minimal impact on community structure. Dispersal effectively transports community members to localities such that vacant niche space is rare. Under these conditions, coexistence is possible only through fitness trade-offs at the local scale (Siepielski & McPeek 2010). Species abundances and niche breadths are expected to decline with increasing diversity and to be negatively correlated through time among coexisting species (e.g., Ricklefs 2012). Propagule addition has only transient effects (e.g., Harrison et al. 2010), and community composition remains relatively constant in the absence of environmental change (Clark & McLachlan 2003).

Regional and Local Processes Combined

Regional and local processes are combined in priority-effect models, in which different sequences of colonization lead to different communities that may be uninvasible at equilibrium (e.g., Sutherland 1974). Local species composition is partly determined by chance arrival from the regional pool, but diversity is still limited by niche space and fitness trade-offs (Chase & Myers 2011). The more frequently equilibrium is disrupted by disturbance, the more that community structure will tend to mirror the regional pool, but some local combinations of species will be more persistent than others (Weiher et al. 2011). Priority effects also increase beta diversity and thus retain species in the pool that would otherwise be excluded in all localities if community composition were not dependent on colonization sequence.

Regional and local processes are also combined in models that consider limited dispersal among localities and spatially explicit competition where species are most likely to compete with close
neighbors. These models predict stronger pool effects when dispersal is low, because microsites are less likely to be taken by the best competitor (e.g., Hurtt & Pacala 1995). Local diversity and abundance do not converge to their deterministic equilibria even when competition and niche differentiation are strong. The effects of dispersal limitation are expected to be greatest for large pools and diverse communities, because a high proportion of the constituent species will be rare (Chase & Myers 2011).

Dispersal to localities also interacts with local processes in stochastic niche theory (Tilman 2004). Competition with residents prevents many colonists from establishing, but successful establishment occurs when propagule arrival coincides with a disturbance that makes surplus resources available. Although the probability of establishment is enhanced by niche differences between colonizers and residents, there is no fixed limit to species similarity, and chance dispersal and disturbance events can have long-lasting effects on community structure.

Because it is not always clear in the literature, we reiterate that openness to regional species pools is not identical to neutrality or pure dispersal assembly. In turn, the regional species pool hypothesis is not contradicted by the existence of niche differences (e.g., Shipley et al. 2012), by environmental filtering of species into local sites (e.g., Myers & Harms 2011), or by relationships between environmental gradients and community composition (e.g., Myers et al. 2012). The key question is not whether niches exist but whether niche space is generally filled by resident species or whether it is frequently made available by disturbances, by weak or variable interactions, or by limited dispersal among localities.

Diversity Gradients

Diversity-productivity gradients exemplify a community pattern that is consistent with either local or regional explanations. Productivity may determine the availability of local niche space (Hurlbert & Jetz 2010), in which case less-productive sites show saturation (full resource use, low invasibility) at smaller numbers of species than more productive sites (e.g., Price & Kirkpatrick 2009). Alternatively, local productivity may determine how many species from the regional pool are capable of colonizing a locality (e.g., Chase & Myers 2011, Belmaker & Jetz 2012). In this case, local diversity is predicted to increase with productivity within generally productive regions, where the regional pool will contain few species tolerant of low productivity. In contrast, local diversity will decrease with increasing productivity within generally unproductive regions, where most species in the pool are adapted to low productivity (Pärtel et al. 1996; Belmaker & Jetz 2012; Lessard et al. 2012a,b).

Geographical gradients in the processes governing diversity are likely to influence the openness of local communities to regional influences. Results from experimental (e.g., Freestone & Osman 2011) and observational (e.g., Witman et al. 2004) analyses suggest that the proportion of regional species able to coexist in localities increases with latitude, consistent with local communities being controlled by stronger biotic interactions in the tropics (Schemske et al. 2009) or lower thermal tolerances of individuals in aseasonal environments (Ghalambor et al. 2006). However, tropical species are likely to be at lower densities than temperate species and may go stochastically extinct from small areas, generating a similar pattern. Moreover, other studies have concluded that tropical forests are more strongly shaped by neutral processes than temperate forests, as demonstrated by stronger effects of distance than environment on beta diversity (e.g., Myers et al. 2012). Tropical forests support more rare, transient, dispersal-limited species than temperate ones (Hubbell 2001, Myers et al. 2012), but very high levels of diversity may make it difficult to distinguish the roles of niche versus dispersal assembly processes (Chisholm & Pacala 2010).
WHY DO REGIONAL SPECIES POOLS VARY?

Regional species pools are ultimately shaped by speciation, immigration, and extinction (Figure 1). These processes lead to the accumulation of species over a time span that may be defined by the origin of its water bodies, inshore habitats, or land surfaces (e.g., the emergence of oceanic islands) or by major climatic and biotic changes (e.g., glacial retreat). The effects of these processes can be represented by the general expression

\[ A = \int_{t=0}^{\text{age}} [\bar{\lambda}(t) + \bar{\delta}(t) - \bar{\mu}(t)] \, dt, \]

where \( A \) is the rate of accumulation of species in the region and \( \bar{\lambda}(t) \), \( \bar{\delta}(t) \), and \( \bar{\mu}(t) \) are, respectively, the average rates of speciation, immigration, and extinction for lineages in the region (Rabosky 2009). These are given as functions of time because they may change with regional age. The expected number of species in the region is then

\[ S = e^A. \]

Assuming that there are no intrinsic limits to regional diversity, the number of species is expected to be higher in older regions, because there is more time for diversification and immigration, and in regions where immigration and net diversification are high. Species abundances and niche breadths should be independent of increasing diversity and should be uncorrelated among coexisting species (e.g., Ricklefs 2012). However, if niche space fills as a region accumulates species, immigration or net diversification will decline as diversity increases, analogous to density-dependent population growth. Regional diversity will become independent of regional age and will be set by ecological limits (Cornell 2013).

The roles of regional age, diversification rates, and ecological limits in determining regional diversity have been addressed by a growing number of phylogenetic and paleontological studies. A recent review of these analyses found some evidence supporting ecological limits to regional diversity over evolutionary or biogeographic time, especially for small taxa in geographically restricted regions such as islands. However, species accumulation generally slows but does not stop completely as diversity increases (Cornell 2013). Regional age and diversification rates therefore remain potent sources of differences in regional species diversity.

Diversification Rate

Diversification rates as causes of variation in regional and local diversity are exemplified by North American damselflies (McPeek & Brown 2000). Local diversity is 4–5 times higher in lakes with fish rather than dragonflies as top predators; phylogenetic evidence implicates higher diversification rates in the damselfly clades adapted to fish lakes, although the greater age of the fish-lake habitat may also play a role. Small to nonexistent niche differences among coexisting fish-lake species suggest that diversification was due to nonadaptive modes of speciation (McPeek & Brown 2000). In another example, angiosperm diversity at local and regional scales is higher in Eastern Asia than in Eastern North America because of higher physiographic complexity in the former (Qian & Ricklefs 2000). Higher climatic stability in Eastern Asia may also contribute to this pattern.

