Molecular Phylogenetics and Evolution 135 (2019) 222–229

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogenetic delineation of regional biota: A case study of the Chinese flora

Jianfei Ye[a,b,c], Limin Lu[a], Bing Liu[a,d], Tuo Yang[a], Jinlong Zhang[e], Haihua Hu[a,b], Rong Li[f], Anming Lu[a], Huiyuan Liu[a], Lingfeng Mao[g,*], Zhiduan Chen[a,d,*]

[a] State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] Beijing Botanical Garden, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[d] Sino-African Joint Research Centre, Chinese Academy of Sciences, Wuhan 430074, China
[e] Flora Conservation Department, Kadoorie Farm and Botanic Garden, Lam Kam Road, Tai Po, New Territories, Hong Kong, China
[f] Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China
[g] College of Biology and the Environment, Nanjing Forestry University, Nanjing 210037, China

ARTICLE INFO

Keywords:
Angiosperms
Chinese flora
Biogeographical regionalization
Phylogenetic beta diversity
Spatial turnover

ABSTRACT

Biogeographical regionalization schemes have traditionally been constructed based on taxonomic endemism of families, genera, and/or species, and rarely incorporated the phylogenetic relationships between taxa. However, phylogenetic relationships are important for understanding historical connections within and among biogeographical regions. Phylogeny-based delineation of biota is a burgeoning and fruitful field that is expected to provide novel insights into the conservation of regional diversity and the evolutionary history of biota. Using the Chinese flora as an example, we compared regionalization schemes that were based on: (1) taxonomic endemism, (2) taxonomic dissimilarity, and (3) phylogenetic dissimilarity. Our results revealed general consistency among different regionalization schemes and demonstrated that the phylogenetic dissimilarity approach is preferable for biogeographical regionalization studies. Using the phylogenetic dissimilarity approach, we identified five phytogeographical regions within China: the Paleotropic, Holarctic, East Asiatic, Tethyan, and Qinghai–Tibet Plateau Regions. The relationship of these regions was inferred to be: (Paleotropic, ((East Asiatic + Holarctic) + (Tethyan + Qinghai–Tibet Plateau)).

1. Introduction

Biogeographical regionalizations show how the diversity of life on Earth is geographically organized, reflect how physical and biological forces have shaped biotic distributions (Kreft and Jetz, 2010), and can provide a spatially explicit framework for studying large-scale biodiversity patterns (Crisp et al., 2009). Wallace's (1876) map of global terrestrial zoogeographical regions laid the foundation for modern biogeography (Ladle and Whittaker, 2011). Since then, various biomes and biogeographical regionalization schemes have been proposed (Takhtajan, 1978; Wu 1979; Thorne, 1987; Holt et al., 2013; Zhang et al., 2016b; Slik et al., 2018).

Traditionally, biogeographical regionalization was usually proposed using the taxonomic endemism approach (Fig. 1), which is based on qualitative evidence, such as the geographical ranges of important taxa, especially endemic taxa (Wallace, 1876; Handel-Mazzettii, 1931; Engler and Diels, 1936; Takhtajan, 1978; Wu, 1979; Wu et al., 2010). Quantitative approaches based on taxonomic (e.g., species, generic, and familial) dissimilarity have recently been used for biogeographical regionalization (Fig. 1) (González-Orozco et al., 2014; Zhang et al., 2016b; He et al., 2017). By using distribution records and statistical approaches, biogeographical studies are more quantitative rather than qualitative. Instead of relying on experts' experience and knowledge, regionalization based on quantitative approaches are much more objective and reproducible. The availability of distributional data, novel multivariate statistical techniques, and enhanced computational power has resulted in numerous large-scale quantitative biogeographical regionalization analyses (Table 1, see Kreft and Jetz 2010). Kreft and Jetz (2010) provided a general framework for biogeographical regionalization based on species distributions. Studies based on taxonomic dissimilarity generally assume that all members of a taxonomic level, such as species, are identical and non-hierarchical, whereas phylogeny-based

* Corresponding authors at: State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, 20 Nanxincun, Xiangshan, Beijing 100093, China. (Z. Chen). College of Biology and the Environment, Nanjing Forestry University, 159 Longpan Road, Nanjing 210037, China. (L. Mao).
E-mail addresses: maolingfeng2008@163.com (L. Mao), zhiduan@ibcas.ac.cn (Z. Chen).

https://doi.org/10.1016/j.ympev.2019.03.011
Received 4 February 2019; Received in revised form 16 March 2019; Accepted 18 March 2019
Available online 22 March 2019

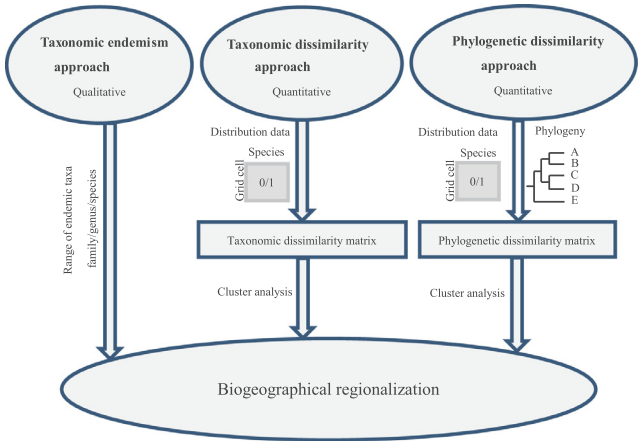

Fig. 1. Comparison of the taxonomic endemism approach, taxonomic dissimilarity approach, and phylogenetic dissimilarity approach to infer biogeographical regionalization.

regionalization incorporates phylogenetic relationships in analyses (Holt et al., 2013; Daru et al., 2016; Li et al., 2015; Slik et al., 2018). The phylogenetic relatedness among taxa in different areas can be quantified based on phylogenetic trees (Graham and Fine, 2008). Additionally, phylogenetic trees have been used to detect the turnover of phylogenetic patterns, which reflects the relative importance of evolutionary and ecological factors in shaping current diversity patterns across broad spatial scales (Graham and Fine, 2008). Several studies have inferred biogeographical regionalization incorporating phylogenies at different geographical scales and within different taxa (Holt et al., 2013; Kubota et al., 2014; Hattab et al. 2015; Jønsson and Holt, 2015; Li et al., 2015; Daru et al., 2016, 2017a, 2017b; Slik et al., 2018). Phylogeny-based regionalizations have also provided rigorous and objective classifications of biota, and enhanced our knowledge of biodiversity (Holt et al., 2013).

