RESEARCH PAPER

Species co-occurrence and phylogenetic structure of terrestrial vertebrates at regional scales

Chuan Yan[1†], Yan Xie[2†], Xinhai Li[3], Marcel Holyoak[4] and Zhibin Zhang[1]*

[1]*State Key Laboratory of Integrated Management on Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China,* [2]*Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China,* [3]*National Zoological Museum of China, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China,* [4]*Department of Environmental Science and Policy, University of California, 1 Shields Avenue, Davis, CA 95616, USA*

*Correspondence: Zhibin Zhang, State Key Laboratory of Integrated Management on Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China.
E-mail: zhangzb@ioz.ac.cn
[†]The first two authors made equal contributions.

ABSTRACT

Aim To determine how taxonomic level and spatial scale affect the phylogenetic structure of species assemblages across four classes of terrestrial vertebrates.

Location Mainland China.

Methods Using species distribution data from Mammalia, Aves, Reptilia and Amphibia, including 2153 species from 2105 counties and 1632 species from 295 nature reserves across mainland China, we analysed the phylogenetic structure of co-occurring species at multiple taxonomic levels (class, order, family and genus) and spatial scales.

Results We found that phylogenetic clustering and unstructured patterns were more frequent than phylogenetic overdispersion in all groups. There was a higher frequency of phylogenetic clustering within classes and orders than within families and genera, while spatial scale had little effect on the frequency of phylogenetic clustering. Birds and mammals showed less frequent clustering patterns than amphibians and reptiles.

Main conclusions Phylogenetic clustering in terrestrial vertebrates was predominant over overdispersion at regional scales and higher taxonomic levels. Our results suggest that regional ecological and evolutionary factors, such as environmental filtering and speciation relative to extinction or colonization rates, are important in determining species assemblages of animals.

Keywords
Competitive exclusion, environmental filtering, neutral process, phylogenetic structure, spatial scale, species assemblage, taxonomic level, vertebrates.

INTRODUCTION

It has been a great challenge to reveal the ecological and evolutionary processes that drive species assembly in nature. An accumulation of phylogenetic information has allowed researchers to address this question by examining the phylogenetic structure of natural species assemblages (Webb *et al.*, 2002). Two main phylogenetic structures identified are phylogenetic clustering (closely related species tend to co-occur) and phylogenetic overdispersion (closely related species tend not to co-occur) relative to a null or neutral expectation. The mechanisms proposed in the literature to explain these structures include the ecological processes of environmental filtering and competitive exclusion and the evolutionary processes of speciation and extinction (Cavender-Bares *et al.*, 2009; Cardillo, 2011).

Environmental filtering predicts that assemblages will be phylogenetically clustered (Webb *et al.*, 2002; Emerson & Gillespie, 2008) provided that the relevant traits and habitat preferences show a phylogenetic signal, i.e. that closely related species are more similar (Den Boer, 1979; Cavender-Bares *et al.*, 2006). The competitive exclusion hypothesis predicts that competition prevents the coexistence of closely related species because they frequently share similar ecological traits and use similar resources (Elton, 1946; Hardin, 1960), which generally results in phylogenetic overdispersion (Emerson & Gillespie, 2008; but see exceptions by Mayfield and Levine, 2010). A third, more recent, view arises from spatial models of neutral communities (Hubbell, 2001), predicting phylogenetically unstructured patterns in local communities (Hubbell, 2001; Kembel & Hubbell, 2006). Lastly, Cardillo (2011) proposed that rapid spe-

ciation and/or slow extinction rates, coupled with limited mobility, or density-dependent rates, would produce phylogenetic clustering.

Consideration of spatial scale is important for interpreting the phylogenetic structure of species assemblages because the processes driving species assembly may vary between local and regional scales (Gómez *et al.*, 2010). At local scales, classic niche theory emphasizes the crucial roles of environmental filters and competitive exclusion (Cavender-Bares *et al.*, 2009). However, regional assembly is expected to be determined by at least three mechanisms. First, at large geographical scales, environmental filtering and interspecific interaction still affect the phylogenetic structure of species assemblages (Cardillo, 2011). Most species distribution models and environmental niche models assume that environmental filtering determines species geographical ranges. However, the extent to which environmental filtering can be reflected in phylogenetic structure depends on the discrepancy between species fundamental and realized niches and the strength of phylogenetic signal in species realized niches (Cardillo, 2011). Secondly, interspecific interactions, including conspecific attraction and interspecific territoriality, have been shown to affect the geographical ranges of animals (Brown *et al.*, 2000; Gotelli *et al.*, 2010). Thirdly, community assembly can be affected by historical events in the form of speciation and extinction, dispersal and colonization (Ricklefs, 1987, 2015). Speciation and extinction rates within clades can either vary with biological traits or environments (Weir & Schluter, 2007) or be density dependent (Phillimore & Price, 2008). Cardillo (2011) proposed that rapid speciation and/or slow extinction rates would produce a phylogenetic clustering structure at large geographical scales. Dispersal ability has been linked to geographical range size, colonization and competition of species (Kneitel & Chase, 2004). Dispersal and colonization history also play crucial roles in community assembly. Species with similar adaptive traits tend to colonize environments at the initial stage, which results in phylogenetically clustered communities, provided the relevant traits are conserved. As more species colonize over time, species interactions become more important and phylogenetic overdispersion is expected to emerge (Emerson & Gillespie, 2008).