Regional Age

Time for diversification as a cause of variation in regional and local diversity is illustrated by tree frogs (Hylidae), which reach extraordinary richness in Amazonia. Both regional and local diversity are strongly related to time since colonization and are unrelated to climate, diversification rates, or
morphological disparity (Wiens et al. 2011). Time for diversification is considered to be at least a secondary contributor to regional and local diversity in mangroves (Ricklefs et al. 2006). Climatic age, i.e., similarity of the present climate to a historic climate, appears to explain regional and local diversity in Californian plants (Anacker & Harrison 2012; see Regional Area and Productivity).

Time for immigration has been studied less extensively but probably plays a role in limiting diversity in some regions. For example, many European tree species have not immigrated back into climatically suitable habitats since the last glacial maximum because of geographical constraints on dispersal (Svenning & Skov 2004). Local and regional tree diversity patterns in central and Northern Europe can thus be largely explained by insufficient time for immigration (Svenning & Skov 2010).

Regional Area and Productivity

Regional area and productivity are influences whose mechanisms are not clear a priori; they may affect regional diversity through their effects on either diversification rates or regional carrying capacity (more individuals, niches or habitat types; reviewed by Hurlbert & Jetz 2010, Cornell 2013). Regional area effects are illustrated by gall wasps (Cynipidae) on oaks in the California Floristic Province (Cornell 1985). Each wasp is a specialist parasite on one or a few species of oaks (Quercus). The geographic ranges of individual oak species can thus be considered “regions” for their gall wasp faunas. The Californian oaks that are more broadly distributed support larger “regional pools” of gall wasp species. In turn, the regional pool size for each oak species predicts the number of wasp species found on local oak populations (Cornell 1985).

Regional productivity effects are shown by plants in the California Floristic Province, where plant productivity is under strong control by rainfall. Regions of California with wetter climates are more diverse because they support older and more species-rich evolutionary lineages, consistent with their origins in the wetter Eocene past (Anacker & Harrison 2012). This regional pattern trickles down to cause richer local communities in wetter regions. The regional-to-local direction of causality is likely because there is no direct link between climate and species richness or phylogenetic diversity at the local scale (Harrison et al. 2006, Anacker & Harrison 2012).

Immigration Rate

Immigration as a contributor to regional and local diversity is illustrated by corals in the western Pacific Ocean (Cornell et al. 2008). Near Indonesia there are > 600 species, but regional diversity declines to < 150 species in the Central Pacific Ocean. Local diversity at the 10-m transect scale closely parallels this gradient (Cornell et al. 2008). Species composition across the gradient shows disjunct, nested patterns consistent with historical geology and limited immigration (Keith et al. 2013). Prevailing currents in the Indonesian archipelago apparently make it a species accumulator (Connolly et al. 2003). Habitat area, heterogeneity, and environmental conditions tend to reinforce the effects of limited immigration on the coral diversity gradient (Connolly et al. 2003, Keith et al. 2013).

The Importance of Range Size

Regional diversification will not filter down to influence local communities if it is driven by allopatric speciation with little subsequent expansion among localities. If all species remain where they arrived or arose, then the species pool concept would be meaningless and communities would comprise exclusively endemic species. In eastern North American pond turtles, regional richness is higher in the south because of greater time for diversification, but local richness is
not higher in southern regions; instead, range sizes are smaller, and species are allopatrically distributed (Stephens & Wiens 2003). In many cases, however, shifting environments and dispersal barriers among localities and the evolution of key innovations allow populations to expand their distribution and become part of the species pool for other localities (Agosta & Klemens 2008). Indeed, positive correlations between divergence time and range overlap suggest that species tend to arise in allopatry and subsequently spread into sympatry with their relatives and other species (e.g., Fitzpatrick & Turelli 2006, Pigot & Tobias 2013). Regional diversification also will not filter down to affect local communities if it is driven by ecological speciation into novel habitats that are highly spatially segregated; however, such shifts may be rarer than ecological speciation events within a habitat (e.g., McPeek & Brown 2000).

**HOW ARE REGIONAL SPECIES POOLS DEFINED AND MEASURED?**

“Regions” and “localities” are perhaps easiest to define in the case of large mainlands that provide colonists to small islands and are little affected by immigration from the islands. In other cases, the appropriate definitions may be more ambiguous. Conceptually, the local community is typically considered to be the scale at which species interactions can actually or potentially affect abundance and diversity (Srivastava 1999; Figure 1); implicitly, these effects are manifested within an “ecological” time period, such as a few years to a decade. It has been argued that the local community is too vaguely delimited to be useful (Ricklefs 2008). However, regions and regional species pools have an even broader range of definitions. To assess the prevalence of alternative ways of defining and measuring regional species pools, we comprehensively reviewed the literature and identified 63 studies that used regional richness to predict local richness (Table 1; also Supplemental Table 1, follow the Supplemental Material link from the Annual Reviews homepage at http://www.annualreviews.org). We summarize the major issues below, moving from conceptual issues to more practical considerations.

**Unfiltered Versus Filtered Pools**

The species pool can be defined as all the species present in a region that can disperse to a focal locality regardless of their ability to tolerate the prevailing environmental conditions (e.g., Gough et al. 1994; see Table 1 for 13 studies using unfiltered pools). This definition is well suited for studies that search for patterns in local community assembly, including alpha and beta diversity, because it lends itself to identifying the abiotic and biotic filters that admit or exclude species from particular localities. However, this definition is less well suited to direct tests for regional species pool effects, because species that are not adapted to the envelope of environmental variables that characterize a set of communities cannot contribute to the diversity and composition of those communities (Belmaker & Jetz 2012). Also, if the regional species pool is defined to include many habitat specialists incapable of inhabiting focal sites, and if species-rich regions have more different types of habitat, “pseudosaturation” may result—a false appearance of local limits to species richness (Srivastava 1999).