However, sampling density and resolution of the phylogeny are important factors that affect the resolution of biogeographical regionalization. For example, phylogenetic regionalization based on woody plants does not represent patterns for the entire flora (Kubota et al., 2014; Daru et al., 2016; Slik et al., 2018). Phylogenetic regionalization based on family-level phylogenies may be useful for rough or preliminary phylogenetic regionalization, but phylogenies at the generic or species level are required for finer-scale analyses, especially for flora with a complex evolutionary history, such as the Chinese flora. To date, most of the phylogenies for higher plants (i.e., Embryophyta) used in phytogeographical regionalization were constructed using the software Phylomatic (Webb and Donoghue, 2005) based on the Angiosperm Phylogeny Group classification system; however, these phylogenies could not sufficiently reflect the phylogenetic relatedness within each family. For a whole regional flora, including both woody and herbaceous taxa, phytogeographical regionalization based on higher-resolution phylogenies at the generic or species level has rarely been explored (Forest et al., 2007). Therefore, it is important to explore the feasibility of using phylogenetic methods for phytogeographical regionalization by conducting a suitable case study for a region with available generic- or species-level phylogenies.

Several phytogeographical regionalization schemes have been proposed for China (e.g., Handel-Mazzettii, 1931; Wu, 1979; Wu et al., 2010; Sun and Wu, 2013; Zhang et al., 2016b), of which the most comprehensive study was carried out by Wu et al. (2010) (see Sun and Wu, 2013 for the revised English version). Wu et al. (2010) divided the Chinese flora into four kingdoms, seven subkingdoms, 24 regions, and 49 subregions mainly based on the distribution of endemic families, genera, and species, respectively. Since then, this regionalization scheme has been widely used for biogeographical and biodiversity

Table 1
Summary of the five phytogeographical regions of the Chinese flora.

Phytogeographical region	Area in km^2 (number of grid cells)	Richness (family/genus/species)	Mean/minimum/maximum elevation (m) and standard deviation	Mean/minimum/maximum annual temperature (°C) and standard deviation	Mean/minimum/maximum annual precipitation (mm) and standard deviation	Floristic distinctiveness (FD)
Paleotropic region	210,492 (24)	217/1605/7134	173.83/29/400/96.55	21.95/19.87/24.29/1.18	1643.35/1328.18/2008.10/194.57	0.21
Holarctic region	2,120,962 (218)	170/1254/6878	731.96/6/1954/490.10	4.1/−6.09/13.70/4.39	497.21/166.20/1030.02/147.54	0.11
East Asiatic region	2,349,449 (239)	233/2304/20317	616.29/2/3527/599.16	16.0/3.55/22.13/2.34	1221.85/516.60/2679.93/356.02	0.05
Tethyan region	2,165,458 (222)	128/912/5412	1810.55/322/5442/1176.35	5.62/−8.60/12.74/5.36	111.20/17.46/450.62/91.31	0.15
QTP region	2,363,729 (240)	205/1896/16220	4247.64/1203/5324/920.67	0.34/−7.62/18.56/5.28	435.32/24.14/2583.46/357.33	0.08

studies throughout China (e.g., Huang et al., 2016; Xu et al., 2017). Chen et al. (2016) reconstructed a well-resolved generic phylogeny for Chinese vascular plants using four chloroplast genes (*atp*B, *mat*K, *ndh*F, and *rbc*L) and one mitochondrial gene (*mat*R). As the distribution records for each species at the county level are also available (Lu et al., 2018), the Chinese flora provides an ideal model for testing phylogenetic regionalization.

In this study, we compared the qualitative and quantitative approaches of phytogeographical regionalization using species distribution records and a phylogenetic tree at the genus level. We aimed to provide a work flow for phytogeographical regionalization and help elucidate the formation and recognition of phytogeographical regions in China. Specific questions included: (1) what are the significance and extent of differences among regionalization schemes based on endemism, taxonomic dissimilarity, and phylogeny, and (2) do phylogenetic dissimilarity approaches have merit for phytogeographical regionalization?

2. Materials and methods

2.1. Study area

As one of the most biodiverse countries worldwide (Mittermeier et al., 1997), China has more than 29,000 angiosperm species, which accounts for nearly 10% of all flowering plants on Earth (Huang et al., 2013). China has a broad latitudinal range, from 18°N in the south to 53.5°N in the north, and covers over 9,600,000 km^2 (Appendix S1: Fig. S1). Its territory encompasses a variety of landscapes, including seashores; tropical, subtropical, and temperate forests; grasslands; deserts; and the extremely elevated Qinghai–Tibet (Xizang) Plateau (QTP).

2.2. Distribution records

The species distribution data were exhaustively compiled from the national and regional flora records, checklists, and herbarium records. Each entry was assigned to a county (there are 2377 counties in China) with an average land area of approximately 4000 km^2. All the names were standardized following *Flora of China* (Wu et al., 1994–2013). Alien and non-angiosperm species were excluded, which left a total of 1,517,128 entries that represented 27,515 species and 2845 genera. Even though our species distribution database is mostly complete, some small geographical gaps in specimen collection still exist and some historical boundaries have been changed during the last few decades. Thus, to reduce sampling bias caused by unequal sampling areas, we up-scaled and assigned the distribution records to 100-km × 100-km grid cells. We assigned the species to each grid cell in all counties that overlapped with each grid cell. Grid cells were excluded if land area occupied less than 50% of the entire area of the cell (i.e., < 5,000 km^2). Following these steps, we generated a genus-level presence–absence matrix for the remaining 943 grid cells.

2.3. Phylogenetic tree

We used the phylogeny reconstructed by Chen et al. (2016); that included 6096 species and 3114 genera in 323 families were included in this phylogeny, which covered more than 93% of vascular plant genera native to China. After excluding 523 genera not found in the distribution records, the phylogeny used in the subsequent analysis included 2591 angiosperm genera (ca. 90% of angiosperm genera in China). Divergence time estimation was conducted using penalized likelihood in treePL (Smith and O'Meara, 2012); calibration details, including parameters, follow Lu et al. (2018).

2.4. Beta diversity metrics

To quantify spatial turnover of floristic composition, we calculated taxonomic turnover between each pair of grid cells. We used the turnover component of the Sørensen dissimilarity index (β_{sim}; Baselga, 2012) to measure taxonomic turnover. β_{sim} is calculated as:

$$\beta_{sim} = 1 - \frac{a}{\min(b,c) + a}$$

where 'a' is the number of shared genera between two grid cells, and 'b' and 'c' are the numbers of genera unique to each grid cell. β_{sim} ranges from 0 (if generic composition is identical between grid cells) to 1 (if there are no shared taxa). Importantly, β_{sim} is independent of differences in species richness among grid cells (Baselga, 2012; Leprieur and Oikonomou, 2014) and therefore provides an unbiased estimation of compositional turnover.

