

**ORIGINAL RESEARCH**

# Environmental and evolutionary drivers of diversity patterns in the tea family (Theaceae s.s.) across China

Mide Rao<sup>1,2</sup>  | Manuel J. Steinbauer<sup>3</sup> | Xiaoguo Xiang<sup>4</sup> | Minggang Zhang<sup>5</sup> |  
 Xiangcheng Mi<sup>2</sup>  | Jintun Zhang<sup>1</sup> | Keping Ma<sup>2</sup>  | Jens-Christian Svenning<sup>3,6</sup> 

<sup>1</sup>Key Laboratory of Biodiversity Sciences and Ecological Engineering, Ministry of Education, College of Life Sciences, Beijing Normal University, Beijing, China

<sup>2</sup>State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China

<sup>3</sup>Section for Ecoinformatics and Biodiversity, Department of Bioscience, Aarhus University, Aarhus C, Denmark

<sup>4</sup>State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China

<sup>5</sup>Institute of Loess Plateau, Shanxi University, Taiyuan, China

<sup>6</sup>Center for Biodiversity Dynamics in a Changing World (BIOCHANGE), Aarhus University, Aarhus C, Denmark

**Correspondence**

Xiangcheng Mi, State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China.

Email: mixiangcheng@ibcas.ac.cn  
and

Jintun Zhang, Key Laboratory of Biodiversity Sciences and Ecological Engineering, Ministry of Education, College of Life Sciences, Beijing Normal University, Beijing, China.

Email: zhangjt@bnu.edu.cn

**Funding information**

Natur og Univers, Det Frie Forskningsråd, Grant/Award Number: 6108-00078B; Villum Fonden, Grant/Award Number: 16549; National Key Technologies Research and Development Program, Grant/Award Number: 2017YFC0503803; National Natural Science Foundation of China, Grant/Award Number: 31170494; National Natural Science Foundation of China, Grant/Award Number: 31670212; EU's Marie Skłodowska-Curie action, Grant/Award Number: 707491; Strategic Priority Research Program of Chinese Academy of Sciences, Grant/Award Number: XDPB020306; Carlsbergfondet, Grant/Award Number: CF14-0148; Chinese Scholarship Council Scholarship, Grant/Award Number: 201506040162; European Research Council, Grant/Award Number: ERC-2012-StG-310886-HISTFUNC

**Abstract**

Subtropical forest is recognized as an important global vegetation type with high levels of plant species richness. However, the mechanisms underlying its diversity remain poorly understood. Here, we assessed the roles of environmental drivers and evolutionary dynamics (time-for-speciation and diversification rate) in shaping species richness patterns across China for a major subtropical plant group, the tea family (Theaceae s.s.) (145 species), at several taxonomic scales. To this end, we assessed the relationships between species richness, key environmental variables (minimum temperature of the coldest month, mean annual precipitation, soil pH), and phylogenetic assemblage structure (net related index) by using non-spatial and spatial linear models. We found that species richness is significantly related to environmental variables, especially soil pH, which is negatively related to species richness both across the whole family and within the major tribe Theaeae (116 species). Family-level species richness is unrelated to phylogenetic structure, whereas species richness in tribe Theaeae was related to phylogenetic structure with U-shaped relationship, a more complex relation than predicted by the time-for-speciation or diversification rate hypotheses. Overall, these results suggest that both environmental and evolutionary factors play important roles in shaping species richness patterns within this subtropical plant family across China, with the latter mainly important at fine taxonomic scales. Most surprisingly, our findings show that soils can play a key role in shaping macro-scale diversity patterns, contrary to often-stated assumptions.