Direct tests for species pool effects therefore usually define the pool to include only species that can both disperse to and potentially persist in a focal locality, in other words, a filtered pool (Srivastava 1999; see Figure 1 and Table 1 for 50 studies using filtered pools). Filtered pools are defined at ecological timescales and assume that evolutionary transitions to other habitat types are not occurring. The regional species pool for, say, a wetland locality may consist of all species observed at that type of wetland within a region (Pártel et al. 1996, 2011); the pool for a coral reef locality consists of all species observed in the same habitat type (reef crest, slope, or flat) within
<table>
<thead>
<tr>
<th>Citation (study system)</th>
<th>Pool definition</th>
<th>Pool delineation</th>
<th>Region (data source)</th>
<th>Locality (data source)</th>
<th>Max. regional scale</th>
<th>Min. local scale</th>
<th>Regional effect shape, strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belmaker et al. 2008 (Global; fish, corals)</td>
<td>Filtered</td>
<td>Biogeographic</td>
<td>Biogeographic region (field samples)</td>
<td>Coral reef sites (inventory)</td>
<td>100,000–10,000,000 km²</td>
<td>15 m²</td>
<td>Both curvilinear, strong</td>
</tr>
<tr>
<td>Lavenne et al. 2000 (Africa, Asia, South America; primates)</td>
<td>Filtered</td>
<td>Biogeographic</td>
<td>Biogeographic region (field samples)</td>
<td>Forest site (inventory)</td>
<td>10,000–80,000 km²</td>
<td>1 m²</td>
<td>Both curvilinear, strong</td>
</tr>
<tr>
<td>Lawton et al. 1993 (Global; bracken fern insects)</td>
<td>Filtered</td>
<td>Biogeographic</td>
<td>Biogeographic region (field samples)</td>
<td>Bracken fern (field samples)</td>
<td>10,000–200,000 km²</td>
<td>300–3,000 m²</td>
<td>Linear, medium</td>
</tr>
<tr>
<td>Wimberly et al. 2004 (Global; epifaunal invertebrates)</td>
<td>Filtered</td>
<td>Biogeographic</td>
<td>Biogeographic region (field samples)</td>
<td>Rock wall segment (field samples)</td>
<td>450–1,000,000 km²</td>
<td>15–60 m²</td>
<td>Linear, strong</td>
</tr>
<tr>
<td>Grace et al. 2011 (Oregon in USA; vascular plants)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Siskiyou plant community (field samples)</td>
<td>Vegetation site (field samples)</td>
<td>1 km²</td>
<td>10,000 m²</td>
<td>Linear, weak</td>
</tr>
<tr>
<td>Pärtel et al. 1996 (Estonia; vascular plants)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Estonian plant community (field samples)</td>
<td>Vegetation site (field samples)</td>
<td>5,000 km²</td>
<td>2,500 m²</td>
<td>Linear, medium</td>
</tr>
<tr>
<td>Stohlgren et al. 2008 (Central USA; exotic/native vascular plants)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Landscape (field samples)</td>
<td>Vegetation plot (field samples)</td>
<td>10,000 km²</td>
<td>100–1,000 m²</td>
<td>Linear, strong; linear, weak</td>
</tr>
<tr>
<td>Calvete et al. 2004 (Spain; parasite plants)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Host population (field samples)</td>
<td>Single host (field samples)</td>
<td>15–60 m²</td>
<td>0.25 m²</td>
<td>Curvilinear, strong</td>
</tr>
<tr>
<td>Cornell 1985 (California in USA; gall wasps)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Host distribution (inventory)</td>
<td>Host population (field samples)</td>
<td>10,000–1,000,000 km²</td>
<td>10,000 m²</td>
<td>Linear, strong</td>
</tr>
<tr>
<td>Dawah et al. 1995 (United Kingdom; gall wasps on grasses)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Host distribution (field samples)</td>
<td>Grassland site (field samples)</td>
<td>150,000 km²</td>
<td>10,000 m²</td>
<td>Linear, strong</td>
</tr>
<tr>
<td>Hawkins &amp; Compton 1992 (South Africa; fig wasps, fig wasp parasites)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Host distribution (field samples)</td>
<td>Single tree (field samples)</td>
<td>Unknown</td>
<td>25 m²</td>
<td>Curvilinear, strong; linear, strong</td>
</tr>
<tr>
<td>Kennedy &amp; Guégan 1999 (United Kingdom; eel parasites)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Host population (field samples)</td>
<td>Single host (field samples)</td>
<td>25 km²</td>
<td>0.01 m²</td>
<td>Curvilinear, strong</td>
</tr>
<tr>
<td>Krasnov et al. 2006 (California in USA, Slovakia, Israel; flea)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Host distribution (inventory)</td>
<td>Single host (field samples)</td>
<td>200,000 km²</td>
<td>0.01 m²</td>
<td>Curvilinear, strong</td>
</tr>
<tr>
<td>Morand et al. 1999 (Global; fish ectoparasites)</td>
<td>Filtered</td>
<td>Community type</td>
<td>Host population (field samples)</td>
<td>Single host (field samples)</td>
<td>25 km²</td>
<td>0.01 m²</td>
<td>Curvilinear, strong</td>
</tr>
<tr>
<td>Harrison et al. 2006 (California in USA; serpentine vascular plants)</td>
<td>Filtered</td>
<td>County/ecoregion</td>
<td>Intracounty serpentine outcrops (inventory)</td>
<td>Serpentine outcrops (field samples)</td>
<td>1,000–15,000 km²</td>
<td>10,000 m²</td>
<td>Not given</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Citation (study system)</th>
<th>Pool definition</th>
<th>Pool delineation</th>
<th>Region (data source)</th>
<th>Locality (data source)</th>
<th>Max. regional scaleb</th>
<th>Min. local scaleb</th>
<th>Regional effect, shape, strengthb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rivadeneira et al. 2002 (Chile; intertidal herbivores)</td>
<td>Filtered</td>
<td>Dispersion field</td>
<td>Seascapes (field samples + inventory)</td>
<td>Rocky intertidal transect (field samples)</td>
<td>111 km²</td>
<td>1 m².d</td>
<td>Linear, weak</td>
</tr>
<tr>
<td>Stendera &amp; Johnson 2005 (Sweden; lake/stream macroinvertebrates)</td>
<td>Filtered</td>
<td>Ecoregion</td>
<td>Ecoregion (field samples)</td>
<td>Shoreline or stream site (field samples)</td>
<td>10,000–200,000 km²</td>
<td>100 m²</td>
<td>Saturated, weak; curvilinear, weak</td>
</tr>
<tr>
<td>Duncan et al. 1999 (New Zealand; vascular plants)</td>
<td>Filtered</td>
<td>Habitat patch</td>
<td>Forest light gap (field samples + census)</td>
<td>Quadrat (field samples)</td>
<td>0.0001–0.00095 km³</td>
<td>0.2 m²</td>
<td>Curvilinear, medium</td>
</tr>
<tr>
<td>Gubbaid et al. 2005 (Brazil; termites)</td>
<td>Filtered</td>
<td>Habitat patch</td>
<td>Forest remnant (field samples)</td>
<td>Sampling unit (field samples)</td>
<td>0.00024–0.0006 km³</td>
<td>0.25 m³</td>
<td>Saturated, weak</td>
</tr>
<tr>
<td>Laws &amp; Obiri 2003 (South Africa; vascular plants)</td>
<td>Filtered</td>
<td>Habitat patch</td>
<td>Forest light gap (field samples + inventory)</td>
<td>Quadrat (field samples)</td>
<td>0.000032–0.000236 km³</td>
<td>0.25 m³</td>
<td>Curvilinear, weak</td>
</tr>
<tr>
<td>Soares et al. 2001 (Brazil; ants)</td>
<td>Filtered</td>
<td>Habitat patch</td>
<td>Forest remnant (field samples)</td>
<td>Quadrat (field samples)</td>
<td>0.04–0.3 km³</td>
<td>1 m³</td>
<td>Curvilinear, medium</td>
</tr>
<tr>
<td>Starzomski et al. 2008 (British Columbia; moss arthropods—August 2004)</td>
<td>Filtered</td>
<td>Habitat patch</td>
<td>Moss patch (inventory)</td>
<td>Moss plug (inventory)</td>
<td>0.000000099 km³</td>
<td>0.0079 m³</td>
<td>Saturated, weak</td>
</tr>
<tr>