Phylogenetic turnover ($p\beta_{sim}$) was quantified using phylogenetic beta diversity and is adapted from β_{sim}; the proportion of shared phylogenetic branch lengths of the dated phylogenetic tree between grid cells is substituted for the proportion of shared genera (Graham and Fine, 2008). $P\beta_{sim}$ ranges from 0 (if genera are identical and share the same branch lengths) to 1 (if genera do not share similar branch lengths). We then generated pairwise distance matrices for $p\beta_{sim}$ and β_{sim} between all grid cells. To evaluate the sensitivity of our results, we also compared the β_{sim} and $p\beta_{sim}$ values with the turnover component obtained from the Jaccard's dissimilarity index (β_{jtu} and $p\beta_{jtu}$; Baselga, 2012).

Because $p\beta_{sim}$ is likely to be related to β_{sim}, we used a null model to test if grid cell assemblages were more or less phylogenetically similar than expected based on the number of taxa. A null distribution was generated for each grid cell by randomizing the tips of the phylogeny 999 times; this randomized the phylogenetic turnover between grid cells while preserving the underlying taxonomic turnover differences between grid cells. In each iteration, $p\beta_{sim}$ was calculated for each compared pair of grid cells. Null distributions were used to calculate standardized effect sizes (SES) of $p\beta_{sim}$ ($SES.p\beta_{sim}$), where the mean of the null distribution was subtracted from the observed $p\beta_{sim}$ and divided by the standard deviation of the null distribution. A positive SES indicates higher phylogenetic turnover than expected based on random sampling of the regional flora; a negative SES indicates the opposite.

Analysis was performed in R 3.1.0 (R Core Team, 2016) using the 'betapart' package (Baselga and Orme, 2012). The average β_{sim}, $p\beta_{sim}$, and $SES.\beta_{sim}$ (*P* lacking) values between the focal grid cell and all other grid cells were mapped to visualize the distinctiveness of each grid cell assemblage with respect to the other. The floristic distinctiveness (FD) of phytogeographical regions is equal to the mean $p\beta_{sim}$ between the focal region and all other regions (Holt et al., 2013).

2.5. Cluster analysis

To identify spatial clusters, we compared the output from eight hierarchical clustering algorithms based on the β_{sim} and $p\beta_{sim}$ pairwise distance matrices: (1) single linkage (SL), (2) complete linkage (CL), (3) unweighted pair-group method using arithmetic averages (UPGMA), (4) unweighted pair-group method using centroids (UPGMC), (5) weighted pair-group method using arithmetic averages (WPGMA), (6) weighted pair-group method using centroids (WPGMC), and (7) and (8) two variants of Ward's minimum variance (ward.D and ward.D2). To assess the degree of data distortion in these models, we used Sokal and Rohlf's (1962) cophenetic correlation coefficient and Gower's (1983) distance. The cophenetic correlation defines the relationship between the terminals of a dendrogram with the original distance matrix for all clustering methods, and has a value between 0 and 1 (poor and strong correlation, respectively). Gower's (1983) distance measures the sum of squared differences between original and cophenetic distances. A cluster method that produces the minimum Gower distance is considered a good clustering model for the distance matrix (Borcard et al., 2011).

A reasonable number of clusters for the phytogeographical regions

was determined by applying the Kelley–Gardner–Sutcliffe penalty function; this function maximizes the differences between groups and the cohesiveness within groups, and its minimum value corresponds to the optimal number of clusters (Kelley et al., 1996). The Kelley–-Gardner–Sutcliffe penalty function was implemented in the 'maptree' package (White and Gramacy, 2012) in R.

2.6. Ordination analysis

To illustrate the spatial turnover and relationships between the clusters in an alternative way, all grid cells were also plotted in a two-dimensional ordination plot using the neighbor-joining algorithm by non-metric multidimensional scaling (NMDS). NMDS is considered the most robust unconstrained method for reducing dimensions (Minchin, 1987). The degree of NMDS ordination is characterized by the stress value (from 0 to 1) and should be minimized. We performed NMDS ordination using the β_{sim} and $p\beta_{sim}$ pairwise distance matrices for clusters that represented phytogeographical regions in the 'vegan' package (Oksanen et al., 2016). We used 100 random starts to find a stable solution and avoid local minima. The ordination was rotated and rescaled to facilitate comparison (Kreft and Jetz, 2010).

3. Results

3.1. Beta diversity

β_{jtu} showed a strong correlation with β_{sim} (Mantel test; $r = 0.99$, $P < 0.001$, 999 permutations). Similarly, there was a strong correlation between $p\beta_{jtu}$ and $p\beta_{sim}$ (Mantel test; $r = 0.99$, $P < 0.001$, 999 permutations). Here, we focused on the results of the turnover component of Sørensen index (β_{sim} and $p\beta_{sim}$); $p\beta_{sim}$ was strongly correlated with β_{sim} (Mantel test; $r = 0.96$, $P < 0.001$, 999 permutations). However, SES.$p\beta_{sim}$ was moderately correlated with β_{sim} (Mantel test; $r = 0.53$, $P < 0.001$, 999 permutations). The average taxonomic turnover, phylogenetic turnover, and SES of phylogenetic turnover for the focal grid cell and all other 942 grid cells are shown in Fig. 2, and a close correlation between phylogenetic and taxonomic heterogeneity was detected. The highest taxonomic and phylogenetic turnover occurred in the southeastern coast of China and Hainan Island, and southwestern Yunnan also had the highest taxonomic turnover. The lowest taxonomic and phylogenetic turnover occurred in the Hengduan Mountains, Qin Mountains, Taihang Mountains, and Yan Mountains (Fig. 2a, b). The null model test revealed that southernmost China exhibited the highest regional phylogenetic turnover. The lowest phylogenetic turnover detected by the null model test occurred in central QTP, the Qin Mountains, the Taihang Mountains, and the Yan Mountains (Fig. 2c). Among the 943 grid cells, the SES were positive in 846 grid cells and negative in 97 grid cells, and SES values in 63 grid cells significantly differed from the null expectation (i.e., SES values > 1.96) (Fig. 2c).

3.2. Phytogeographical regionalization

Among the eight clustering algorithms, UPGMA consistently provided the best fit between the dendrogram and the original distance matrices for β_{sim} and $p\beta_{sim}$, whereas SL and ward.D consistently performed the worst (Appendix S1: Table S2). Therefore, we adopted the UPGMA method to define the phytogeographical regions.