**KEYWORDS**

China, evolution, phylogeny, soil pH, species richness, Theaceae

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Subtropical forests are important terrestrial ecosystems that contribute importantly to both biodiversity and ecosystem functioning (Fang & Yoda, 1991; Wu & Wu, 1998). This is especially true in China, which contains the largest subtropical forest area in the world (Song, Chen, & Wang, 2005; Zhu, 2013; Zhu et al., 2008). Although many studies have explored drivers of high diversity in the tropics (Mittelbach et al., 2007; Wiens & Donoghue, 2004), what drives diversity patterns within the large subtropical region in China remains poorly understood.

Climate is typically expected to be the primary driver of species diversity patterns at large geographic scales (Field et al., 2009; Hawkins et al., 2003). Increasing intensity of frost has been found to determine the decline in woody species richness with latitude in China, with annual precipitation also explaining much of the variation in woody species richness (Wang, Fang, Tang, & Lin, 2011). Subtropical China is characterized by the monsoon climate with mean annual temperature ranging from 15 to 20°C, mean annual precipitation (MAP) ranging from 900 to 2,000 mm, and a frost period ranging from 64 to 100 days (Wang, Kent, & Fang, 2007). We thus expect that species diversity patterns within this region are shaped by both rainfall and temperature. However, environmental variables related to edaphic properties, such as soil pH, are suggested as important predictors for species richness at local scales (Palpurina et al., 2017; Pausas & Austin, 2001; Zellweger et al., 2016). Subtropical China is mainly covered with red soils, which are highly weathered, nutrient-deficient, and acidic with high accumulation of aluminum (Al) and iron (Fe) (Wilson, He, & Yang, 2004). Soil pH strongly affects the exchangeable base cations (calcium [Ca], magnesium [Mg], potassium [K], sodium [Na]), mobilizing exchangeable Al and affecting the availability of nutrients such as K and phosphorus (P), which are crucial for tree growth (Lieb, Darrouzet-Nardi, & Bowman, 2011). Soil pH displays a strong trend, increasing from the southeast to the northwest across China (Xiong & Li, 1987), which suggests that it could play a role in large-scale diversity patterns. However, the role of soil pH in shaping species richness at large spatial scales has rarely been studied neither generally nor for China specifically (but see Azevedo, Zelm, Hendriks, Bobbink, & Huijbregts, 2013).