The pattern of phylogenetic structure may depend on the taxonomic level at which communities are defined. Communities that are broadly defined taxonomically tend to show stronger signals of niche conservatism, while those from narrowly defined communities frequently show species interaction (Cavender-Bares *et al.*, 2006). For instance, Cavender-Bares *et al.* (2006) showed that communities including a single phylogenetic lineage, such as *Quercus*, *Pinus* or *Ilex*, tended to show phylogenetic overdispersion or null structure, while the inclusion of all seed plants resulted in phylogenetic clustering. Phylogenetic clustering may be dominant at higher taxonomic levels (including all species from taxa with a high rank in a taxonomic hierarchy), while phylogenetic overdispersion may be dominant at lower taxonomic levels (Cavender-Bares *et al.*, 2006).

To date, most studies have been conducted at local spatial scales (with a sampling area of a few hectares) and limited taxonomic groups. Notable exceptions include several studies of mammals and birds, conducted at regional scales (Lovette & Hochachka, 2006; Cooper *et al.*, 2008; Gómez *et al.*, 2010; Cardillo, 2011; Barnagaud *et al.*, 2014). Recent studies suggest that regional ecological and evolutionary processes also play very important roles in structuring species assemblage (Ricklefs, 1987, 2015; Cardillo, 2011). The frequencies of differing phylogenetic structures reported in the literature may be due to differences in studied taxa which differ greatly in their ecological traits, dispersal ability, life history and evolutionary history (Emerson & Gillespie, 2008; Cavender-Bares *et al.*, 2009; Vamosi *et al.*, 2009). Using multiple taxa would not only help to test the consistency of phylogenetic structures across different species groups, but would also provide insight into the roles of ecological factors (e.g. thermoregulation, habitat requirements, dispersal ability, etc.) and evolutionary history (e.g. speciation, extinction, colonization, etc.) in structuring species assemblages.

In this study, we used a large set of co-occurrence data for terrestrial vertebrates in mainland China to test several questions: (1) how does phylogenetic structure (clustering, overdispersion, unstructured) of co-occurring species differ between birds, mammals, amphibians and reptiles; (2) how does phylogenetic structure vary with the taxonomic level of observation (class, order, family and genus); and (3) how do the observed phylogenetic structures depend on the spatial scale of observation?

METHODS

Study area and species co-occurrence data

Mainland China spans over 60° in longitude and 35° in latitude and mostly lies in temperate climate zones, with its south in the subtropical zone and north in the subarctic zone. Its wide coverage and the tremendous differences in climate, topography and human impact result in great variation in the composition of species assemblages and community diversity across China.

Species distribution data from four classes (Mammalia, Aves, Reptilia and Amphibia) were obtained from the China Species Information Service (CSIS) (Xie *et al.*, 2009). These distribution data were compiled from many published resources including China Animal Annals for vertebrates, provincial animal annals, journal papers and survey reports from the 1920s onwards (Xie *et al.*, 2009). The database has up-to-date, comprehensive and detailed information on animal distribution of terrestrial vertebrates at county (*n* = 2153 species) and nature reserve (*n* = 1632 species) levels across mainland China (Mammalia = 431, Aves = 1094, Reptilia = 354, Amphibia = 274 in 2105 counties; Mammalia = 354, Aves = 905, Reptilia = 227, Amphibia = 146 in 295 nature reserves). County areas ranged from 13 to 208,000 km^2 with 75% > 1257 km^2 (Fig. 1a, c), while nature reserve areas ranged from 1 to 298,000 km^2 with 75% < 471 km^2 (Fig. 1b, d). The spatial scales explored here are best considered as variation within the regional scale (Gómez *et al.*, 2010; Cardillo, 2011). It should be noted that counties and nature reserves are likely to receive different levels of human disturb-
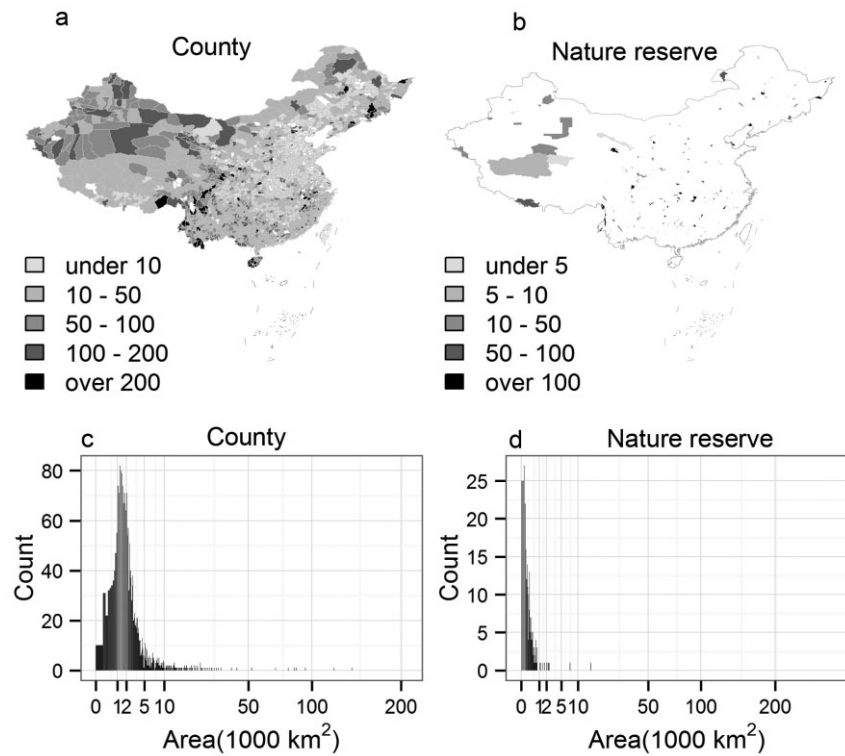
Figure 1 Species distribution data at county and nature reserve levels. (a) Number of species in different counties. (b) Number of species in different nature reserves. (c) Frequency distribution of county area (1000 km²). (d) Frequency distribution of nature reserve area (1000 km²).

ance. Thus, we avoid making inferences about the effects of spatial scales by comparing counties with nature reserves because human impact might be confounded with differences in spatial scale between nature reserves and counties; instead we took the approach of looking at spatial scale within either nature reserves or counties.