The UPGMA clustering of grid cell assemblages based on both β_{sim} and $p\beta_{sim}$ dissimilarity matrices yielded five major phytogeographical regions in China (Fig. 3a and 4a). To maintain the stability of names, we labeled each region based on the four kingdom names from Wu et al. (2010) with which it most coincided: (1) Holarctic, (2) Tethyan, (3) East Asiatic, or (4) Paleotropic regions. The fifth region was called the QTP region. We delineated the boundaries for the five regions in our phylogenetic regionalization scheme and highlighted the characteristics

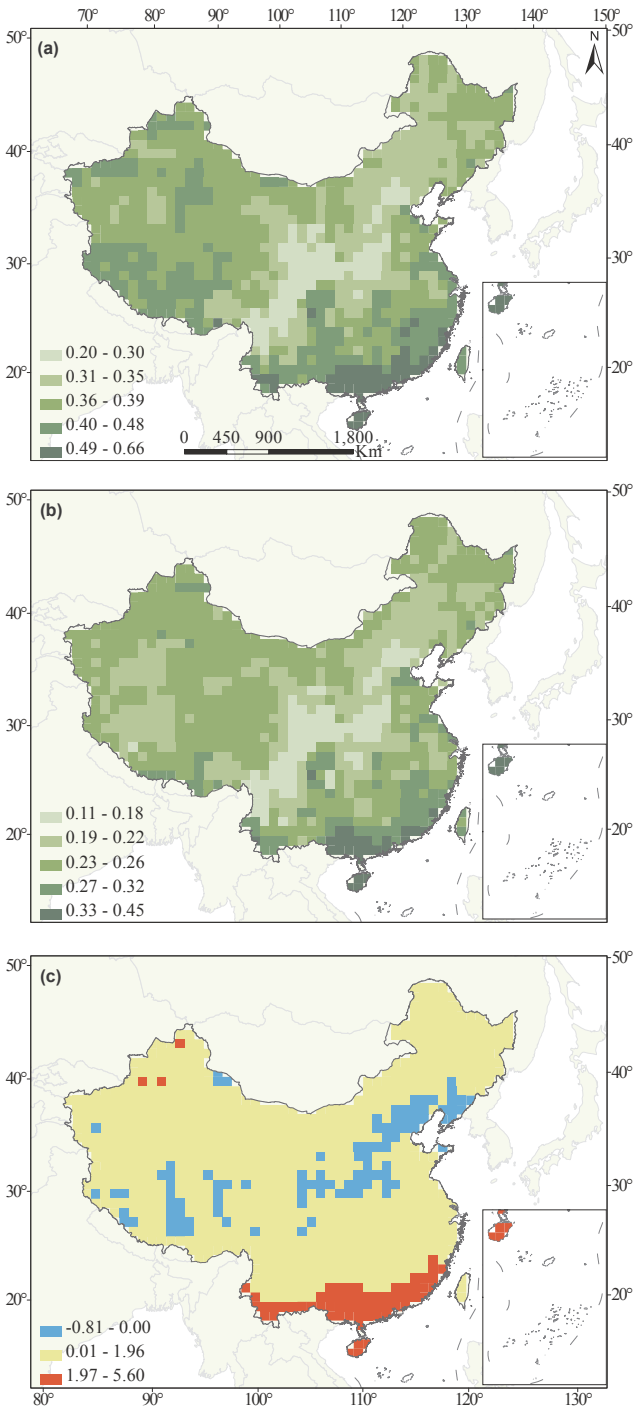

Fig. 2. Patterns of spatial turnover of the Chinese flora. Numbers indicate the average values of β_{sim} (a), $p\beta_{sim}$ (b), and SES.$p\beta_{sim}$ (c) for the focal grid cell and all other grid cells. The maps were generated using ArcGIS 10.0 in Albers projection.

of each region (Appendix S1: S1).

The two regionalization schemes derived from taxonomic and phylogenetic approaches in this study produced broadly similar results, although there was a noticeable difference in the size of the East Asiatic region (Fig. 5). In the phylogenetic regionalization scheme, the East Asiatic region was larger, because the boundary between the East Asiatic region and the Holarctic region was extended northward, and the Paleotropic region extended southward. In the dendrogram derived from the β_{sim} dissimilarity matrix (Fig. 3b), grid cells located in eastern China were grouped into three clusters (the Paleotropic, East Asiatic,

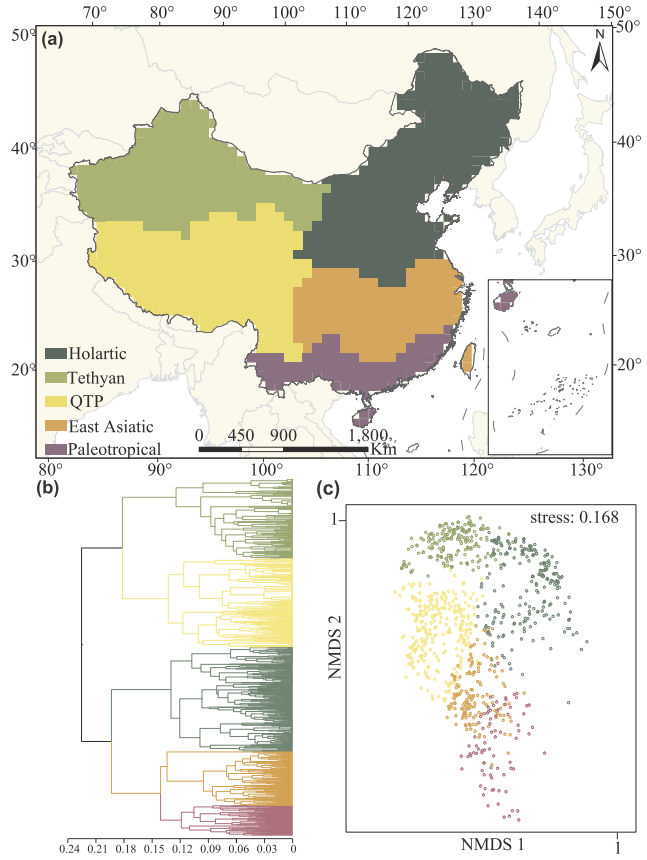

Fig. 3. Map (a) and dendrogram (b) based on UPGMA hierarchical clustering and NMDS ordination of grid cell assemblages based on the β_{sim} dissimilarity matrix (c). Five phytogeographical regions are highlighted in the dendrogram and displayed in the map, and are represented by the same colors in the NMDS ordination. The map was generated using ArcGIS 10.0 in Albers projection.