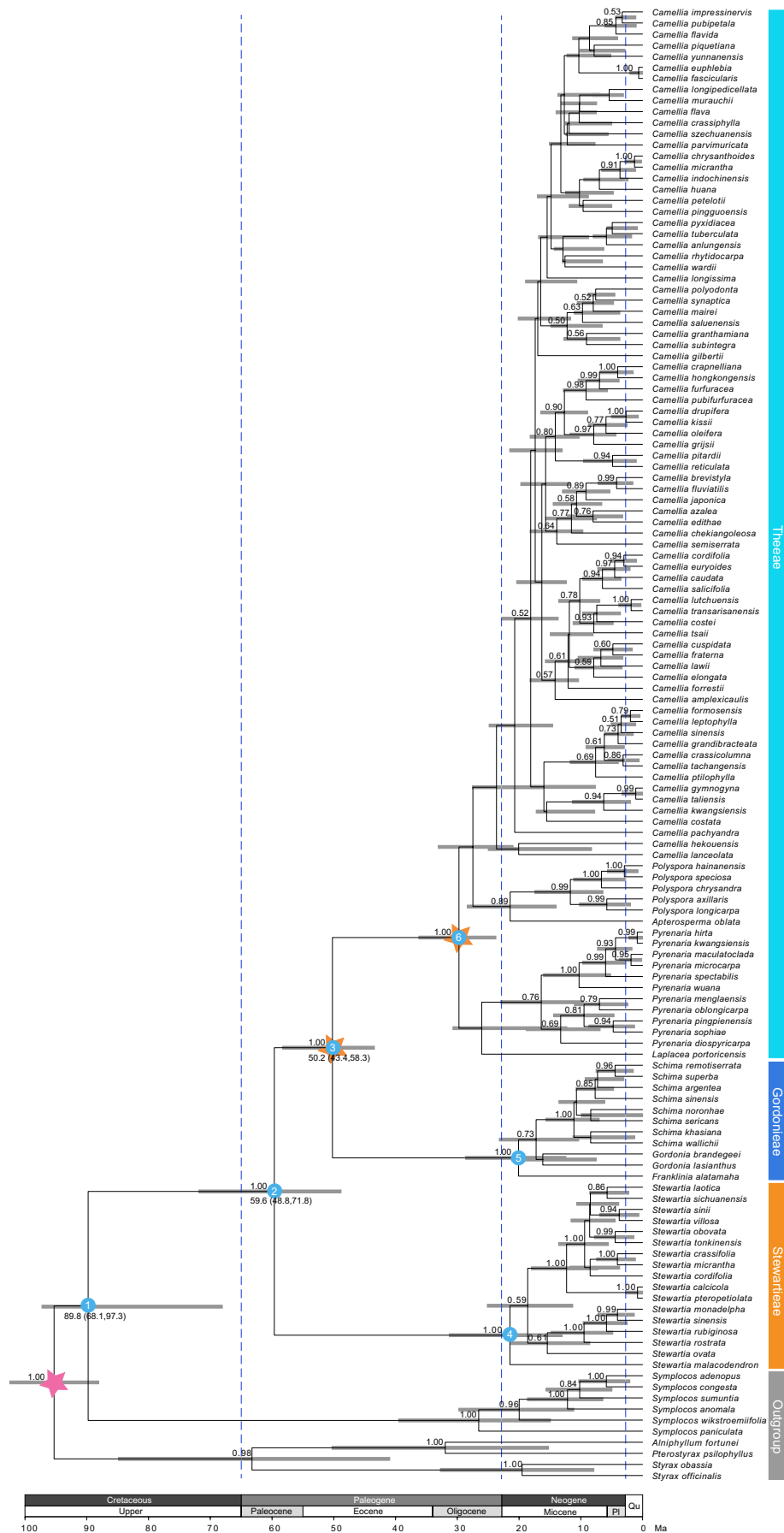
Patterns of species richness at large spatial scales should ultimately be associated with evolutionary and biogeographic processes that directly influence the number of species, namely, speciation, extinction, and dispersal (Ricklefs, 2004; Wiens, Parra-Olea, Garcia-Paris, & Wake, 2007). Two general explanations have been proposed to explain geographical patterns in species richness

(Wiens, 2011). First, species-rich areas may have had more time for speciation to accumulate species due to longer occupancy in certain environmental conditions (Qian, Wiens, Zhang, & Zhang, 2015; Stephens & Wiens, 2003). Second, more species-rich areas may have experienced higher diversification (i.e., speciation minus extinction) rates in certain environmental conditions (Qian et al., 2015; Svenning, Borchsenius, Bjorholm, & Balslev, 2008). Environmental factors could play different roles underlying these two explanations: Under the first hypothesis, known as time-for-speciation, environmental factors are associated with limitations to the colonization of new habitats due to niche conservatism (Qian, Zhang, Zhang, & Wang, 2013; Wiens & Donoghue, 2004; Wiens et al., 2010). Under the second hypothesis, called diversification rate, environmental factors are responsible for species richness patterns through their influence on the diversification rate (Qian et al., 2015; Wiens, 2011). Evidence for time-for-speciation has been found for various taxonomic groups (Kerckhoff, Moriarty, & Weiser, 2014; Li et al., 2009; Wiens, Graham, Moen, Smith, & Reeder, 2006). The diversification rate hypothesis also has been supported by several studies (Cardillo, Orme, & Owens, 2005; Pyron & Wiens, 2013; Svenning et al., 2008). Consequently, these mechanisms are not mutually exclusive and their relative importance may vary by clade and region. Relatively few studies have tested these two hypotheses simultaneously (Marin & Hedges, 2016; Qian et al., 2015; Svenning et al., 2008), particularly in subtropical regions.