<td>Winkler &amp; Kampichler 2000 (Austria; collembola)</td>
<td>Filtered</td>
<td>Habitat patch</td>
<td>Grassland remnant (field samples)</td>
<td>Quadrat (field samples)</td>
<td>0.0023–0.0375 km³</td>
<td>2.25 m³</td>
<td>Curvilinear, weak</td>
</tr>
<tr>
<td>Srivastava 1999 (North America; sunfish parasites)</td>
<td>Filtered</td>
<td>Host distribution</td>
<td>Host distribution (inventory)</td>
<td>Discrete water body (field samples)</td>
<td>58,500 km²</td>
<td>5 km²</td>
<td>Linear, medium</td>
</tr>
<tr>
<td>Ekroos &amp; Kuusarait 2012* (Finland; grasslands)</td>
<td>Filtered</td>
<td>Landscape</td>
<td>Seminatural grassland (field samples)</td>
<td>Grasland field (field samples)</td>
<td>0.25 km²</td>
<td>0.02 km²</td>
<td>Linear, medium</td>
</tr>
<tr>
<td>Freeman &amp; Harrison 2006 (California in USA; vascular plants)</td>
<td>Filtered</td>
<td>Landscape</td>
<td>Serpentine seep (inventory)</td>
<td>Serpentine seep quadrat (field samples)</td>
<td>0.02 km²</td>
<td>1 m².d</td>
<td>Linear, weak</td>
</tr>
<tr>
<td>Griffiths 1997 (North America; fish)</td>
<td>Filtered</td>
<td>Landscape</td>
<td>Multiple lakes (inventory)</td>
<td>Lake (field samples)</td>
<td>100,000–1,000,000 km²</td>
<td>0.1–100,000 km²</td>
<td>Curvilinear, medium</td>
</tr>
<tr>
<td>Kristiansen et al. 2011 (Western Amazon; palms)</td>
<td>Filtered</td>
<td>Landscape</td>
<td>Forest (field samples)</td>
<td>Transect (field samples)</td>
<td>120–40,000 km²</td>
<td>2.500 m²</td>
<td>Linear, strong</td>
</tr>
<tr>
<td>Sharlin 2000 (Global; planktonic lake crustaceans)</td>
<td>Filtered</td>
<td>Landscape</td>
<td>Multiple lakes (field samples)</td>
<td>Lake (field samples)</td>
<td>133–12,000,000 km²</td>
<td>0.1–20 km²</td>
<td>Linear, medium</td>
</tr>
<tr>
<td>Belote et al. 2009* (Appalachi in USA; predisturbance vascular plants)</td>
<td>Filtered</td>
<td>Large plot</td>
<td>Forest plot (field samples)</td>
<td>Forest plot (field samples)</td>
<td>0.02 km²</td>
<td>1 m².d</td>
<td>Saturated, weak</td>
</tr>
<tr>
<td>Borges &amp; Brown 2004 (Acrob; spiders, sucking insects)</td>
<td>Filtered</td>
<td>Large plot</td>
<td>Pasture (field samples)</td>
<td>Pasture plot (field samples)</td>
<td>0.0009 km²</td>
<td>0.02 km²</td>
<td>Curvilinear, weak, linear, medium</td>
</tr>
<tr>
<td>Collins et al. 2002* (Kansas in USA; grazed-ungrazed vascular plants)</td>
<td>Filtered</td>
<td>Large plot</td>
<td>Grassland plot (field samples)</td>
<td>Grassland quadrat (field samples)</td>
<td>0.0025 km²</td>
<td>0.1 m²</td>
<td>Linear, medium; curvilinear, strong</td>
</tr>
<tr>
<td>Kallimanis et al. 2006 (Greece; vascular plants)</td>
<td>Filtered</td>
<td>Large plot</td>
<td>Forest (field samples)</td>
<td>Forest quadrat (field samples)</td>
<td>0.0041 km²</td>
<td>0.0041 m³.d</td>
<td>Curvilinear, weak</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Citation (study system)</th>
<th>Pool definition</th>
<th>Pool delineation</th>
<th>Region (data source)</th>
<th>Locality (data source)</th>
<th>Max. regional scaleb</th>
<th>Min. local scaleb</th>
<th>Regional effect, shape, strengthb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Munguia 2004 (Atlantic Ocean; mobile-sséine epifaunal invertebrates after 12 days)</td>
<td>Filtered</td>
<td>Large plot</td>
<td>Subtidal patch (field samples)</td>
<td>Single shell (field samples)</td>
<td>0.000025 km²</td>
<td>0.01 m²</td>
<td>Both curvilinear, strong</td>
</tr>
<tr>
<td>Valone &amp; Hoffman 2002 (Arizona in USA; summer/winter desert vascular plants)</td>
<td>Filtered</td>
<td>Large plot</td>
<td>Vegetation plot (field samples)</td>
<td>Quadrat (field samples)</td>
<td>0.05 km²</td>
<td>0.25 m²</td>
<td>Linear weak; linear strong</td>
</tr>
<tr>
<td>Canning-Clode et al. 2009 (Atlantic Ocean; epifaunal invertebrates after 12 months)</td>
<td>Filtered</td>
<td>Locality-centered (propagate arrival)</td>
<td>Biogeographic region (field samples)</td>
<td>Settlement site (field samples)</td>
<td>400,000–2,000,000 km³</td>
<td>1 km²</td>
<td>Linear; not given</td>
</tr>
<tr>
<td>Roilos 2001 (Finland; bees)</td>
<td>Filtered</td>
<td>Map unit</td>
<td>Grid cell (atlas)</td>
<td>Farm (field samples)</td>
<td>25,000 km²</td>
<td>1 km²</td>
<td>Curvilinear, weak</td>
</tr>
<tr>
<td>Heino et al. 2003 (Finland; stream macroinvertebrates)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>Drainage basin (field samples)</td>
<td>Stream riffle (field samples)</td>
<td>756–7,742 km²</td>
<td>100 m²</td>
<td>Curvilinear, strong</td>
</tr>
<tr>
<td>Hugueny &amp; Paugy 2009 (Ivory Coast; fish)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>Drainage basin (inventory)</td>
<td>Stream pool (field samples)</td>
<td>800–97,000 km²</td>
<td>10,000 m²</td>
<td>Linear, strong</td>
</tr>
<tr>
<td>Marchant et al. 2006 (Australia; stream bank/channel invertebrates)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>Drainage basin (field samples)</td>
<td>Stream site (field samples)</td>
<td>1,400–24,000 km²</td>
<td>500 m²</td>
<td>Curvilinear medium; curvilinear strong</td>
</tr>
<tr>
<td>Oberdorff et al. 1998 (France; stream diatoms, richest depositional habitats)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>Drainage basin (field samples)</td>
<td>Stream site (field samples)</td>
<td>25–144 km²</td>
<td>1,000 m²</td>
<td>Linear, medium</td>
</tr>
<tr>
<td>Passy 2009 (North America; stream diatoms, richest depositional habitats)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>Drainage basin (field samples)</td>
<td>Stream site (field samples)</td>
<td>12,532–181,583 km²</td>
<td>1 m²</td>
<td>Both linear, medium</td>
</tr>
<tr>
<td>Soininen et al. 2009 (Finland; stream diatoms)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>Drainage basin (field samples)</td>
<td>Stream riffle (field samples)</td>
<td>5,000,000 km²</td>
<td>200 m²</td>
<td>Linear, weak</td>
</tr>
<tr>
<td>Vaughn 1997 (Central USA; stream mussels)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>Drainage basin (field samples)</td>
<td>River site (field samples)</td>
<td>295–105,000 km²</td>
<td>400 m²</td>
<td>Linear, medium</td>
</tr>
<tr>
<td>Angermeier &amp; Winston 1998 (Virginia in USA; native/non-native stream fish)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>Drainage basin (inventory)</td>
<td>Stream site (field samples)</td>
<td>3,500–25,000 km²</td>
<td>300–3,000 m²</td>
<td>Curvilinear, weak; linear, medium</td>
</tr>
<tr>
<td>Gröndahl &amp; Heino 2012 (Finland; stream macroinvertebrates— all taxa)</td>
<td>Filtered</td>
<td>Physiographic</td>
<td>2-km stream segment (field samples)</td>
<td>Stream riffle (field samples)</td>
<td>0.03 km²</td>
<td>50 m²</td>
<td>Linear, strong</td>
</tr>
<tr>
<td>Cornell et al. 2006 (Western Pacific Ocean; coralinvertebrates)</td>
<td>Filtered</td>
<td>Seascape</td>
<td>Archipelago (field samples)</td>
<td>Reef transect (field samples)</td>
<td>3,000–40,000 km²</td>
<td>10 m²</td>
<td>Linear, strong</td>
</tr>
<tr>
<td>Freestone &amp; Osman 2011 (Atlantic Ocean; epifaunal invertebrates)</td>
<td>Filtered</td>
<td>Seascape</td>
<td>Seacap (field samples)</td>
<td>Settlement site (field samples)</td>
<td>10,000 km²</td>
<td>1 km²</td>
<td>Saturated, strong</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