Co-occurrence metrics

We used two metrics based on presence–absence matrices to indicate species co-occurrence in this study. The first one is a co-occurrence index (CI) based on proportional similarity (Schoener, 1970), $CI_{ij} = 1 - 0.5\Sigma|p_{ik} - p_{jk}|$, where p_{ik} is the proportion of occurrences of species i in site k. The second one is a metric equivalent to a standardized version of the checkerboard score (DOI), $DOI_{ij} = (p_{ij} - p_i p_j) / p_i p_j$, where p_i is the proportion of sites where species i occurs and p_{ij} is the proportion of sites where species i and j co-occur (Hardy, 2008). Higher values of CI and DOI indicate a higher probability of co-occurrence, and lower values indicate the opposite. Using these two indices would make our results conservative in detecting the phylogenetic structures, since high Type 1 errors have been suggested to occur in previous studies (Kembel, 2009).

Phylogenetic distance of paired species

The phylogenetic distance of paired species was calculated from up-to-date global genetic supertrees constructed in previous studies. The supertrees for mammals and amphibians were provided by Fritz *et al.* (2009) and Pyron & Wiens (2011), respectively. The supertree for birds was constructed using phylogeny

subsets tools (http://birdtree.org/) (Jetz *et al.*, 2012). The proportions of species included in these phylogenetic supertrees in the dataset were as follows: Mammalia = 417/431, Aves = 886/ 1094, Amphibia = 155/274 in counties and Mammalia = 338/ 354, Aves = 693/905, Amphibia = 96/146 in nature reserves. We found no available supertree for the whole class Reptilia, and instead we used a supertree of order Squamata that allowed us to analyse its phylogenetic pattern from order to genus levels. We note that updating phylogenetic information for specific species may give different values of phylogenetic distance, but this might have a minor influence on our main results considering the large dataset we used.

Identification of phylogenetic structure

Identification of phylogenetic structures was conducted for species belonging to the four classes separately. There are two main categories of methods for measuring community phylogenetic structure (Kembel, 2009). The first is to measure the relatedness of species occurring together in samples (Webb *et al.*, 2002; Villalobos *et al.*, 2013), and the second is to measure the correlation between the co-occurrence and phylogenetic distance of paired species (Cavender-Bares *et al.*, 2004; Hardy, 2008).

The first category of methods are direct approaches to compute a phylogenetic structure of the community within a sample site or region, and include Webb's net relatedness index (NRI), which is a continuous measure of phylogenetic clustering or overdispersion. Its raw or relative variations rather than threshold values are useful measures of assemblage structure. NRI has the advantage of reflecting the phylogenetic structure of

species within a sample area, but it has two limitations. First, at lower taxonomic levels (e.g. genera), the sample size (species number) in a county or nature reserve would probably be too small to test the significance of the phylogenetic structure, which may be why it has seldom been used for studying the effects of taxonomic level in the literature. Second, the significance test of NRI needs to define a species pool based on some geographical area surrounding the focal assemblage, and the choice of species, although believed to reflect different evolutionary scenarios, largely affects the frequency of phylogenetic structures (Cardillo, 2011).

In this study, following previous work (Cavender-Bares *et al.*, 2006; Hardy, 2008; Kembel, 2009), we studied phylogenetic structure by measuring the correlations between species co-occurrence metric and phylogenetic distance of paired species. The approach has a direct ecological connection to pairwise species coexistence by measuring the species co-occurrence probability in space. The sample size of this approach for a given species pair is determined by sampling sites (i.e. the number of counties in this study), not limited by the number of species in each assemblage. Thus, this approach has the advantage of reflecting the differences in phylogenetic structure across different taxa, and can easily be used for studying both taxonomic level and spatial scale (see Cavender-Bares *et al.*, 2006). The significance of the correlation coefficient was judged at $P < 0.05$ by ranking it with those from 999 null models. Among previously proposed null models, the randomization of community matrices while fixing row and column sums has been given most credibility (Gotelli, 2000; Kembel, 2009). We applied this null model using a fast and unbiased algorithm (curveball algorithm) for large community matrices (Strona *et al.*, 2014). Based on null model analysis, we classified different outcomes as phylogenetic clustering (C) for significant negative correlations between phylogenetic distance and species co-occurrence metrics, phylogenetic overdispersion (O) for significant positive correlations or unstructured (U) for non-significant correlations. Our method combining distance-based metrics (CI and DOI) and the null model randomizing species-by-site matrix in the current study was chosen after reviewing the relevant literature (Gotelli, 2000; Hardy, 2008; Kembel, 2009; Strona *et al.*, 2014). All analyses were performed using R v.3.1.1 (R Development Core Team, 2013).

Testing the effects of taxonomic level and spatial scale on phylogenetic structure

To test the effects of taxonomic level, we examined phylogenetic structures by separating assemblages of species across spatial scales, classes, orders, families and genera. Then we summarized the frequencies of the three phylogenetic structures (clustered versus overdispersed versus null patterns) at the taxonomic level of order, family or genus from four classes, and compared the frequency of the phylogenetic structures among them to test the effects of taxonomic level.