The tea family (Theaceae s.s.) is a dominant woody constituent of subtropical forests in eastern Asia (Wu, 1995). Theaceae s.s. includes three major lineages: tribes Theeae, Gordonieae, and Stewartieae, with Theeae being the most species-rich. Theaceae species are concentrated in subtropical forests and tropical mountain areas, with most genera being evergreen broad-leaved shrubs and trees (Luna & Ochoterena, 2004). Their geographic distribution and characteristic traits reflect their limited tolerance of frost (Sakai & Weiser, 1973). Furthermore, Theaceae species usually require ample amounts of water and prefer acidic soils (Ming, 2000). There is a rich fossil record for Theaceae, beginning from the late Cretaceous through the Paleogene and the Neogene (Prince, 2009), which shows that Theaceae was a conspicuous component of the vegetation across North America, Europe, and Asia prior to the late Neogene cooling and the Quaternary glaciations (Grote & Dilcher, 1989). Today, Theaceae s.s. diversity is concentrated in southeast Asia and in southern China and is absent from Europe (Grote & Dilcher, 1989).

In this study, we combine species richness, environmental and phylogenetic data for Theaceae s.s. to address the following

**FIGURE 1** Timing of the diversification of Theaceae (133 species). Chronogram is derived from the maximum clade credibility tree by BEAST analyses. Gray bars indicate 95% highest posterior density (HPD) intervals of the age estimates. The red and orange stars represent the root and fossil calibration nodes separately (see Section 2). The blue circles with numbers indicate the nodes of interest (see Section 6). Numbers above the branch indicate Bayesian inference (BI) posterior probability (PP) values. Only nodes with PP support >0.5 are shown. Numbers below the branch indicate the estimated divergent median ages for the nodes, where values in parentheses represent 95% HPD. The ages of stratigraphic boundaries were obtained from the International Chronostratigraphic Chart (Cohen, Finney, Gibbard, & Fan, 2013) (PI, Pliocene; Qu, Quaternary)



questions: (a) How are the patterns of species richness in Theaceae s.s. distributed along environmental gradients across China? (b) What causes these patterns? Is the higher species richness in some environments caused by greater time for building up species or/and faster net rates of diversification in certain environments? (c) Do the causes of the species richness patterns differ at the family and tribe levels?

## 2 | MATERIALS AND METHODS

### 2.1 | Species distribution and environmental data

We compiled presence/absence data for Theaceae s.s. (hereafter Theaceae for simplicity) (following *Flora of China*, Ming & Bartholomew, 2007) in 100 km × 100 km cells across China using the Chinese Vascular Plant Database, in which plant occurrence locations were recorded at the county level. This database is based on: (a) a specimen's locality from the National Specimen Information Infrastructure ([www.nsii.org.cn](http://www.nsii.org.cn)); (b) published and provincial floras, including *Flora Republicae Popularis Sinicae* (FRPS: *Flora of China*, Chinese version); as well as (c) public checklists, and species surveying reports for all 145 Chinese tea family species (117 in Theaceae, 13 in Gordonieae, 15 in Stewartieae). We removed border cells with an area <5,000 km<sup>2</sup>. Because the sample sizes for Gordonieae and Stewartieae were too small, analyses were only run on for the full dataset, comprising the whole family (Theaceae), and for the biggest tribe (Theaceae). To compute the phylogenetic assemblage structure (see below), cells with fewer than three Theaceae or Theaceae species were excluded, yielding a final subset of 233 cells and 145 species for the family-wide dataset and of 230 cells and 116 species for the tribe-level dataset.

Environmental data were sampled for all grid cells in each subset. Nineteen bioclimatic variables were obtained from WorldClim (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005), but we mainly used minimum temperature of the coldest month (MINT) and MAP. We excluded other variables due to their weak predictive abilities or because they were highly correlated with MINT/MAP in our preliminary analysis. Soil pH was obtained from SoilGrids (Hengl et al., 2014). All data were obtained at a resolution of 30 arc seconds.