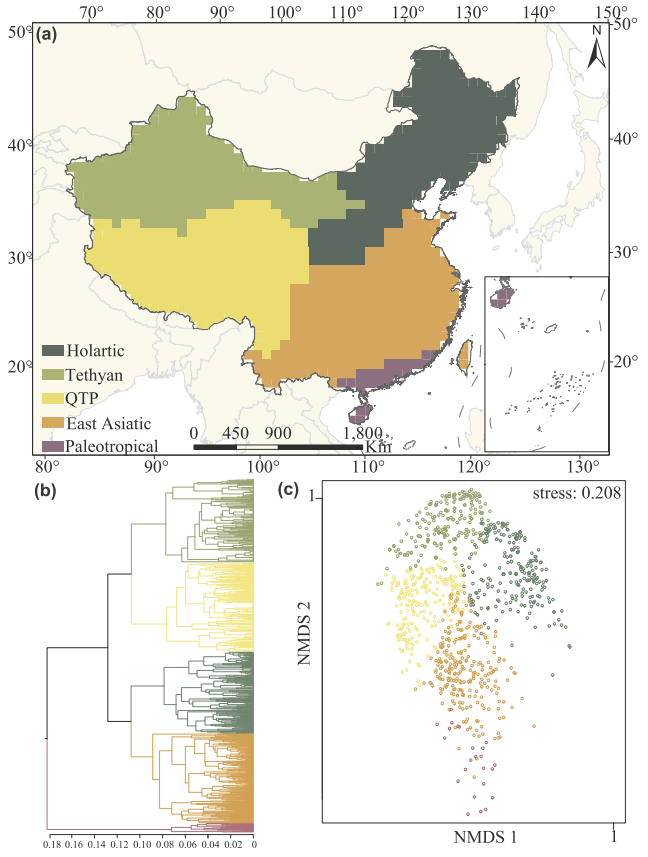

Fig. 4. Map (a) and dendrogram (b) resulting from UPGMA hierarchical clustering and NMDS ordination of grid cell assemblages based on the $p\beta_{sim}$ dissimilarity matrix (c). The five phytogeographical regions are highlighted in the dendrogram and displayed in the map, and are represented by the same colors in the NMDS ordination. The map was generated using ArcGIS 10.0 in Albers projection.

and Holarctic regions) and the other grid cells in western China grouped into two clusters (Tethyan and QTP regions). In the dendrogram based on the $p\beta_{sim}$ dissimilarity matrix (Fig. 4b), grid cells located in southeastern China were grouped into the Paleotropic region, which formed the Southern cluster, and the other grid cells were grouped into four regions, which formed the Northern cluster. The Northern cluster was further split into two major subclusters, with the East Asiatic and the Holarctic regions forming an Eastern subcluster, and the QTP and Tethyan regions forming a Western subcluster.

The stress values were relatively low (β_{sim} matrix, 0.168; $p\beta_{sim}$: 0.208), which indicated relatively good projection of dissimilarity matrices. However, the NMDS ordination results showed continuous rather than discontinuous phytogeographical transitions between the clustered regions (Figs. 3c, 4c). We also investigated the FD of the five phylogeny-based phytogeographical regions, and found that the Paleotropic region had the highest FD (FD = 0.20). In contrast, the East Asiatic region had the lowest FD (FD = 0.05; Table 1). Moreover, the Paleotropic region had a significantly higher phylogenetic turnover than expected (SES.$p\beta_{sim}$ > 1.96).

4. Discussion

In this study, we used a molecular phylogenetic tree for angiosperm genera and their distribution records to examine spatial phylogenetic turnover and identified five phytogeographical regions in China (Fig. 3a). Among these regions, the Paleotropic region had the most evolutionarily distinct flora, with genera that had significantly higher phylogenetic turnover than expected (SES.$p\beta_{sim}$ > 1.96, Fig. 2c). The East Asiatic region has preserved boreotropical taxa that were once common in the Northern Hemisphere but went extinct from other areas during the Quaternary glaciation (Wolfe, 1975; Wen, 1999; López-Pujol and Ren, 2010). The Holarctic region has Boreal-Tertiary flora, and it is closely related to the East Asiatic region. The Tethyan region has Tethys Sea flora, which is evolutionarily distinct from the floras of the Holarctic and East Asiatic regions, and is more closely related to the QTP region flora. The QTP region is the largest phytogeographical region in China and has one of the world's richest floras, with numerous endemic species (Huang et al., 2016).

For the taxonomic endemism approach, phytogeographical regionalization schemes are mainly based on the geographic ranges of endemic taxa. The evidence and criteria used to divide regions are qualitative, and sometimes heavily rely on experts' opinions (Zhang et al., 2016a). The number of recognized regions and their boundaries often vary among researchers and are therefore inconsistent (Handel-Mazzettii, 1931; Engler and Diels, 1936; Takhtajan, 1978; Wu, 1979; Wu et al., 2010). In contrast, our regionalization based on clustering algorithms that used β_{sim} and $p\beta_{sim}$ could reveal the optimal number of regions and yield clear boundaries. We herein compared the regionalization schemes obtained based on taxonomic dissimilarity and phylogenetic dissimilarity matrices with those obtained by the taxonomic endemism approach. We discuss the merits of using a phylogenetic approach, which include obtaining an optimal regionalization scheme with clear regional boundaries and elucidating the evolutionary history of flora.

4.1. Comparison between taxonomic endemism and phylogenetic dissimilarity approach

Our regionalization based on β_{sim} is quite different from previous endemism-based regionalization schemes. The Paleotropic region in endemism-based regionalization schemes only includes the southernmost part of China; however, in our regionalization based on β_{sim}, this region had a wider range (Appendix S1: Fig. S4b and 4c). Compared with Wu et al.'s (2010) endemism-based East Asiatic Kingdom, the East Asiatic region based on β_{sim} was smaller and only included the eastern subtropical forest region. The Holarctic, East Asiatic, and Paleotropic regions clustered together in the β_{sim} dendrogram (Fig. 3), whereas the Paleotropic Kingdom was sister to other regions in Wu et al.'s (2010) endemism-based regionalization.

Despite these differences, the regionalization schemes based on taxonomic endemism and phylogenetic dissimilarity were comparable. Our two major clusters in the $p\beta_{sim}$ dendrogram (Southern and Northern) were generally consistent with the divisions described in previous studies. The Chinese flora was divided into two kingdoms in previous studies: the Paleotropic Kingdom, which includes the southernmost part of China, and the Holarctic Kingdom, which includes the rest of China (Engler and Diels, 1936; Takhtajan, 1978; Wu, 1979). However, the boundary of the two major parts of the Chinese flora in our phylogenetic regionalization scheme substantially differed from that of a traditional endemism-based scheme (Appendix S1: Fig. S4).