To test the effects of spatial scale, we first measured the phylogenetic structure of taxa from counties and nature reserves separately; second, we measured the phylogenetic structure of taxa in counties or nature reserves with different area levels. We divided counties or nature reserves into four levels (Q1–Q4) according to their area using the *quantile* function in R (for counties, Q1–Q4 = 0.013–1.257, 1.258–1.984, 1.985–3.349, 3.35–207.975, respectively; for nature reserves, Q1–Q4 = 0.001–0.061, 0.062–0.192, 0.193–0.471, 0.472–298, respectively; all areas in 1000 km^2). Because there were too few species from some taxa to perform analysis, we restricted analyses to those with at least eight species (see 'Sample size' in Table S1 in the Supporting Information). We also measured the phylogenetic structure of taxa in counties or nature reserves with all area levels as comparison (defined as 'All' spatial scales hereafter).

The analyses described above provided data on phylogenetic structures of different taxa from four classes at various taxonomic levels and spatial scales. Because the frequency of phylogenetic overdispersion is too low in many categories, which gives unstable statistical inference in logistic regressions (Hosmer & Lemeshow, 2004), we downgraded the results to 'clustering' versus 'not clustering'. We applied generalized linear models (GLMs) to evaluate the factors (including main and pairwise interactive effects) in affecting the frequency of phylogenetic clustering. The response variable was a binary variable indicating presence/absence of phylogenetic clustering, and the predictor category variables included the class (including Mammalia, Aves, Reptilia and Amphibia), taxonomic level (including order, family and genus levels), spatial scale (data from 'All' or Q1–Q4 spatial scales were modelled separately) and site (including county and nature reserve). Inferences about their effects on the frequency of phylogenetic clustering were based on the best-fitting models selected by the corrected Akaike information criterion.

RESULTS

Phylogenetic structures in terrestrial vertebrates

At the class level, all correlations between species co-occurrence metrics (CI and DOI) and phylogenetic distance were significantly negative ($P = 0.001$–0.028) for data from all counties or nature reserves (Fig. 2, Table S2 in Supporting Information). Considering spatial scales Q1–Q4 separately did not change the overall pattern. These results suggest that phylogenetic clustering was consistently observed in co-occurring species at the class level.

GLMs showed significant effects of class (Mammalia, Aves, Reptilia and Amphibia) on the frequency of phylogenetic clustering (Table S3). When phylogenetic structures of multiple taxonomic levels and spatial scales were pooled for four classes, we found that the proportion of clustering patterns reflected by CI ranked as Reptilia (63%) > Amphibia (62%) > Mammalia (56%) > Aves (32%) at the county level. At the nature reserve level, the proportion of clustering patterns reflected by CI ranked as Amphibia (83%) > Aves (44%) > Reptilia (33%) > Mammalia (26%) (Fig. 3, Tables S4 & S5). Similar results were found using DOI, except for the rank of Reptilia (Table S4, Fig. S1). Hence,
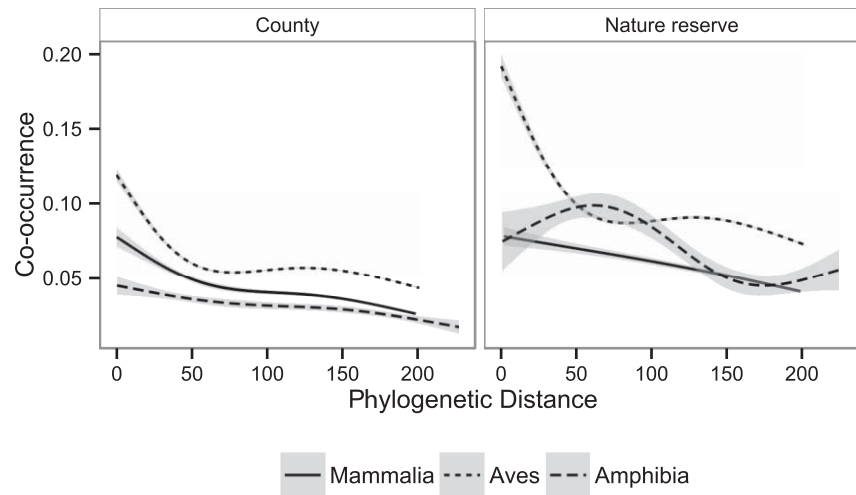
Figure 2 Relationship between co-occurrence (CI) probability and phylogenetic distance of pairwise species at county (a) and nature reserve (b) level, smoothed by cubic smoothing splines. The relationship between co-occurrence (CI) and phylogenetic distance for Reptilia is missing because a phylogenetic supertree is not available for the whole class of Reptilia.
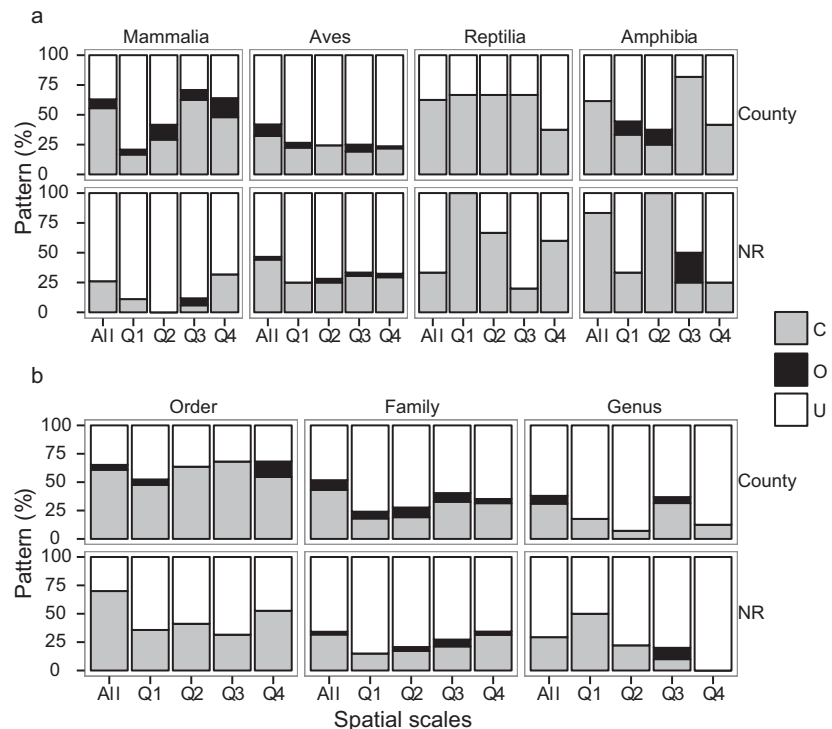