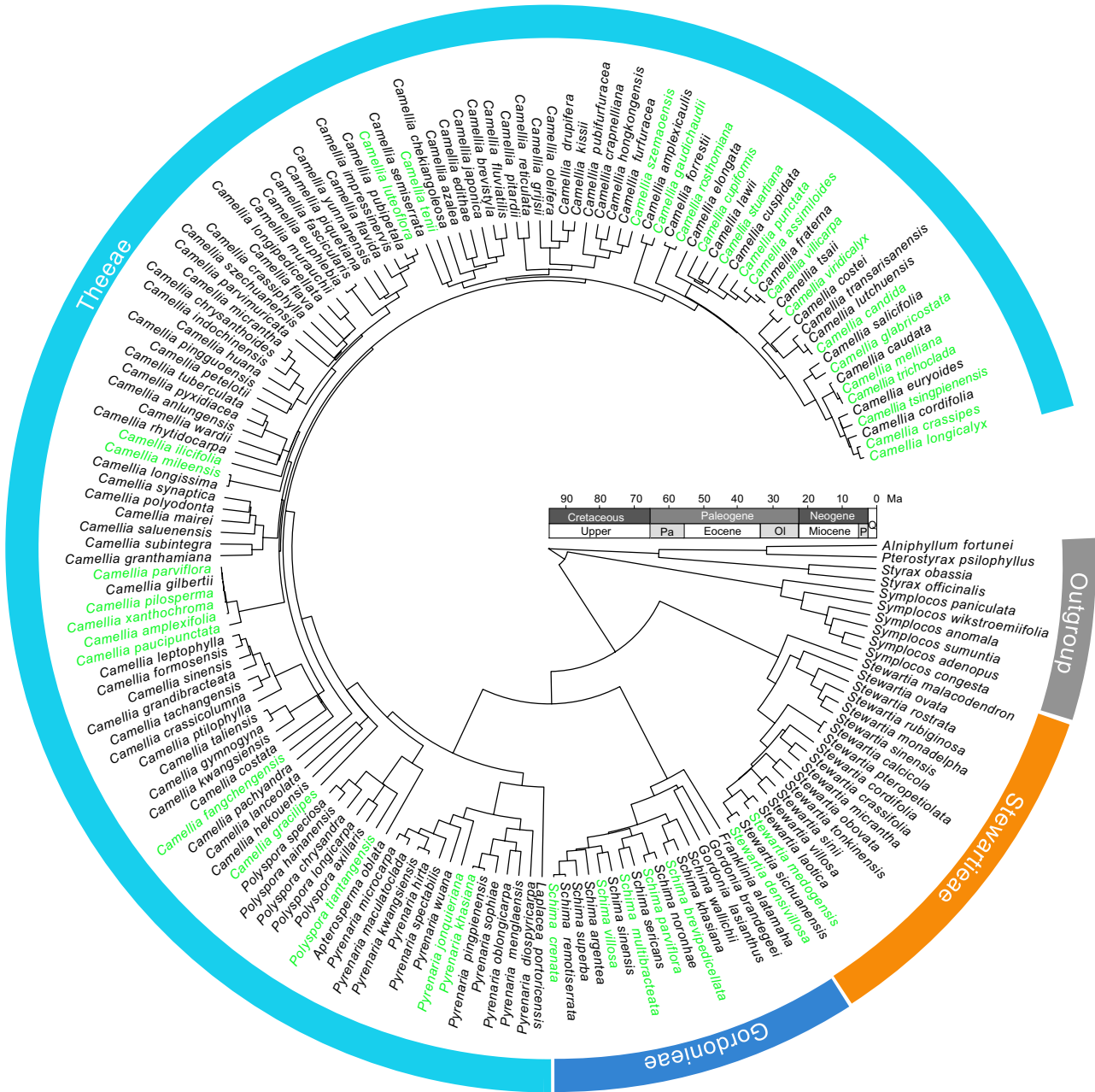
### 2.2 | Phylogenetic analysis

We sampled a total of 133 species with relevant sequences from GenBank, including 123 ingroup and 10 outgroup species, which were selected from closely related families in Styracaceae and Symplocaceae (see Supporting Information Table S1). The taxonomy of species in China follows Ming and Bartholomew (2007), whereas the taxonomy of species outside of China was based on The Plant List (<https://www.theplantlist.org>). Ten chloroplast DNA sequences (*atpB-rbcL*, *atpI-atpH*, *matK*, *matK-trnK*, *psbA-trnH*, *rbcL*, *rbcL-accD*, *rpl16*, *rpl32-trnL*, and *trnL-trnF*) and the nuclear internal transcribed spacer (ITS) region were used for analyses (GenBank accession numbers are given in Supporting Information Table S1).

Sequences alignment was performed in Clustal X (Thompson, Gibson, Plewniak, Jeanmougin, & Higgins, 1997) and then adjusted manually in BioEdit (Hall, 1999). The cpDNA and ITS regions were analyzed separately. The homogeneity between the cpDNA and ITS was tested using the incongruence length difference test (Farris, Källersjö, Kluge, & Bult, 1994), which was implemented in PAUP v4.0b10 (Swofford, 2003). The result ( $p = 0.01$ ) suggested the incongruence between ITS and cpDNA. According to Nishii et al. (2015), incongruences supported by bootstrap values higher than 75% and/or 0.95 posterior probabilities would be regarded as significant. A comparison of the topologies of the ITS with the cpDNA trees for the analyses revealed no strong conflict in relationships between the main clades. Thus, the combined cpDNA and ITS matrix was used to estimate the divergence times.

