Research

Phylogenetic structure of mammal assemblages at large geographical scales: linking phylogenetic community ecology with macroecology

Marcel Cardillo*
Centre for Macroevolution and Macroecology, Research School of Biology, Australian National University, Canberra, Australian Capital Territory 0200, Australia

Phylogenetic community ecology seeks to explain the processes involved in the formation of species assemblages by analysing their phylogenetic structure, and to date has focused primarily on local-scale communities. Macroecology, on the other hand, is concerned with the structure of assemblages at large geographical scales, but has remained largely non-phylogenetic. Analysing the phylogenetic structure of large-scale assemblages provides a link between these two research programmes. In this paper, I ask whether we should expect large-scale assemblages to show significant phylogenetic structure, by outlining some of the ecological and macroevolutionary processes that may play a role in assemblage formation. As a case study, I then explore the phylogenetic structure of carnivore assemblages within the terrestrial ecoregions of Africa. Many assemblages at these scales are indeed phylogenetically non-random (either clustered or overdispersed). One interpretation of the observed patterns of phylogenetic structure is that many clades underwent rapid biome-filling radiations, followed by diversification slowdown and competitive sorting as niche space became saturated.

Keywords: assembly rules; diversification; extinction; macroevolution; null models; speciation

1. INTRODUCTION

Since a pioneering paper was published by Webb et al. [1] a decade ago, analysis of the phylogenetic structure of species assemblages has grown rapidly into the prosperous research programme of phylogenetic community ecology [2–5]. Phylogenetic community ecology does not pose novel problems or raise new issues; it simply uses new kinds of data and new analytical tools to ask the same fundamental question that community ecologists have been struggling with for decades. This question is: if the set of species found in a particular place at a particular time is a limited subset of the species found within the broader region, are the processes that determine this subset predictable and deterministic, or neutral and stochastic? Attempts to answer this question have focused largely on tests of whether assemblages show significant structure: that is, whether assemblages consist of species that are non-random subsets of the regional pool of species, with respect to taxonomic composition or ecological attributes [6]. If assemblages are non-random, then the nature of the non-randomness may provide clues to the ecological and evolutionary processes underlying assemblage formation.

The idea that these processes might be revealed in the patterns of phylogenetic relatedness among species grew out of two suppositions. One was that species with similar ways of obtaining resources from the environment may compete, leading to exclusion of the weaker competitor [7]. The other supposition was that closely related species, having been derived from a recent common ancestor, are often ecologically similar, so may compete more strongly than more distant relatives [8]. If the principle of competitive exclusion operates in multispecies assemblages, we might expect to see fewer closely related species coexisting in a particular area than we would in the absence of interspecific competition. Early efforts to test this hypothesis systematically were based on analysing the ratio of species numbers to genus numbers (S/G ratios) within assemblages. The expectation was that in a competition-structured assemblage, S/G ratios should be lower than would be the case if the assemblage was simply a random subset of species drawn from the regional species pool (Jaccard 1922, cited in [9]). The early emphasis was on competition as a structuring process, although the role of habitat filtering in elevating the coexistence of close relatives has also been recognized [10].
With the rapid increase in availability of phylogenies in the 1990s and 2000s, these kinds of tests were extended to exploit the rich source of information represented by phylogenies. A phylogeny provides a more precise estimate of the degree of relatedness of two species than a taxonomic classification, lending more power (at least in principle) to tests based on relatedness. Analysing the phylogenetic structure of an assemblage and comparing it to null expectations is thus a natural extension of the analysis of S/G ratios. One early analysis of phylogenetic assemblage structure was presented by Haydon et al. [11], who simulated the radiation of a clade on an island archipelago. They measured phylogenetic structure of the simulated assemblages using the ratio of the mean branch length separating pairs of occurring species to the mean branch length separating all pairs of species, across the archipelago. A few years later, Webb [1] presented two similar metrics (net relatedness index (NRI) and nearest taxon index (NTI)) to measure phylogenetic structure of forest tree assemblages. Modified forms of these two metrics have been widely used since, and this paper, together with a follow-up review [5], was largely responsible for kickstarting the current wave of interest in phylogenetic community structure.