Figure 3 The proportion (%) of phylogenetic patterns (using co-occurrence index, CI) summarized by four classes of Aves, Mammalia, Amphibia and Reptilia (a) and taxonomic levels of order, family and genus (b) at different spatial scales (All and Q1–Q4) within counties or nature reserves (NR). Similar results were obtained using the standardized version of checkerboard score (DOI) (Fig. S1). C, phylogenetic clustering; O, phylogenetic overdispersion; U, unstructured pattern; Q1–Q4, representing four spatial scales (from small to large) defined by the area of counties or nature reserves.

overall, amphibian and reptile assemblages were phylogenetically clustered more frequently than mammals and birds.

Effects of taxonomic level on phylogenetic structures

At the order, family and genus levels, a large proportion of clustering (C) and unstructured (U) patterns were observed, and only a small proportion of cases were overdispersed (O) (Fig. 3, Tables S5 & S6). GLMs showed significant effects of taxonomic level (order, family and genus) on the proportion of phylogenetic clustering (Table S3). The pooled proportion of clustering patterns showed the rank of order > family > genus (for

co-occurrence metric CI; respectively 61%, 43%, 31% at county level and 70%, 32%, 29% at nature reserve level; results were similar for DOI; see Table S6). The proportion of unstructured patterns showed the opposite rank, with order < family < genus (for CI, respectively 35%, 48%, 62% in at county level and 30%, 66%, 71% at nature reserve level; similarly for DOI; se Table S6, Fig. S1). These results suggest that at higher taxonomic levels a higher proportion of phylogenetic clustering was observed.

Effects of spatial scale on phylogenetic structures

Spatial scale (county versus nature reserve; Q1–Q4) did not significantly affect the frequency of phylogenetic clustering (see

Table S3). For co-occurrence metric CI the pooled proportion of clustering structures were 25% for Q1, 28.6% for Q2, 40.4% for Q3 and 32.4% for Q4, at the county level, with equivalent figures of 26%, 25%, 23% and 32% for quartiles at the nature reserve level (see results for DOI in Table S6). These results suggest that the effect of spatial scale is minor within regional scales.

DISCUSSION

Phylogenetic structures and underlying driving forces

Recent regional studies of animals have raised novel hypotheses for explaining phylogenetic structure (Cardillo, 2011), but a thorough examination in multiple taxa and at multiple spatial scales is lacking. For the first time we investigated the phylogenetic structure in four classes of terrestrial vertebrates at multiple spatial scales and taxonomic levels. We provide clear evidence that at regional scales, phylogenetic clustering is predominant among different classes, taxonomic levels and spatial scales, while overdispersion is rare. Birds and mammals showed less frequent clustering patterns than amphibians and reptiles. The extent to which frequent clustering patterns were observed also increased with taxonomic level (genera to orders), but did not vary with spatial scale. Our results suggest that the ecological and evolutionary processes in regional species assemblages are different from those in local-scale assemblages (Cardillo, 2011).

That we frequently observed phylogenetic clustering but rarely observed overdispersion at regional scales can be well explained by environmental filtering and/or evolutionary processes. It is generally believed that closely related species compete for similar resources at local scales, which often results in overdispersion (Keddy & Weiher, 1999). While the large area covered by regional scales would allow coexistence of closely related species through the inclusion of more diverse habitats, closely related species tend to also have similar abiotic determinants, resulting in similar geographical ranges and a high probability of co-occurrence.

High phylogenetic clustering could also be caused by a high speciation rate and low extinction rate (Cardillo, 2011). It is suggested that species with larger geographical ranges speciate faster and loss of area increases extinction rates (Rosenzweig, 2001). Human impacts have been extensive in China, and local extinction rates of species may be high, but our data showed frequent phylogenetic clustering which might not be explained by the high speciation–low extinction hypothesis. Alternatively, the speciation and colonization rate may also influence the regional assembly processes. In China, there are many large mountains and rivers which might limit the spatial spread of species, and consequently colonization rates. A low colonization rate may cause new species to arise near to their closely related ancestors, which would result in more phylogenetic clustering patterns.