Our five regions of the Chinese flora delineated by the phylogenetic dissimilarity approach were also comparable to the four kingdoms of Wu et al.'s (2010) regionalization scheme (Appendix S1: Fig. S4b), although there were three major differences. First, Wu et al.'s (2010) East Asiatic Kingdom (which included three subkingdoms) did not cluster together in our dendrogram (Fig. 4a, b); the western part of Wu et al.'s (2010) East Asiatic Kingdom (IIIE and IIIF) roughly corresponded to our QTP region, whereas its eastern part (IIID) almost included our Paleotropic and East Asiatic regions, and the northern boundary extended into the Holarctic region (Appendix S1: Fig. S4b and S4e). Second, southeastern Xizang, southwestern Yunnan, and Taiwan were placed in the Paleotropic Kingdom in Wu et al.'s (2010) scheme, but were placed in the East Asiatic region in our scheme (Appendix S1: Fig. S4b and S4e). Third, Wu et al.'s (2010) Eurasian Forest Subkingdom (IA) was divided into two clusters (Tethyan and Holarctic regions) in our clustering dendrogram (Appendix S1: Fig. S4b and S4e). The phylogenetic dissimilarity approach can be used to test the naturalness of Wu et al.'s (2010) four kingdoms and their relationships (Ebach and Parenti, 2015). Using the phylogenetic approach allowed us to provide deeper insight to bioregionalization. Our results show that Wu et al.'s (2010) Chinese phytogeographical regionalization scheme could be improved by considering the phylogenetic relationships.

4.2. Comparison between taxonomic dissimilarity and phylogenetic dissimilarity approach

Like the phylogenetic dissimilarity approach, the taxonomic dissimilarity approach is quantitative, and the results of regionalization are reproducible and testable. In this study, both of the approaches we used yielded five regions. The shapes and boundaries of the five regions are broadly similar, which probably resulted from the high correlation ($r = 0.96$) between β_{sim} and $p\beta_{sim}$. However, the range of regions and inter-regional relationships in the taxonomic and phylogenetic regionalization schemes differed. These differences demonstrated that the phylogenetic dissimilarity approach captured the evolutionary history of the Chinese flora.

The East Asiatic/Holarctic boundaries differed between the regionalization schemes produced by the taxonomic and phylogenetic dissimilarity approaches. The East Asiatic/Holarctic boundaries in our phylogenetic regionalization extended to Henan and Shandong Provinces (Fig. 5; Appendix S1: Fig. S1), in which there are some early-

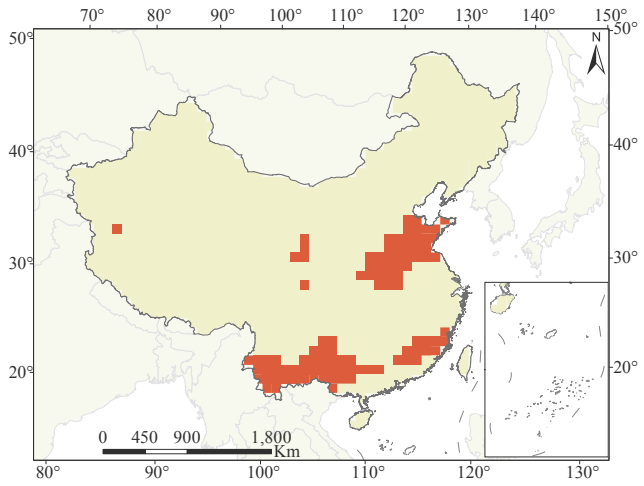

Fig. 5. Differences between our taxonomic and phylogenetic dissimilarity-based regionalization schemes. Red cells indicate areas where the two schemes did not match. The maps were generated using ArcGIS 10.0 in Albers projection. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

diverging lineages of basal angiosperms and basal core eudicots (e.g., Calycanthaceae, Lardizabalaceae, Lauraceae, Magnoliaceae, and Menispermaceae). The East Asiatic/Holarctic boundaries differed between taxonomic and phylogenetic approaches, probably because of the presence–absence of old lineages; the old lineages mainly occur in the subtropical evergreen forest and are characteristic taxa of the East Asiatic region. The phylogenetic regionalization extended the boundary of the East Asiatic/Holarctic region northward, and areas that share an evolutionary history with the East Asiatic region were included. Another noticeable difference between the two approaches was the location of the boundary between the East Asiatic and Paleotropic regions (Fig. 5). Yunnan, Guizhou, and Guangxi were placed in the East Asiatic region in the phylogenetic regionalization, but in the Paleotropic region in the taxonomic regionalization. Because these areas have more endemic, particularly paleoendemic genera (Ying et al., 1994), they are thought to have been a refuge during the Pleistocene glacial period (López-Pujol et al., 2011) and have closer floristic affinity with the East Asiatic region than with the Paleotropic region (Wen, 1999; López-Pujol and Ren, 2010). Therefore, the phylogenetic regionalization included more areas from north and southwest China in the East Asiatic region compared with the taxonomic regionalization. These results clearly reflect the evolutionary history of the flora.

We investigated the evolutionary distinctiveness of floristic regions and found that the Paleotropic region had the highest mean $p\beta_{sim}$ (mean $p\beta_{sim}$ between Paleotropic and all other floristic regions = 0.205; Table 1), which emphasizes the high evolutionary uniqueness and independence of the Paleotropic region. The null model test further supported that the Paleotropic region had significantly higher phylogenetic turnover than expected ($SES.p\beta_{sim} > 1.96$). The regional relationships in the phylogenetic regionalization scheme can be explained by the tropical niche conservatism hypothesis (TCH). The TCH suggests that many angiosperm clades originated and diversified in tropical areas when the Earth's climate was warm and wet, and the dispersal of lineages from tropical to temperate (or subtropical) regions was limited by the lineages' tolerance to cold (Wiens et al., 2006; Kerkhoff et al., 2014; Zanne et al., 2014). In addition, the TCH predicts major breaks in the geographical distribution of clades (i.e., high phylogenetic turnover) that correspond to the shifts in climate from tropical to temperate, particularly in cold areas (Wiens and Donoghue, 2004). Among the five regions delineated by the phylogenetic regionalization scheme, the Paleotropic region had the highest FD (Table 1) and significantly higher phylogenetic turnover than expected

(Fig. 2c). This finding supports the idea that there was limited migration of the tropical lineages in South China and is consistent with TCH predictions. Therefore, phylogenetic regionalization was better than taxonomic regionalization for reflecting the evolutionary history of tropical lineages in South China based on their migration and establishment outside of the tropics.