To date, most analyses of phylogenetic structure (with a few exceptions: [12,13]) have focused on ‘local-scale’ communities: that is, species found within areas considered to be within the dispersal capabilities of all individuals, for example, 0.16 ha rainforest plots [1] or 5 × 10 m fynbos plots [14]. The emphasis on local scales is probably because inter-specific competition remains a process of particular interest for community ecologists, and it is within small areas that the influence of competition is expected to be most evident. What about assemblages of species delimited within large geographical regions? So far, the principal role of large-scale assemblages in phylogenetic community ecology has been to provide the regional source pools from which local-scale null models are constructed. Little attention has been directed in describing and understanding the phylogenetic structure of the regional pools themselves. On the other hand, the macroecology research programme is aimed explicitly at understanding the structure of regional and continental assemblages. However, the emphasis in macroecology has been to analyse structure in terms of species-level traits, such as body size, population density and geographical range size, rather than phylogenetic relatedness [15,16]. Hence, analyses of the phylogenetic structure of large-scale assemblages helps to form a link between the two ecological research programmes of phylogenetic community ecology and macroecology.

In local-scale assemblages, phylogenetic structure has been interpreted primarily as the result of ecological sorting processes: as well as competition, the mechanism most commonly invoked is habitat filtering, where close relatives are brought together by shared habitat preferences [1]. At larger scales, it becomes necessary to consider macroevolutionary as well as ecological processes as determinants of the structure of assemblages [13,15]. But do ecological sorting processes like competitive exclusion or habitat filtering still play a role at large scales? Or are these kinds of processes rendered unimportant by habitat heterogeneity and opportunities for species to segregate spatially within large geographical regions? The answer has implications for understanding not only the formation of regional assemblages and geographical patterns of biodiversity, but also for the formation of local communities. For example, a competition-structured regional species pool could reduce the power of null models to detect the signature of competition at local scales [17].

In this paper, I examine the phylogenetic structure of mammal assemblages at scales larger than the local community. I begin by outlining some of the processes that may drive the formation of assemblages at these larger scales, and ask whether we should expect to see phylogenetic non-randomness in assemblage structure at these scales. Using African carnivores (order Carnivora) as a case study, I then explore patterns in the phylogenetic structure of assemblages within ecoregions, contiguous areas of structurally and ecologically similar habitat that range in size from 49 km\(^2\) to over 4.6 million km\(^2\). I ask whether ecoregion carnivore assemblages, in general, tend to show non-random phylogenetic structure, whether the choice of source pool makes a substantial difference to the detection of significant structure, and whether phylogenetic structure varies with ecoregion area. Answers to these questions may offer clues to the processes underlying assemblage formation at large geographical scales.

2. MECHANISMS PRODUCING NON-RANDOM ASSEMBLAGES AT LARGE GEOGRAPHICAL SCALES

When testing hypotheses about the processes structuring assemblages at local scales, the usual first step is to test if the assemblage actually shows significant structure, by comparing it to a set of null assemblages generated by randomly drawing species from a defined regional species pool, or shuffling species among sites within the regional pool [6]. The second step is to ask whether any structure (i.e. non-randomness with respect to a given measure of assemblage composition) is consistent with the predictions of the hypothesized process of assemblage formation. For large-scale assemblages, such as the set of species found within an ecoregion, a reserve or a geopolitical boundary, the same procedure can be applied. In these cases, we should ask whether the assemblage is a non-random subset of the species found (for example) in the broader biome, ecoclimatic region or continent. Below, I outline some of the mechanisms that could produce phylogenetic non-randomness of mammal assemblages at large geographical scales. Phylogenetic non-randomness may arise because the phylogeny itself carries the signal of historical events (macroevolutionary processes of speciation and extinction), or because phylogenetic relatedness is a proxy for ecological similarity (ecological sorting processes of habitat filtering and interspecific competition).
(a) Trait- or environment-linked speciation and extinction

Diversification rates in mammals appear to be geographically heterogeneous [18], either because rates are linked to biological properties of particular clades, which inhabit different regions, or because rates are linked to features of the environment. If rates of speciation or extinction are associated with heritable biological traits such as life-history speed or population density [19,20], a phylogeny may become increasingly imbalanced over time, with species increasingly concentrated within a few large clades. Depending on the extent to which rapidly diversifying clades occupy restricted geographical regions, the result will be geographical as well as phylogenetic heterogeneity in speciation or extinction rates. Hence, regions with clades that are speciating rapidly or going extinct slowly will contain a high proportion of closely related species, and assemblages in these regions should be phylogenetically clustered compared with the broader biome or continental source pool. Alternatively, speciation and extinction rates may be associated with environmental features such as topographic diversity or available energy, without necessarily being biased towards certain clades [21,22]. Again, elevated speciation rates or reduced extinction rates in these kinds of environments should produce phylogenetically clustered assemblages, compared with the source pool.