In our study we found a high proportion of unstructured patterns. A lack of structure might result from neutral processes, phylogenetic niche lability (niche conservatism by phylogenetic force versus ecological divergence favoured by natural selection) or a lack of phylogenetic resolution (Losos *et al.*, 2003; Hardy & Senterre, 2007). We adopted two co-occurrence metrics and a null model that fixed frequency and site and thus should control Type 1 error rates (Gotelli, 2000; Kembel, 2009). We also made our detection on the phylogenetic structures conservative. However, we cannot exclude the possibility of higher Type 2 error rates (Kembel, 2009); thus unstructured patterns could result from either weak or no phylogenetic signal or from neutral community processes. In general, we found more frequent unstructured patterns at lower taxonomic levels (families and genera) than at higher ones (orders and classes). This was probably caused by the difference in sample size (number of pairwise species) between higher and lower taxonomic levels. It would be more difficult to detect significant phylogenetic patterns at lower taxonomic levels with a smaller sample size, and also the range of phylogenetic variation (equivalent to effect size in a statistical test) is smaller at lower taxonomic levels.

Phylogenetic structures of terrestrial vertebrates

The observed high probability of unstructured phylogenetic patterns in terrestrial vertebrates matches many previous studies conducted at regional scales (Losos *et al.*, 2003; Cardillo *et al.*, 2008; Cardillo, 2011), which suggest that additional factors besides phylogenetic distance influence species assemblages. We also found a high level of phylogenetic clustering in mammals, which is inconsistent with the dominant phylogenetic unstructured patterns of global island mammal assemblages found by Cardillo *et al.* (2008). Cardillo (2011) and Cooper *et al.* (2008) found frequent unstructured patterns and phylogenetic overdispersion in carnivores and squirrels at global and continental scales, while these two taxa showed phylogenetic clustering in our study. This difference implies that the mechanism structuring species assemblages may be different between global and regional scales. Alternatively it may be because of geographical variation between China and other regions; this needs further investigation. The higher level of clustering than overdispersion observed in birds in our study is consistent with the results for birds at the scale of zoogeographical region, which was suggested to result from sympatric and parapatric speciation under environmental constraints (Barnagaud *et al.*, 2014). Previous studies found that wood warblers (at multiple 39.4-km transects) in North America and antbirds (at ecoregion scales) in Amazonian forests predominantly showed unstructured and overdispersed phylogenetic structures (Lovette & Hochachka, 2006; Gómez *et al.*, 2010). However, more phylogenetic clustering was found in tropical hummingbird communities of Ecuador in sampling areas of 0.07–25 km^2 (Graham *et al.*, 2009). The reasons for the differences among studies in birds may be complex; for instance low niche conservatism (Gómez *et al.*, 2010), repeatedly shifting geographical ranges induced by climate cycles (Lovette & Hochachka, 2006) or human alteration of distributions.

To our knowledge there have been no studies examining the difference in phylogenetic structures between four classes of terrestrial vertebrates. We found that reptiles and amphibians

tended to have more phylogenetic clustering patterns than birds and mammals. These patterns are consistent with differing strengths of environmental filtering among these groups. One possible explanation is that environmental filtering might have stronger effects on heterotherms like reptiles and amphibians than homeotherms like birds and mammals. Distribution ranges of heterotherms are more influenced by climate (such as temperature and water availability) than that of homeotherms (Sunday *et al.*, 2012). Homeotherms like mammals and birds also have more flexibility in activity periods (both daytime and seasonal) than amphibians and reptiles (Hut *et al.*, 2012). Broad differences among vertebrate taxa in environmental effects, behaviour, dispersal and evolutionary history merit further study in relation to phylogenetic structure.

Effects of the taxonomic level

Previous studies of local plant communities suggest that phylogenetic structure shifts from overdispersion to clustering when taxonomic levels increase (Cavender-Bares *et al.*, 2006). The general pattern was suggested to reflect a reconciliation of contradictory results driven by niche conservatism and species interactions in local communities (Cavender-Bares *et al.*, 2006). By using a wider range of taxa in animals at regional scales we also found the frequency of phylogenetic clustering increased with taxonomic level, but found no obvious change of overdispersion patterns with taxonomic level. Therefore, our results only partly support the observation from plant local communities (Cavender-Bares *et al.*, 2006). The absence or rare overdispersion patterns seen at the taxonomic levels in our study may be because at a regional scale closely related species have a high probability of coexisting within a large sample area although they are less likely to coexist within a sample area at a local scale. Provided there are phylogenetic signals in species traits and habitat preference, it can be hypothesized that increasing taxonomic or phylogenetic scale leads to increasing conservatism because traits within clades are less variable than traits among clades, while density-dependent species interaction should be more intense among close relatives within clades (Cavender-Bares *et al.*, 2009). Thus, our results could imply that lower and higher taxonomic levels might show a changed signal of phylogenetic niche conservatism but not apparently strengthen or reduce the signal of species interaction in structuring animal assemblages. Alternatively, the relatedness and sample size (number of pairwise species) at lower taxonomic levels was small, thus the probability of detecting the significant phylogenetic pattern would become smaller. Our results suggest that taxonomic level also affects the probability of observing particular phylogenetic patterns.

Effects of spatial scale

Spatial scale is another important factor in influencing phylogenetic structure (Keddy & Weiher, 1999; Cavender-Bares *et al.*, 2006; Emerson & Gillespie, 2008; Kooyman *et al.*, 2011). Some studies suggest that phylogenetic clustering is more often detected at large spatial scales, and that phylogenetic overdispersion is more frequently detected at small spatial scales (Cavender-Bares *et al.*, 2004; Emerson & Gillespie, 2008; Kooyman *et al.*, 2011). Conversely, there is a shift from phylogenetic overdispersion to clustering with decreasing geographical scale in antbird assemblages because niche conservatism is low (Gómez *et al.*, 2010). Vamosi *et al.* (2009) found that spatial scale of previous studies (which were mostly conducted at local scales) was not a significant predictor of phylogenetic structure. In general, we found more clustering but very few overdispersion structures at regional scales, contrary to many studies in which overdispersion was common at local scales. However, within regional scales, we found only a minor effect of spatial scale (from Q1 to Q4), which is consistent with the observation in Carnivora by Cardillo (2011). Assuming the probability of phylogenetic overdispersion shifting to clustering along a continuum of spatial scales (Fig. S2), these results demonstrate that a difference exists between local and regional scales but that the spatial scale within each of local or regional scales has little effect, suggesting that spatial scale from local to regional scales may have a strong nonlinear effect.