| Citation (study system) | Pool definition | Region (data source) | Locality (data source) | Max. regional scale | Min. local scale | Regional effect shape, strength
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Russell et al. 2006 (West coast in USA; intertidal invertebrates)</td>
<td>Filtered</td>
<td>Seascape</td>
<td>Rocky intertidal seascape (field samples)</td>
<td>20 km²</td>
<td>0.25 m²</td>
<td>Curvilinear, medium</td>
</tr>
<tr>
<td>Cornwell 1996 (Global; corals)</td>
<td>Unfiltered</td>
<td>Biogeographic</td>
<td>Biogeographic region (inventory)</td>
<td>Reef site (field samples)</td>
<td>10,000–1,000,000 km²</td>
<td>Linear, weak</td>
</tr>
<tr>
<td>Findley &amp; Findley 2001 (Atlantic/Indian Oceans; coral reef fishes)</td>
<td>Unfiltered</td>
<td>Biogeographic</td>
<td>Biogeographic region (inventory)</td>
<td>Reef transect (field samples)</td>
<td>400 m²</td>
<td>Curvilinear, medium</td>
</tr>
<tr>
<td>Srivastava 1999 (Global; tropical forest birds)</td>
<td>Unfiltered</td>
<td>Dispersion field</td>
<td>Map unit (atlas)</td>
<td>Vegetation site (field samples)</td>
<td>200,000–1,000,000 km²</td>
<td>0.15 km²</td>
</tr>
<tr>
<td>Alvarado &amp; Rodriguez 2004 (North America; nonvolant mammals, bats)</td>
<td>Unfiltered</td>
<td>Dispersion field</td>
<td>Map unit (atlas)</td>
<td>Map point (atlas)</td>
<td>160,000 km²</td>
<td>1 km²</td>
</tr>
<tr>
<td>Belmaker &amp; Jetz 2012 (Global; birds, amphibians, mammals)</td>
<td>Unfiltered</td>
<td>Dispersion field</td>
<td>Map unit (atlas)</td>
<td>Site (inventory)</td>
<td>31,400 km²</td>
<td>400 km²</td>
</tr>
<tr>
<td>Richardson et al. 1995 (Australia; Bankia vascular plants)</td>
<td>Unfiltered</td>
<td>Dispersion field</td>
<td>Map unit (atlas)</td>
<td>Vegetation site (field samples)</td>
<td>215 km²</td>
<td>10,000 m²</td>
</tr>
<tr>
<td>Stevens &amp; Willig 2002 (New World; bats)</td>
<td>Unfiltered</td>
<td>Dispersion field</td>
<td>Map unit (atlas)</td>
<td>Bat community (field samples)</td>
<td>Unknown</td>
<td>1,000 km²</td>
</tr>
<tr>
<td>White &amp; Hurlbert 2010 (North America; birds—5-year timescale)</td>
<td>Unfiltered</td>
<td>Dispersion field</td>
<td>Map unit (atlas)</td>
<td>BBS survey route (field samples)</td>
<td>325,000 km²</td>
<td>20 km²</td>
</tr>
<tr>
<td>Lessard et al. 2012b (North America; ants)</td>
<td>Unfiltered</td>
<td>Dispersion field</td>
<td>Map unit (atlas)</td>
<td>Sampling site (field samples)</td>
<td>3,000,000 km²</td>
<td>0.1 km</td>
</tr>
<tr>
<td>Ricklefs 2000 (Caribbean; birds)</td>
<td>Unfiltered</td>
<td>Island</td>
<td>Island (field samples)</td>
<td>Vegetation site (field samples)</td>
<td>180–75,000 km²</td>
<td>100,000 m²</td>
</tr>
<tr>
<td>Terborgh &amp; Faaborg 1986 (Caribbean rainforest; sclerophyll birds)</td>
<td>Unfiltered</td>
<td>Island</td>
<td>Island (inventory)</td>
<td>Vegetation tract (field samples)</td>
<td>4.8–30,000 km²</td>
<td>100,000 m²</td>
</tr>
<tr>
<td>Bini et al. 2000 (Neotropical; snakes)</td>
<td>Unfiltered</td>
<td>Map unit</td>
<td>Grid cell (atlas)</td>
<td>Grid cell (atlas)</td>
<td>1,102,500 km²</td>
<td>22,500 km²</td>
</tr>
<tr>
<td>Caley &amp; Schluter 1997 (Global; multiple taxa)</td>
<td>Unfiltered</td>
<td>Map unit</td>
<td>Grid cell (atlas + inventory)</td>
<td>Grid cell (atlas and inventory)</td>
<td>250,000 km²</td>
<td>2,500 km²</td>
</tr>
</tbody>
</table>