(b) Density-dependent speciation and extinction

Although some models of cladogenesis assume unlimited accumulation of species, it is also possible that diversification rates within clades are density-dependent, such that speciation rates slow or extinction rates rise as an equilibrium or saturation level of species richness is approached [23–28]. The likely effects of saturation on patterns of phylogenetic assemblage structure are difficult to predict, but high extinction rates could result in longer apparent branches, on average [29], which may lead to increased phylogenetic dispersion. Within a given region, if large clades are more likely to have become saturated than smaller clades [25], density-dependent diversification could serve to limit the degree to which the regional assemblage is phylogenetically clustered within the broader species pool. More precise predictions about the possible influence of density-dependent diversification on phylogenetic assemblage structure could be obtained from simulations.

(c) Habitat filtering and geographical range limits

Any assemblage can only consist of species whose environmental tolerance limits allow them to maintain a population in the relevant site or region. Hence, if adaptations to particular environmental conditions are shared by closely related species, an assemblage may become phylogenetically clustered compared with the broader region, a process referred to as habitat filtering [1,5]. At the level of local-scale communities, habitat filtering relates to the movement of individuals within a landscape: an animal may only establish a breeding territory in a particular location if conditions are suitable. At large geographical scales, habitat filtering can be thought of in terms of the differential expansion of species’ geographical ranges and the abiotic determinants of range limits. However, although environmental conditions determine which species cannot be present in a particular region, they do not necessarily predict which species will be found there. So, the degree to which habitat filtering can account for phylogenetic structure of assemblages at large scales will depend largely on (i) the discrepancy between species’ fundamental and realized niches (as expressed in the geographical range), and (ii) the strength of phylogenetic signal in species’ realized niches.

In mammals, the discrepancy between fundamental and realized niches can be considerable. Shifts in species’ geographical ranges following deglaciation of northern latitudes in the Late Pleistocene and through the Holocene appear to have been mostly idiosyncratic, with members of the same Pleistocene assemblage responding in different ways to climate and habitat shifts [30,31]. This suggests that there is not always a tight association between species assemblages and particular environments, at least in regions that have experienced major and rapid climate shifts. It is possible that in more climatically stable regions at lower latitudes, there is less of a disjunction between fundamental and realized niches, and a stronger association between assemblages and environments. If this is the case, the influence of habitat filtering on the phylogenetic structure of large-scale assemblages may be weaker in the tropics compared with high latitudes.

However, there is also reason to believe that phylogenetic signal in the realized niche of mammal species is low, even in climatically stable regions. On the one hand, geographical ranges may be limited by hard boundaries such as a large river or an abrupt change in habitat type, and the ability of species to breach these kinds of boundaries will be determined at least partly by traits shared among closely related species. However, more subtle environmental gradients can also limit the expansion of a species’ range. Indeed, a dispersal barrier may be defined as a selective gradient where a species’ available genetic variation is insufficient to allow populations at the range margin to adapt to changing conditions across the gradient [32]. Furthermore, the capacity for peripheral populations to adapt and spread may be limited by gene flow from denser populations at the core of the species’ range, where selective pressures may be different [32]. To the extent that available genetic variation and strength of gene flow within species’ ranges do not show strong phylogenetic signal, there is an idiosyncratic, non-phylogenetic component to species’ capacity for range expansion. This may limit the degree to which the match of species’ distributions to habitats is predictable from phylogeny, thus limiting influence of habitat filtering on phylogenetic structure of large-scale assemblages.