Species assemblages of larger regional scales are expected to show more evidence of phylogenetic clustering because speciation rate increases sharply with area (Losos & Schluter, 2000) and extinction is slowed down because large regions accommodate large populations (Purvis *et al.*, 2005; Cardillo, 2011). Under dispersal limitation, a larger spatial scale contains more species that diverge from one ancestral taxon, which would exhibit phylogenetic clustering. In this study, we noted that the environmental differences between counties and nature reserves might interfere with spatial scale, thus the effect of spatial scale on the frequency of phylogenetic structure was mainly inferred from the tests of areas of levels Q1–Q4 within nature reserves or counties. However, we found little evidence for spatially dependent change of phylogenetic structure.

Environmental heterogeneity is expected to be important in maintaining species co-occurrence within a given area (Kneitel & Chase, 2004), and may affect phylogenetic structure. In our study, the sampling area of different administrative regions or nature reserves varies to a large extent. The larger the sampling area, the larger the spatial heterogeneity of the environment would be within that space (Cavender-Bares *et al.*, 2006). In general, we found that sample area (Q1–Q4 in nature reserves or counties) had little effect on the observed frequency of phylogenetic structures. Our results suggest that heterogeneity is not important in affecting the observed frequency of phylogenetic structure at regional scales.

CONCLUSION

In a growing number of studies on the phylogenetic structure of co-occurring species, phylogenetic patterns serve as proxies for community assembly processes that link evolution, habitat filtering and functional trait dispersion (Emerson & Gillespie, 2008; Gerhold *et al.*, 2015). Compared with local communities, the mechanism structuring regional-scale assemblage may include in

addition to ecological processes the macroevolutionary processes of speciation and extinction (Cardillo, 2011). Our study at regional scales indicates that phylogenetic clustering and unstructured patterns were much more frequent than phylogenetic overdispersion in Chinese terrestrial vertebrates at multiple taxonomic levels and (regional) spatial scales. The higher degree of phylogenetic clustering in co-occurring amphibians and reptiles than mammals and birds is consistent with stronger environmental filtering and/or a low colonization rate. Taxonomic level, instead of spatial scale, exhibited a major effect on the probability of phylogenetic structures being observed. Our results suggest that environmental filtering, speciation relative to extinction or colonization may be the major driving forces for community assembly at regional scales. Future efforts should undertake to quantify functional traits of animals and incorporate them into analyses of phylogenetic structure. There is also a pressing need for studies of the co-occurrence patterns and phylogenetic structure of terrestrial vertebrates at small spatial scales.

ACKNOWLEDGEMENTS

REFERENCES

Barnagaud, J.Y., Daniel Kissling, W., Sandel, B. *et al.* (2014) Ecological traits influence the phylogenetic structure of bird species co-occurrences worldwide. *Ecology Letters*, **17**, 811–820.

Brown, J.H., Fox, B.J. & Kelt, D.A. (2000) Assembly rules: desert rodent communities are structured at scales from local to continental. *The American Naturalist*, **156**, 314–321.

Cardillo, M. (2011) Phylogenetic structure of mammal assemblages at large geographical scales: linking phylogenetic community ecology with macroecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2545–2553.

Cardillo, M., Gittleman, J.L. & Purvis, A. (2008) Global patterns in the phylogenetic structure of island mammal assemblages. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 1549–1556.

Cavender-Bares, J., Ackerly, D.D., Baum, D.A. & Bazzaz, F.A. (2004) Phylogenetic overdispersion in Floridian oak communities. *The American Naturalist*, **163**, 823–843.

Cavender-Bares, J., Keen, A. & Miles, B. (2006) Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology*, **87**, S109–S122.

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

Cooper, N., Rodriguez, J. & Purvis, A. (2008) A common tendency for phylogenetic overdispersion in mammalian assemblages. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 2031–2037.

Den Boer, P.J. (1979) Exclusion or coexistence and the taxonomic or ecological relationship between species. *Netherlands Journal of Zoology*, **30**, 278–306.

Elton, C. (1946) Competition and the structure of ecological communities. *Journal of Animal Ecology*, **15**, 54–68.

Emerson, B.C. & Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution*, **23**, 619–630.

Fritz, S.A., Bininda-Emonds, O.R.P. & Purvis, A. (2009) Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters*, **12**, 538–549.

Gerhold, P., Cahill, J.F., Winter, M., Bartish, I.V. & Prinzing, A. (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, **29**, 600–614.

Gómez, J.P., Bravo, G.A., Brumfield, R.T., Tello, J.G. & Cadena, C.D. (2010) A phylogenetic approach to disentangling the role of competition and habitat filtering in community assembly of Neotropical forest birds. *Journal of Animal Ecology*, **79**, 1181–1192.

Gotelli, N.J. (2000) Null model analysis of species co-occurrence patterns. *Ecology*, **81**, 2606–2621.

Gotelli, N.J., Graves, G.R. & Rahbek, C. (2010) Macroecological signals of species interactions in the Danish avifauna. *Proceedings of the National Academy of Sciences USA*, **107**, 5030–5035.