*Supplemental Material*

*Some studies did more than one test for a total of 193 separate correlations, 81 of which are indicated here. More complete information is presented in Supplemental Table 1.

*The maximum regional scale and minimum local scale examined are shown. Scale sometimes estimated by inspection.

*Strength taken as R² value (<0.3, weak; 0.3–0.5, medium; >0.5, strong) or sometimes estimated from inspection. Shape taken from original paper. A recent reclassification (Szava-Kovats et al. 2013) found fewer linear shapes than shown here.

*Indicates that other scales were examined and included in the supplementary table.

*Least disturbed locality. Other disturbance levels are listed in Supplemental Table 1.
a region (Cornell et al. 2008); regional species pools for heterogeneous plant communities might consist of all species found on the same soil type and within the same elevational band (Grace et al. 2011). Experiments can also be used to determine habitat affinities, but such testing is labor-intensive and usually impractical. Habitat affinities can also be determined by niche modeling or habitat occupancy data (Svenning & Skov 2004, Lessard et al. 2012a).

Biotic and abiotic filters may interact to exclude regional species from local sites, making the relevant species pool for any given site smaller than the abiotically filtered pool. For example, the specialization of plant species to particular soils could arise through interactions between soil fertility and competitive abilities (e.g., Elmendorf & Moore 2008) or between soil fertility and the impacts of herbivory (e.g., Fine et al. 2004). As long as these interactions consistently lead to the absence of a species from a soil type, the species is generally not considered a member of the regional pool for localities on that soil. Conversely, if the species is sometimes able to colonize a particular soil, depending on which competitor species happen to be present, it remains part of the regional pool for that soil. Clearly, only experiments combining addition of propagules that have been determined to tolerate the local abiotic conditions with manipulations of the resident community can fully decouple the contributions of biotic and abiotic filtering (e.g., Burns & Strauss 2011, Myers & Harms 2011).

Scales of Space and Time

The number of species available to colonize a locality will clearly increase with the length of time considered. Four studies in our survey examined the effects of timescale on species arrival at a locality and confirmed this expectation (Munguia 2004, Starzomski et al. 2008, Canning-Clode et al. 2009, White & Hurlbert 2010; see Supplemental Table 1). Availability will also vary with species- and site-specific factors influencing rates of arrival per unit time. These relationships can be expressed as follows:

\[ Z = \sum_{x=1}^{S} (1 - (1 - P_x(d, c))^T), \]

where \( Z \) is the number of species arriving over a period of \( T \) years that can inhabit the site, \( S \) is the total number of species in the region capable of inhabiting the site, and \( P_x \) is the annual probability of arrival of species \( x \), which depends on its dispersal ability \( (d) \) and landscape connectivity \( (c) \).

Ideally, regional species pools would be defined with reference to the effects of time, dispersal, and landscape structure. In practice, few options exist for explicitly measuring these variables and including them in species pool delineations (Lessard et al. 2012a). One problem that occurs as a result is pseudosaturation, that is, when the effective species pool in larger regions is overestimated because it does not take dispersal limitation into account (Srivastava 1999). The more common, if somewhat oblique, way to deal with the issue of timescales is by evaluating a range of regional spatial scales that are expected to correspond with different timescales. In one proposed scheme, the regional pool contains all species inhabiting a vegetation type, the actual pool or local pool contains those within a localized cluster of stands of that vegetation, and the community pool consists of those within a single stand (Pärtel et al. 1996). Nine other studies have applied a similar approach of examining a range of regional spatial scales (Table 1, Supplemental Table 1), and 12 have examined a range of local scales (Table 1, Supplemental Table 1).

Another very indirect way to deal with variable probabilities of arrival is to weight species in the region by their frequency of occurrence, on the assumption that more common species are more likely to disperse to a focal site. These frequencies are then used in generating a null model for predicting local community assembly (Lessard et al. 2012a).
Locality-Centered Definitions

Rather than defining the pool as a fixed geographic area, the dispersion field approach considers the regional species pool for each site to consist of all species whose geographic ranges overlap with that site or whose ranges overlap with an area of fixed size centered on that site (summarized by Carstensen et al. 2013). Seven studies in our survey use the dispersion field approach. Often, atlases or other large-scale spatial inventories are the sources of regional data. Such data generally include some species that cannot tolerate the environment in the focal site because they don’t include information on habitat affinities (Table 1). However, inventories integrate species occurrences over many years and are thus likely to provide a more complete estimate of the pool than other methods. Distributions must be mapped using sufficiently large grain sizes, because false-presence errors leading to overestimates of the species pool are larger at smaller grains (Belmaker & Jetz 2012). In most studies that use the dispersion field definition of the regional pool, local richness is obtained from habitat-specific field samples or inventories (Table 1). However, the atlas-based, dispersion field approach has occasionally been used to define local richness (Bini et al. 2000).

Another modification of the dispersion-field method uses community occupancy data rather than geographic ranges to delimit regional species pools. The pool for each focal site includes all species found at all sites that share any species with that site. This species pool may be further refined by weighting each species by its community similarity to the focal site (the compositional overlap between the sites it occupies and the focal site), and/or by its incidence across all sites (as a measure of its abundance, and therefore its probability of dispersing to the focal site). Because the pool is defined using community survey data, such refinements can deliver a somewhat more filtered habitat-specific pool (Carstensen et al. 2013).

Propagule arrival experiments are also, in effect, a locality-centered way to define regional species pools. An example would be experimental settlement panels used to identify the benthic marine species capable of dispersing to and selecting a hard-substrate environment. The panel is examined at regular intervals to assess all the species that settle and establish at a given point locality (Canning-Clode et al. 2009). Sufficient time must be allowed for rare species to settle to ensure that estimates of the pool are reasonably complete.

Geographic Units or Community Types as Replicates

An alternative to the locality-centered approach is to delineate regional species pools using fixed geographic units such as map units (3 studies), counties or ecoregions (1 study), physiographic units such as river drainages (9 studies), landscapes or seascapes (8 studies), habitat patches (6 studies), or large study plots (6 studies). Diversity data from these larger spatial units are then used to predict community patterns in smaller localities within each unit (Table 1). Challenges to this approach include filtering the regional pool appropriately (as above), choosing an appropriate scale for the locality, and dealing with the potential spatial nonindependence of contiguous geographic units (see Srivastava 1999, Bini et al. 2000, Harrison et al. 2006). The latter challenge can also be addressed by using separate biogeographic regions to define the pools because true biogeographic regions share few if any species (8 studies; Table 1).