(d) Interspecific competition

Competition has long been a major focus of interest for researchers studying the structure and composition of local-scale assemblages. If interspecific competition
is mediated by phylogenetically conserved traits, a competition-structured assemblage is expected to be phylogenetically overdispersed, because competitive exclusion should be more likely among closely related species [1,5]. In mammals, interspecific competition has been shown to reduce population densities, or cause local extinction, of weaker competitors within local sites (e.g. [33–35]). It is therefore reasonable to expect that local-scale mammal assemblages are sometimes structured by competition. But does any influence of competition at the level of local populations scale up to an effect on the structure of large-scale assemblages [34,36]?

Theory predicts that even in the absence of significant abiotic barriers to dispersal, interspecific competition can produce stable geographical range limits that are narrower than would be the case without competition [32,37]. This is supported by numerous examples in mammals. For example, whereas the northern limits of the distribution of arctic foxes (Alopex lagopus) are determined by abiotic factors, the southern limits appear to be determined by competition for food and den sites with red foxes (Vulpes vulpes), which are ecologically similar but larger [38]. In North America, abundances of coyotes (Canis latrans) are lower, and mortality rates higher, in areas with high wolf (Canis lupus) densities [39]. Coyotes have expanded their distribution throughout much of the United States following recent extirpation of wolves, suggesting that their distributions were previously limited by competition with wolves [39].

Rapid and recent range shifts precipitated by invasive species present particularly clear evidence of range boundaries being limited by competition. In Britain, grey squirrels (Sciurus carolinensis) were introduced from North America in the late nineteenth century, and have since spread through most of England. Simultaneously, the range of the closely related native red squirrels (Sciurus vulgaris) has contracted in a pattern complementary to the expansion of the grey squirrels. The two species have similar food requirements, and similar patterns of space use and daily activity, but in broadleaf woodlands grey squirrels appear to have a competitive advantage, as they are able to digest acorns more efficiently and interfere in seed-hoarding by red squirrels [33,40]. It therefore seems very likely that competition has been a major factor in the replacement of red by grey squirrels across Britain [33].

Some macroecological patterns in mammals have been interpreted as the outcome of competition on geographical scales [34], and large-scale, assemblage-level analyses suggest that mammal geographical distributions may be limited by interspecific competition. Letcher et al. [41] showed that closely related mammal species of similar body size have less geographical overlap than expected under a null model. A similar finding was made for granivorous rodents by Bowers & Brown [42]. More recently, Davies et al. [43] found that a large part of the variation in range overlap among carnivore sister species was explained by divergence in carnassial tooth length, a key ecological trait in carnivores. This suggests that competition has either driven the evolutionary divergence of carnassial length between sympatric species, or has led to the exclusion of weaker competitors from their potential distributions.

3. LARGE-SCALE ASSEMBLAGE STRUCTURE OF AFRICAN CARNIVORES

The aim of the analyses presented here was to explore the structure of large-scale assemblages for a single taxonomic guild of mammals. I confined the analysis to African carnivores to avoid the extra complexity likely to be introduced by inter-ordinal or cross-continental comparisons. Carnivores were chosen because their geographical distributions are perhaps better-known than other mammal orders, and their phylogeny is one of the best-resolved, with 79 per cent of nodes resolved [44]. The basic question was whether assemblages tend to show significant phylogenetic structure consistent with any of the processes outlined above. I begin by analysing phylogenetic structure of carnivore assemblages within ecoregions (see below), using two metrics that capture different aspects of the average relatedness among species. I then test the influence of the choice of source pool, biome type and geographical area on the structure of assemblages.

(a) Methods

I used terrestrial ecoregions as the geographical unit for defining carnivore assemblages in Africa, excluding Madagascar. Ecoregions are a way of defining distinct biotas occurring within relatively large, contiguous areas of similar habitat and climate [45]. Ecoregions are grouped into biomes, ecologically similar regions that are not necessarily contiguous. I excluded from the analyses ecoregions found in the mangrove and inland water biomes, and in two biomes (temperate conifer forest and tropical and subtropical dry broadleaf forests) in which there was only one ecoregion (which would mean the assemblage and source pool are the same, see below). I identified the set of carnivore species found within each of the remaining 102 ecoregions using geographical range maps from the Global Mammal Assessment [46].