Graham, C.H., Parra, J.L., Rahbek, C. & McGuire, J.A. (2009) Phylogenetic structure in tropical hummingbird communities. *Proceedings of the National Academy of Sciences USA*, **106**, 19673–19678.

Hardin, G. (1960) The competitive exclusion principle. *Science*, **131**, 1292–1297.

Hardy, O.J. (2008) Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *Journal of Ecology*, **96**, 914–926.

Hardy, O.J. & Senterre, B. (2007) Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. *Journal of Ecology*, **95**, 493–506.

Hosmer, D.W. & Lemeshow, S. (2004) *Applied logistic regression*. John Wiley & Sons, Hoboken, NJ.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, NJ.

Hut, R.A., Kronfeld-Schor, N., Van Der Vinne, V. & De La Iglesia, H. (2012) In search of a temporal niche: environmental factors. *Progress in Brain Research,* **199**, 281–304.

Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. (2012) The global diversity of birds in space and time. *Nature*, **491**, 444–448.

Keddy, P. & Weiher, E. (1999) Introduction: the scope and goals of research on assembly rules. *Ecological assembly rules: perspectives, advances, retreats* (ed. by E. Weiher and P. Keddy), pp. 1–22. Cambridge University Press, Cambridge.

Kembel, S.W. (2009) Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. *Ecology Letters*, **12**, 949–960.

Kembel, S.W. & Hubbell, S.P. (2006) The phylogenetic structure of a Neotropical forest tree community. *Ecology*, **87**, S86–S99.

Kneitel, J.M. & Chase, J.M. (2004) Trade-offs in community ecology: linking spatial scales and species coexistence. *Ecology Letters*, **7**, 69–80.

Kooyman, R., Rossetto, M., Cornwell, W. & Westoby, M. (2011) Phylogenetic tests of community assembly across regional to continental scales in tropical and subtropical rain forests. *Global Ecology and Biogeography*, **20**, 707–716.

Losos, J.B. & Schluter, D. (2000) Analysis of an evolutionary species–area relationship. *Nature*, **408**, 847–850.

Losos, J.B., Leal, M., Glor, R.E. *et al.* (2003) Niche lability in the evolution of a Caribbean lizard community. *Nature*, **424**, 542–545.

Lovette, I.J. & Hochachka, W.M. (2006) Simultaneous effects of phylogenetic niche conservatism and competition on avian community structure. *Ecology*, **87**, S14–S28.

Mayfield, M.M. & Levine, J.M. (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, **13**, 1085–1093.

Phillimore, A.B. & Price, T.D. (2008) Density-dependent cladogenesis in birds. *PLoS Biology*, **6**, e71.

Purvis, A., Cardillo, M., Grenyer, R. & Collen, B. (2005) Correlates of extinction risk: phylogeny, biology, threat and scale. *Phylogeny and conservation* (ed. by A. Purvis, J.L. Gittleman and T. Brooks), pp. 295–316. Cambridge University Press, Cambridge.

Pyron, R.A. & Wiens, J.J. (2011) A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution*, **61**, 543–583.

R Development Core Team (2013) *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/.

Ricklefs, R.E. (1987) Community diversity: relative roles of local and regional processes. *Science*, **235**, 167–171.

Ricklefs, R.E. (2015) Intrinsic dynamics of the regional community. *Ecology Letters*, **18**, 497–503.

Rosenzweig, M.L. (2001) Loss of speciation rate will impoverish future diversity. *Proceedings of the National Academy of Sciences USA*, **98**, 5404–5410.

Schoener, T.W. (1970) Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology*, **51**, 408–418.

Strona, G., Nappo, D., Boccacci, F., Fattorini, S. & San-Miguel-Ayanz, J. (2014) A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications*, **5**, art. no. 4114.

Sunday, J.M., Bates, A.E. & Dulvy, N.K. (2012) Thermal tolerance and the global redistribution of animals. *Nature Climate Change*, **2**, 686–690.

Vamosi, S.M., Heard, S.B., Vamosi, J.C. & Webb, C.O. (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, **18**, 572–592.

Villalobos, F., Rangel, T.F. & Diniz-Filho, J. (2013) Phylogenetic fields of species: cross-species patterns of phylogenetic structure and geographical coexistence. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20122570.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.

Weir, J.T. & Schluter, D. (2007) The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science*, **315**, 1574–1576.

Xie, Y., Zhang, S. & Wang, W. (2009) *Biodiversity atlas of China.* Hunan Education Press, Changsha.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1 The proportion of phylogenetic patterns (using a standardized version of the checkerboard score).
Figure S2 Hypothesis on the probability of phylogenetic overdispersion shifting to clustering along a continuum of spatial scales.
Table S1 The number of species used for measuring phylogenetic structures at different taxonomic levels.
Table S2 Correlations between phylogenetic distance and two species co-occurrence indices (CI and DOI) at the class level.
Table S3 The best-fitting models.
Table S4 Summary of the proportion of phylogenetic patterns pooled from the four classes at various taxonomic and spatial scales.
Table S5 Summary of the phylogenetic patterns observed at various classes, taxonomic (class, order, family and genus) levels and spatial scales.
Table S6 Summary of proportion of phylogenetic patterns pooled from order, family and genus at the four classes and spatial scales.

BIOSKETCH

The research was conducted by the co-authors with a shared interest in animal community assembly. Z.Z. conceived the idea, C.Y. managed data and conducted analyses, and both took responsibility for the writing. X.Y. provided the original data sources. X.L. provided support for data management. M.H. provided valuable comments and writing support.

Editor: Adam Algar