A related approach, already mentioned, uses different community types rather than fixed spatial units as replicate regions (10 studies; Table 1). For example, each species of oak is considered one “region” supporting its own species pool of cynipid gall wasps (Cornell 1985) or each identified vegetation type may be a region supporting its own pool of plant species (Pärtel et al. 1996). Inventory data from multiple oak species or multiple vegetation types—usually based
on large-scale, long-term surveys— are then used to create multiple replicate regional pools, and these are in turn used to predict the average richness (or other attribute) of local gall wasps or plant assemblages within each region. Community type definitions are particularly useful for detecting “dark diversity,” which is the proportion of the filtered pool that is absent from the community (Pärtel et al. 2011). But they may be misleading in some cases, for example, when saturation levels in the different community types are not identical (Srivastava 1999).

Biogeographic species pools may also be constructed using recently developed tools for distance-based clustering (grouping of sites based on species compositional similarity among those sites) or network modularity analysis (grouping of sites based on shared occurrence of species among those sites) as described by Carstensen et al. (2013). Input data can either be independent of local communities, such as atlas grid data, or be the community samples themselves. Potential advantages include a more ecologically realistic, objective estimate of the pool that is statistically independent of locally sampled richness (Carstensen et al. 2013). Because this approach is relatively new, it is not yet possible to evaluate its success in analyzing regional influences on local communities.

**Regional Pools Constructed from Local Samples**

Regional species pools may be estimated using the same samples that are used to derive local community richness rather than independent regional data (33 studies; Table 1). This is the normal approach in null model analyses of local community structure in which the localities provide the single pool of species from which null communities are drawn. Some dark diversity (e.g., rare species) may be missed, however, resulting in underestimation of the regional pool. Solutions are to sample sites until the cumulative species-versus-sites-sampled curve levels off or to use species richness estimators, which extrapolate from local samples to an asymptotic, sampling-independent value of cumulative regional species richness (e.g., Winkler & Kampichler 2000, Borges & Brown 2004, Bittman et al. 2004, Cornell et al. 2008, Grace et al. 2011, Ekroos & Kuussaari 2012). Another problem is that regional and local richness are estimated using the same data and thus the estimates are not statistically independent. One way to avoid this problem is to exclude each local site from the data used to estimate the regional pool for that site (Collins et al. 2002, Valone & Hoffman 2002, Gröndroos & Heino 2012).

**How Are Regional Species Pools Used?**

**Testing for Regional Versus Local Influences on Local Richness**

Regional species pools were used early on to test whether community richness was limited (saturated) by competitive exclusion by asking whether local richness increased linearly (= unsaturated) or curvilinearly (= saturated) with regional richness. Subsequent studies concluded that unsaturated patterns were more common than saturated ones (Lawton 1999), but a long and contentious literature critiqued the approach and in some cases offered improvements (reviewed by Harrison & Cornell 2008). In a recent development, Szava-Kovats et al. (2013) proposed log ratio–based linear regression to overcome the statistical biases of the original method and found that this new test detects more indeterminate outcomes and significantly more evidence for saturation.

A broader question about local richness is whether it is in any way influenced by regional richness, without focusing on constant local richness (saturation) as the alternative hypothesis. Instead, local environmental causes of variation in species richness (environmental filters) can be included in models, and their effects can be contrasted with regional influences. Such multivariate
models can also provide more accurate estimates of regional influences in cases in which local and large-scale influences vary collinearly (White & Hurlbert 2010). They can also sort out whether some environmental filters, such as climate, are acting regionally to determine the size of the pool or locally to determine how many species can coexist (Harrison et al. 2006, White & Hurlbert 2010, Grace et al. 2011, Belmaker & Jetz 2012, Grönonos & Heino 2012). Nineteen studies in our survey examined both local and regional influences on local richness, often with multiple analyses examining different taxa, habitats, time periods, spatial scales, levels of disturbance, and native versus introduced biotas. Eighteen of the analyses showed local influences to be more important, 18 showed regional influences to be more important, and 11 showed both to be about equally important (Supplemental Table 1). As expected, regional influences were more important for disturbed environments (Collins et al. 2002) and for introduced species (Angermeier & Winston 1998); however, few generalizations are possible given that authors often measured different numbers and kinds of local environmental variables.

Testing for Regional Influences on Beta Diversity

Dissimilarity among communities, or beta diversity, was originally conceived as a straightforward result of niche differences and spatial gradients in the environment, but the roles of chance and dispersal limitation among localities have also been shown to be important, especially in the context of testing neutral theory (Anderson et al. 2011). All else being equal, larger regional species pools should generate greater beta diversity as well as higher local species richness. Kraft et al. (2011) showed that beta diversity of trees among 0.01-ha subplots within 0.1-ha forest plots decreased with increasing latitude. But the relationship was considerably weakened when beta diversity was compared to null communities randomly drawn from the regional species pools (i.e., species found in the 0.1-ha plots) to assign species to localities (i.e., 0.01-ha subplots).

Testing for Regional Influences on Phylogenetic Structure

A growing body of literature asks whether species coexisting in local communities are more closely or distantly related to one another than expected by chance (underdispersed or overdispersed), where chance is determined by randomizations of the pool of species found at the set of local sites (Webb et al. 2002). Considerable attention has been given to defining the pool of species (or sites) appropriately, because larger pools tend to lead to underdispersion and vice versa (e.g., Swenson et al. 2006). However, only a few studies consider variation in the phylogenetic structure of regional species pools as an actual (i.e., biological, not statistical) cause of variation in the phylogenetic structure of local communities. Lessard et al. (2012b) showed that ant communities become more phylogenetically clustered with increasing latitude because of geographic variation in regional phylogenetic structure. Anacker & Harrison (2012) showed that phylogenetic diversity in local Californian plant communities is strongly determined at the regional scale by the biogeographic structuring of species pools along a rainfall gradient, whereas functional trait data indicated that phylogenetic diversity is little affected by niche availability or niche partitioning at the local scale.

Testing for Regional Influences on the Outcome of Experiments

Regional species pools can also help to explain idiosyncratic outcomes of ecological experiments replicated across different continents and climatic zones. For example, a watering experiment might have widely disparate effects on local diversity in grassland communities. For each site, the species in the pool could be categorized as mesic or xeric. The experimental outcome could then
be compared with a null model drawing from this regional pool, such that the observed increase in mesic species could be compared with the availability of mesic species in the pool (Lessard et al. 2012a).

CONCLUSIONS

As the appreciation for historical and contingent processes has grown, the traditional ecological paradigm of stable, equilibrial communities in which assembly is controlled by niche space and limiting similarity has been increasingly questioned. Phylogenetic analysis in particular has provided a powerful new tool for testing hypotheses about the large-scale causes of variation in diversity. At the same time, the wealth of mostly small-scale studies of biodiversity and ecosystem function illustrate that localities are where nutrients are taken up and biomass is generated and consumed, making the small scale still a critical focus of study.

Carefully constructed regional species pools can help bridge a formidable gap between scales. Analyses of variation in regional pool size examine the large-scale processes that generate diversity, whereas analyses of local communities within the context of their regional pools can illuminate how large-scale forces interact with local factors to shape the assemblages that we see on the ground (Figure 1).