4.3. Merits of phylogeny-based approach

The above comparisons demonstrate that the phylogenetic dissimilarity approach can provide more information than the taxonomic dissimilarity approach on the relationships between different phytogeographical regions. The phytogeographical regions and relationships between regions derived from the phylogenetic dissimilarity approach reflected more of the evolutionary history of flora than those from the taxonomic dissimilarity approach. Previous studies revealed that the biodiversity patterns in China were shaped by historical events, such as the uplift of the QTP in the early Miocene and the relevant Asian monsoon system (Sun and Wang, 2005; Favre et al., 2015; Lu et al., 2018). Our clustering dendrogram based on $p\beta_{sim}$ matrices reflected the phylogenetic relationships among regional floras (Fig. 2a) and highlights the important role of shared evolutionary histories of regional floras, including origin, diversification, and dispersal, in structuring the distribution of plants in China (Daru et al., 2017a).

In conclusion, compared with the taxonomic endemism and taxonomic dissimilarity approaches, the phylogenetic dissimilarity approach based on a densely sampled and well-resolved regional phylogeny at the genus level more effectively revealed hidden phylogenetic affinities among regions in biogeographical regionalization. We propose a new phytogeographical regionalization of the Chinese flora with five distinctive regions and evolutionary affinities.

Acknowledgements

We are grateful to Dr. Tianyu Hu and Wubing Xu for assistance with data analysis, and Dr. Jihong Huang and Liguang Chen for map preparation. We thank the researchers in Prof. Peihao Peng's lab at Chengdu University of Technology for distribution data collection. We also thank Prof. Ryan Frazier at Arizona State University and Dr. Jianhua Li at Hope College for their assistance with the English language. This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB31000000), the National Natural Science Foundation of China (NSFC 31590822, 31800178, 31870506 and 31461123001), the Sino-Africa Joint Research Center, Chinese Academy of Sciences, CAS International Research and Education Development Program (SAJC201613), Talent Foundation of Nanjing Forestry University (163108129), and Jiangsu Natural Science Foundation (BK20181398 to Lingfeng Mao). We thank Mallory Eckstut, Ph.D., for editing the English text of a draft of this manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.03.011.

References

Baselga, A., 2012. The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. Glob. Ecol. Biogeogr. 21, 1223–1232.
Baselga, A., Orme, C.D.L., 2012. Betapart: an R package for the study of beta diversity. Methods Ecol. Evol. 3, 808–812.
Borcard, D., Gillet, F., Legendre, P., 2011. Numerical ecology with R. Springer.
Chen, Z., et al., 2016. Tree of life for the genera of Chinese vascular plants. J. Syst. Evol. 54, 277–306.
Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P., Linder, H.P., 2009. Phylogenetic biome conservatism on a global scale. Nature 458, 754–756.
Daru, B.H., Michelle, B., Olivier, M., Kowiyou, Y., Hanno, S., Slingsby, J.A., Jonathan, D.T., 2016. A novel phylogenetic regionalization of phytogeographical zones of southern Africa reveals their hidden evolutionary affinities. J. Biogeogr. 43, 155–166.
Daru, B.H., Elliott, T.L., Park, D.S., Davies, T.J., 2017a. Understanding the processes underpinning patterns of phylogenetic regionalization. Trends Ecol. Evol. 32, 845–860.
Daru, B.H., Holt, B.G., Lessard, J.P., Yessoufou, K., Davies, T.J., 2017b. Phylogenetic regionalization of marine plants reveals close evolutionary affinities among disjunct temperate assemblages. Biol. Conserv. 213, 351–356.
Ebach, M.C., Parenti, L.R., 2015. The dichotomy of the modern bioregionalization revival. J. Biogeogr. 42, 1801–1808.
Engler, A., Diels, L., 1936. Syllabus der Pflanzenfamilien. Gebrüder Borntraeger Verlag.
Favre, A., Packert, M., Pauls, S.U., Jahnig, S.C., Uhl, D., Michalak, I., Muellner-Riehl, A.N., 2015. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. Biol. Rev. Camb. Philos. Soc. 90, 236–253.
Forest, F., Grenyer, R., Rouget, M., Davies, T.J., Cowling, R.M., Faith, D.P., Balmford, A., Manning, J.C., Proches, S., van der Bank, M., Reeves, G., Hedderson, T.A.J., Savolainen, V., 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. Nature 445, 757–760.
Graham, C.H., Fine, P.V., 2008. Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. Ecol. Lett. 11, 1265–1277.
González-Orozco, C.E., Ebach, M.C., Laffan, S., Thornhill, A.H., Knerr, N.J., Schmidt-Lebuhn, A.N., Cargill, C.C., Clements, M., Nagalingum, N.S., Mishler, B.D., 2014. Quantifying phytogeographical regions of Australia using geospatial turnover in species composition. PLoS ONE 9 e92558–e92558.
Gower, J.C., 1983. Comparing classifications. In: Felsenstein, J. (Ed.), Numerical taxonomy. Springer, New York, pp. 137–155.
Handel-Mazzettii, H., 1931. Die pflanzengeographische Gliederung und Stellung Chinas. Engl. Bot. Jahrb. 64, 309–323.
Hattab, T., Albouy, C., Lasram, F.B.R., Loc'H, F.L., Guilhaumon, F., Leprieur, F., 2015. A biogeographical regionalization of coastal Mediterranean fishes. J. Biogeogr. 42, 1336–1348.
He, J., Kreft, H., Gao, E., Wang, Z., Jiang, H., 2017. Patterns and drivers of zoogeographical regions of terrestrial vertebrates in China. J. Biogeogr. 44, 1172–1184.
Holt, B.G., Lessard, J.P., Borregaard, M.K., Fritz, S.A., Araújo, M.B., Dimitrov, D., Fabre, P.H., Graham, C.H., Graves, G.R., Jonsson, K.A., Nogués-Bravo, D., Wang, Z., Whittaker, R.J., Fjeldså, J., Rahbek, C., 2013. An update of Wallace's zoogeographic regions of the world. Science 339, 74–79.
Huang, H., Oldfield, S.F., Qian, H., 2013. Global significance of plant diversity in China. In: Hong, D.Y., Blackmore, S. (Eds.), Plants of China. Science Press, Beijing, pp. 7–34.
Huang, J., Huang, J., Lu, X., Ma, K., 2016. Diversity distribution patterns of Chinese endemic seed plant species and their implications for conservation planning. Sci. Rep. 6, 33913.
Jønsson, K.A., Holt, B.G., 2015. Islands contribute disproportionately high amounts of evolutionary diversity in passerine birds. Nat. Commun. 6, 8538.
Kelley, L.A., Gardner, S.P., Sutcliffe, M.J., 1996. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. Protein Eng. 9, 1063–1065.
Kerkhoff, A.J., Moriarty, P.E., Weiser, M.D., 2014. The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. Proc. Natl. Acad. Sci. USA 111, 8125–8130.
Kreft, H., Jetz, W., 2010. A framework for delineating biogeographical regions based on species distributions. J. Biogeogr. 37, 2029–2053.
Kubota, Y., Hirao, T., Fujii, S., Shiono, T., Kusumoto, B., 2014. Beta diversity of woody plants in the Japanese archipelago: the roles of geohistorical and ecological processes. J. Biogeogr. 41, 1267–1276.
López-Pujol, J., Ren, M., 2010. China: a hot spot of relict plant taxa. In: Rescigno, V., Maletta, S. (Eds.), Biodiversity hotspots. Nova Science Publishers, New York, pp. 123–137.
López-Pujol, J., Zhang, F., Sun, H., Ying, T., Ge, S., 2011. Centres of plant endemism in China: places for survival or for speciation? J. Biogeogr. 38, 1267–1280.
Ladle, R., Whittaker, R.J., 2011. Conservation biogeography. Wiley-Blackwell.
Leprieur, F., Oikonomou, A., 2014. The need for richness-independent measures of turnover when delineating biogeographical regions. J. Biogeogr. 41, 417–420.
Li, R., Kraft, N.J.B., Yang, J., Wang, Y., 2015. A phylogenetically informed delineation of floristic regions within a biodiversity hotspot in Yunnan. China. Sci. Rep. 5, 9396.
Lu, L., Mao, L., Yang, T., Ye, J., Liu, B., Li, H., Sun, M., Miller, J.T., Mathews, S., Hu, H., Niu, Y., Peng, D., Chen, Y., Smith, S.A., Chen, M., Xiang, K., Le, C.T., Dang, V.C., Lu, A., Soltis, P.S., Soltis, D.E., Li, J.H., Chen, Z., 2018. Evolutionary history of the angiosperm flora of China. Nature 554, 234–238.
Minchin, P.R., 1987. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 69, 89–107.
Mittermeier, R.A., Robles G.P., Mittermeier, C.G., 1997. Megadiversity: earth's biologically wealthiest nations. CEMEX/Agrupaciaon Sierra Madre.
Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2016. Vegan: community ecology package. R package vers. 2.2–1. https://CRAN.R-project.org/package=vegan.
R Core Team, 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Slik, J.W.F., Franklin, J., Arroyo-Rodríguez, V., Field, R., Aguilar, S., Aguirre, N., Ahumada, J., Aiba, S.I., Alves, L.F., Anitha, K., Avella, A., Mora, F., et al., 2018. Phylogenetic classification of the world's tropical forests. Proc. Natl. Acad. Sci. USA 115, 1837–1842.
Smith, S.A., O'Meara, B.C., 2012. TreePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28, 2689–2690.
Sokal, R.R., Rohlf, F.J., 1962. The comparison of dendrograms by objective methods.