To calculate phylogenetic structure of ecoregion assemblages, I used two metrics, NRI and NTI [5]. NRI measures the standardized effect size of the mean phylogenetic branch length separating each pair of species in the assemblage. NTI is similar but measures only branch lengths separating each species from its nearest relative. NRI is thus a measure of the degree of phylogenetic clustering (species more closely related than expected) or overdispersion (species less closely related than expected) throughout the phylogeny, while NTI reflects clustering or overdispersion at the tips of the phylogeny. Branch lengths were from the mammal supertree of Bininda-Emonds et al. [44]. NRI and NTI were implemented using the functions ‘ses.mpd’ and ‘ses.mntd’, respectively, in the R library Picante [47].

Null models for assemblage structure of each ecoregion were constructed using two levels of source pool: biomes and the entire African continent. As implemented in Picante, NRI and NTI were calculated using distributions of 1000 random values of mean phylogenetic
distances between species, obtained by shuffling species among ecoregions. For biome-level source pools, species were shuffled among ecoregions within the same biome. For continental source pools, species were shuffled among ecoregions across the whole of Africa, excluding the ecoregions omitted from the analysis (as described above). Two-tailed \( p \)-values were obtained by comparing the observed mean pairwise distances (for NRI) and mean nearest neighbour distances (for NTI) with those from the distributions of random matrices. Assemblages with \( p < 0.025 \) were considered significantly clustered, and assemblages with \( p > 0.975 \) were considered significantly overdispersed. The raw data matrices and derived data values are provided in the electronic supplementary material.

**Phylogenetic structure: comparison of biomes and source pools**

Of the 102 terrestrial ecoregions included in the analysis, a large proportion show significant phylogenetic non-randomness in their carnivore assemblages, but the frequency and kind of non-randomness varies depending on the metric and the source pool used to construct null models (figure 1). When biome source pools are used (figure 1a), the majority of ecoregion assemblages have phylogenetic structure consistent with the null model. However, phylogenetic overdispersion is more prevalent than clustering, particularly for NTI, for which 21 ecoregions (20.8\%) are significantly overdispersed and three are clustered. When continental source pools are used (figure 1b), this pattern is reversed: for NRI and NTI, 68 and 26 ecoregions, respectively (67\% and 25.4\%), are significantly clustered, and none are overdispersed.

Can we infer anything about the processes of ecoregion assemblage formation from these general patterns in the frequency of phylogenetic clustering and overdispersion? Tentatively, I believe we can. If each clade originates in a particular area, then diversifies and spreads geographically, then each clade may (i) expand to occupy the contiguous area of suitable habitat represented by an ecoregion; (ii) expand beyond ecoregion boundaries to occupy the area of suitable habitat represented by the biome; or (iii) keep expanding beyond biome boundaries and throughout the continent, by radiating into a diverse range of habitat types. A clade’s geographical expansion may proceed by (iv) gradual diffusion from its centre of origin, so that area occupied by the clade is proportional to its age, or (v) rapid radiation into the available area of suitable habitat, before niches become filled and diversification slows.

The patterns of phylogenetic clustering and overdispersion in African carnivores appear to be consistent with scenarios (ii) and (v) above: that is, a pattern of biome-filling radiations followed by niche saturation and competitive sorting. That these were biome-filling rather than ecoregion-filling radiations is supported by the contrasting patterns obtained under biome-level and continental-level source pools. Under scenarios (i) or (iv) above, we would expect to see many ecoregion assemblages being phylogenetically clustered within biomes, but this is not the case (figure 1a). On the other hand, a large proportion of ecoregions are clustered within the continental fauna (figure 1b), which suggests that scenario (iii) is uncommon.

Differences in patterns of clustering and overdispersion at shallow (NTI) and deeper (NRI) levels in the phylogeny could indicate support for scenario (v) above. The higher frequency of clustered assemblages for NRI compared with NTI, under a continental source pool (figure 1b), suggests that if there were biome-filling radiations, many of them were ancient, with rates of diversification slowing in more recent times, as expected under density-dependent models of clade growth. This slowdown may have been driven by the saturation and increasingly fine subdivision of niche space within ecoregions, and a subsequent increase in interspecific competition. With biome source pools (figure 1a), competitive sorting among recent lineages may be reflected in the greater frequency of overdispersion closer to the tips of the phylogeny (NTI compared with NRI), where we expect competition to be most common.