The regional species pool paradigm depends on there being a significant disparity in timescales so that regions remain relatively constant over timescales during which local communities may undergo substantial turnover. Today’s biosphere is changing under the influences of biological invasions, nutrient deposition, climatic warming, and intensification of many other anthropogenic disturbances. Whether the regional species pool concept remains useful, we suggest, depends on the spatial scale and pace of these changes. To the extent that changes are both widespread and rapid, the concept of a relatively constant or at least very slowly changing regional pool donating species to variable, more rapidly changing local communities may begin to lose its predictive capacity.

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

This work was supported by the College of Agricultural and Environmental Sciences at the University of California, Davis. We thank the following individuals for comments on the manuscript and/or for sharing encouragement and insights: K. Graham, J.-P. Lessard, M. Pärtel, T. Price, D. Srivastava, and S. Strauss.

LITERATURE CITED

Contents

Prescriptive Evolution to Conserve and Manage Biodiversity
Thomas B. Smith, Michael T. Kinnison, Sharon Y. Strauss,
Trevon L. Fuller, and Scott P. Carroll

The Phylogeny and Evolution of Ants
Philip S. Ward

What Are Species Pools and When Are They Important?
Howard V. Cornell and Susan P. Harrison

Biogeomorphic Impacts of Invasive Species
Songlin Fei, Jonathan Phillips, and Michael Shouse

Mutualistic Interactions and Biological Invasions
Anna Traveset and David M. Richardson

The Evolution of Animal Domestication
Greger Larson and Dorian Q. Fuller

Complex Ecological Interactions in the Coffee Agroecosystem
Ivette Perfecto, John Vandermeer, and Stacy M. Philpot

Reversible Trait Loss: The Genetic Architecture of Female Ornaments
Ken Kraaijeveld

The Utility of Fisher’s Geometric Model in Evolutionary Genetics
O. Tenaillon

The Molecular Basis of Phenotypic Convergence
Erica Bree Rosenblum, Christine E. Parent, and Erin E. Brandt

Advances in the Study of Coevolution Between Avian Brood Parasites and Their Hosts
William E. Feeney, Justin A. Welbergen, and Naomi E. Langmore

Ecological Restoration of Streams and Rivers: Shifting Strategies and Shifting Goals
Margaret A. Palmer, Kelly L. Hondula, and Benjamin J. Koch
Warmer Shorter Winters Disrupt Arctic Terrestrial Ecosystems
   Elisabeth J. Cooper

Unifying Species Diversity, Phylogenetic Diversity, Functional Diversity, and Related Similarity and Differentiation Measures Through Hill Numbers
   Anne Chao, Chun-Huo Chiu, and Lou Jost

Trophic Cascades in a Multicausal World: Isle Royale and Yellowstone
   Rolf O. Peterson, John A. Vucetich, Joseph M. Bump, and Douglas W. Smith

Origins of Plant Diversity in the California Floristic Province
   Bruce G. Baldwin

Animal Phylogeny and Its Evolutionary Implications
   Casey W. Dunn, Gonzalo Giribet, Gregory D. Edgecombe, and Andreas Hejnowicz

A Multiscale, Hierarchical Model of Pulse Dynamics in Arid-Land Ecosystems

Population Biology of Aging in the Wild
   Deborah A. Roach and James R. Carey

Gecko Adhesion as a Model System for Integrative Biology, Interdisciplinary Science, and Bioinspired Engineering
   Kellar Autumn, Peter H. Niewiarowski, and Jonathan B. Puthoff

Biodiversity and Ecosystem Functioning
   David Tilman, Forest Isbell, and Jane M. Cowles

On the Nature and Evolutionary Impact of Phenotypic Robustness Mechanisms
   Mark L. Siegal and Jun-Yi Lue

Ecology and Evolution of the African Great Lakes and Their Faunas
   Walter Salzburger, Bert Van Boeckel, and Andrew S. Cohen

Biome Shifts and Niche Evolution in Plants
   Michael J. Donoghue and Erika J. Edwards

Using Ancient DNA to Understand Evolutionary and Ecological Processes
   Ludovic Orlando and Alan Cooper

Resolving Conflicts During the Evolutionary Transition to Multicellular Life
   Paul B. Rainey and Silvia De Monte

Speciation in Freshwater Fishes
   Ole Seehausen and Catherine E. Wagner
New From Annual Reviews:

Annual Review of Statistics and Its Application

Volume 1 • Online January 2014 • http://statistics.annualreviews.org

Editor: Stephen E. Fienberg, Carnegie Mellon University
Associate Editors: Nancy Reid, University of Toronto
Stephen M. Stigler, University of Chicago

The Annual Review of Statistics and Its Application aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

Complimentary online access to the first volume will be available until January 2015.

Table of contents:
- A Systematic Statistical Approach to Evaluating Evidence from Observational Studies, David Madigan, Paul E. Strang, Jesse A. Berlin, Martijn Schuemie, J. Marc Overhage, Marc A. Suchard, Bill Dumouchel, Abraham G. Hartzema, Patrick B. Ryan
- The Role of Statistics in the Discovery of a Higgs Boson, David A. van Dyk
- fMRI Imaging Analysis, F. DuBois Bowman
- Statistics and Climate, Peter Guttorp
- Climate Simulators and Climate Projections, Jonathan Rougier, Michael Goldstein
- Probabilistic Forecasting, Tilmann Gneiting, Matthias Katzfuss
- Bayesian Computational Tools, Christian P. Robert
- Bayesian Computation Via Markov Chain Monte Carlo, Radu V. Craiu, JeRey S. Rosenthal
- Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models, David M. Blei
- Structured Regularizers for High-Dimensional Problems: Statistical and Computational Issues, Martin J. Wainwright
- High-Dimensional Statistics with a View Toward Applications in Biology, Peter Bühlmann, Markus Kalisch, Lukas Meier
- Next-Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data, Kenneth Lange, Jeanette C. Papp, Janet S. Sinischer, Eric M. Sobel
- Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond, Elena A. Erosheva, Ross L. Matsueda, Donatello Telesca
- Event History Analysis, Niels Keiding
- Statistical Evaluation of Forensic DNA Profiler Evidence, Christopher D. Steele, David J. Balding
- Using League Table Rankings in Public Policy Formation: Statistical Issues, Harvey Goldstein
- Statistical Ecology, Ruth King
- Estimating the Number of Species in Microbial Diversity Studies, John Bunge, Amy Willis, Fiona Walsh
- Dynamic Treatment Regimes, Bihbas Chakraborty, Susan A. Murphy
- Statistics and Related Topics in Single-Molecule Biophysics, Hong Qian, S.C. Kou
- Statistics and Quantitative Risk Management for Banking and Insurance, Paul Embrechts, Marius Hofert

Access this and all other Annual Reviews journals via your institution at www.annualreviews.org.