Taxon 11, 33–40.
Sun, H., Wu, Z., 2013. Phytogeographical regions of China. In: Hong, D.Y., Blackmore, S. (Eds.), Plants of China (A Companion to the Flora of China). Science Press and Cambridge University Press, Beijing and Cambridge, pp. 176–204.
Sun, X., Wang, P., 2005. How old is the Asian monsoon system? – Palaeobotanical records from China. Paleogeogr. Paleoclimatol. Paleoecol. 222, 181–222.
Takhtajan, A., 1978. The floristic regions of the world. Academy of Sciences of the U.S.S.R.
Thorne, R.F., 1987. Phytogeography: floristic regions of the world. Science 236, 4797.
Wallace, A.R., 1876. The geographical distribution of animals; with a study of the relations of living and extinct faunas as elucidating the past changes of the Earth's surface. Macmillan Company, London.
Webb, C.O., Donoghue, M.J., 2005. Phylomatic: tree assembly for applied phylogenetics. Mol. Ecol. Notes 5, 181–183.
Wen, J., 1999. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. Annu. Rev. Ecol. Syst. 30, 421–455.
White, D., Gramacy, R.B., 2012. Maptree: Mapping, pruning, and graphing tree models. R package version 1.4-7. https://CRAN.R-project.org/package=maptree.
Wiens, J.J., Donoghue, M.J., 2004. Historical biogeography, ecology and species richness. Trends Ecol. Evol. 19, 639–644.
Wiens, J.J., Graham, C.H., Moen, D.S., Smith, S.A., Reeder, T.W., 2006. Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: treefrog trees unearth the roots of high tropical diversity. Am. Nat. 168, 579–596.
Wolfe, J.A., 1975. Some aspects of plant geography of the northern hemisphere during the late Cretaceous and Tertiary. Ann. Mo. Bot. Gard. 62, 264–279.
Wu, Z., 1979. The regionalization of Chinese flora. Acta Bot. Yunnan. 1, 1–20.
Wu, Z., Raven, P., Hong, D.Y., 1994–2013. Flora of China. Science Press and Missouri Botanical Garden Press, Beijing and St Louis.
Wu, Z., Sun, H., Zhou, Z., Li, D., Peng, H., 2010. Floristics of seed plants from China. Science Press, Beijing.
Xu, H., Cao, M., Wu, Y., Cai, L., Cao, Y., Ding, H., Cui, P., Wu, J., Wang, Z., Le, Z., 2017. Optimized monitoring sites for detection of biodiversity trends in China. Biodiv. Conserv. 5, 1–13.
Ying, T., Zhang, Y., Boufford, D.E., 1994. The endemic genera of seed plants of China. Science press, Beijing.
Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G., McGlinn, D.J., O'Meara, B.C., Moles, A.T., Reich, P.B., Royer, D.L., Soltis, D.E., Stevens, P.F., Westoby, M., Wright, I.J., Aarssen, L., Bertin, R.I., Calaminus, A., Govaerts, R., Hemmings, F., Leishman, M.R., Oleksyn, J., Soltis, P.S., Swenson, N.G., Warman, L., Beaulieu, J.M., 2014. Three keys to the radiation of angiosperms into freezing environments. Nature 506, 89–92.
Zhang, D., Ye, J., Sun, H., 2016a. Quantitative approaches to identify floristic units and centres of species endemism in the Qinghai-Tibetan Plateau, south-western China. J. Biogeogr. 43, 2465–2476.
Zhang, M., Slik, J.W.F., Ma, K., 2016b. Using species distribution modeling to delineate the botanical richness patterns and phytogeographical regions of China. Sci. Rep. 6, 22400.