Comparisons of patterns between different biomes (figure 2) may also be revealing. In the species-rich tropical biomes (tropical and subtropical moist broadleaf forests, and tropical and subtropical grasslands, savannas and shrublands), as well as in the montane grasslands and shrublands biome, the contrast between NRI clustering within the continental source pool (figure 2c) and NTI overdispersion within biome source pools (figure 2b) is particularly strong. If a process of ancient biome-filling radiations followed by more recent competitive sorting is a reasonable explanation for this contrast, as suggested above, then this process seems to have occurred primarily in the tropics. In the remaining, higher latitude, biomes (deserts and xeric shrublands, flooded grasslands and savannahs, and Mediterranean forests, woodlands and scrub), the contrasts between NRI and NTI, and between biome and continental source pools, are less sharp.
There are several reasons to predict that phylogenetic assemblage structure might vary with the size of ecoregions. Speciation rates are expected to increase with the size of regions, because the probability of \textit{in situ} speciation rises sharply in regions above a threshold area [48–50]. At the same time, extinction rates should decrease, because larger regions can accommodate larger populations, and are less vulnerable to environmental disturbances [51]. For both of these reasons, we might expect larger ecoregions to show more evidence of phylogenetic clustering than smaller ones. This association should be evident using null models constructed from both biome and continental source pools, if clades diversify and expand by gradual diffusion from a centre of origin (scenario (iv) above). On the other hand, if clades rapidly expand into the available area of suitable habitat (scenario (v) above), the association between large ecoregions and phylogenetic clustering will be less clear, or even non-existent. The patterns for African carnivores offer no strong evidence that larger ecoregions are more clustered (figure 3). For NRI, there is indeed a weak negative association with ecoregion area when continental source pools are used ($r^2 = 0.03$, $p = 0.06$), but when biome source pools are used the association is weakly positive ($r^2 = 0.06$, $p = 0.01$). There are no significant associations between ecoregion area and NTI. The lack of any strong tendency for large ecoregions to show phylogenetic clustering is consistent with the scenario of radiating clades spreading out rapidly to fill each biome, rather than diffusing gradually from centres of origin.

Another possible consequence of clades expanding rapidly, and saturating available niche space, is that smaller ecoregions may be more likely to show the signature of competition-structuring (phylogenetic overdispersion), since there may be fewer opportunities for similar species to avoid competition by spatial or habitat segregation, compared with large ecoregions. Again, this predicted pattern is not evident in the African carnivores, with overdispersion occurring in ecoregions of many sizes, including some of the largest (figure 3). Perhaps this should not be surprising if, as outlined earlier, competitive replacement can occur across areas larger than any African ecoregion.

\section*{4. CONCLUSIONS}

The analyses presented here are not intended to be an in-depth investigation of the structure and the formation of carnivore assemblages in African ecoregions. Rather, they are a preliminary, broad overview of the patterns of phylogenetic structure at large geographical scales. The
results show that many large-scale assemblages of African carnivores are phylogenetically non-random, and may provide some insight into the processes by which the assemblages were formed. Although the interpretation of the results is speculative, the patterns of clustering and overdispersion at different levels in the phylogeny appear to suggest a model of clade growth characterized by early biome-filling radiations followed by competitive sorting within ecoregions. This scenario fits with what seems be a growing recognition that models of density-dependent, ecologically limited clade growth are a better explanation for many patterns of diversity than models in which species richness increases with clade age or diversification rate [24–28,52].

Phylogenetic community ecology has hitherto focused primarily on local-scale assemblages and ecological processes, and how they are reflected in the assemblage phylogeny. Macroecology, on the other hand, has focused on large-scale assemblages and macroevolutionary processes, but has been largely non-phylogenetic. Analyses of the phylogenetic structure of large-scale assemblages provides a natural link between these two research programmes. One way of developing this link further would be to analyse distributions of ecologically important species-level traits in the context of phylogenetic assemblage structure, as has been done at the local scale (e.g. [53–55]). Another approach would be to simultaneously test pairwise patterns of species co-occurrence, ecological similarity and phylogenetic relatedness [13].

I thank Kate Jones and Kamran Safi for inviting me to contribute to this special issue, and Lindell Bromham and Matthew Phillips for comments on the manuscript. I am supported by an Australian Research Council QEII Fellowship and the Research School of Biology, Australian National University.

REFERENCES
7 Gause, G. F. 1934 *The struggle for existence*. Baltimore, MD: Williams & Wilkins.