Divergence times were estimated using a Bayesian uncorrelated relaxed-clock model in BEAST 2.3.0 (Bouckaert et al., 2014). For BEAST analysis, cpDNA and ITS were assigned GTR+I+Γ, respectively, as determined by the Akaike information criterion (AIC) in Modeltest 3.7 (Posada & Crandall, 1998). Three calibration points were used: the root was set to 96 million years ago (Ma) with a normal prior distribution for the whole tree (88–103 Ma) (Wikström, Kainulainen, Razafimandimbison, Smedmark, & Bremer, 2015) (Figure 1); the stem node and crown node of Theaceae were constrained to a minimum age of 40 Ma (Figure 1) and 20 Ma (Figure 1), respectively, using a lognormal prior distribution with a standard deviation of 1.0, and following the example of Zhang, Kan, Zhao, Li, and Wang (2014). The Yule process was chosen as the speciation process. Markov chain Monte Carlo searches were run for 100,000,000 generations and sampled every 1,000 generations. Convergence and effective sample sizes of all parameters were assessed in Tracer 1.6.0 (Rambaut, Suchard, Xie, & Drummond, 2014). The maximum clade credibility tree was computed using treeAnnotator 2.3.0 (Bouckaert et al., 2014) and is shown in Figure 1 (133 species). Thirty-seven species in China that did not have molecular data were added manually to the tree as follows: (a) We randomly selected 1,000 trees from the BEAST tree set after the burn-in; (b) based on known morphological classification information from the *Flora of China* (Ming & Bartholomew, 2007), we added the 37 species to each tree in their corresponding sections with random branch lengths using the R package “phytools” (Revell, 2012). Finally, we had compiled a set of 1,000 random-addition trees with 170 species (for example tree, see Figure 2).

We calculated the mean phylogenetic net relatedness index (NRI; Webb, 2000) based on the set of 1,000 random-addition trees and quantified the degree of phylogenetic relatedness among species for each 100 km × 100 km cell. NRI was calculated using the mean pairwise distance (MPD), which measures the mean phylogenetic relatedness between all pairs of species occurring in an assemblage:  $NRI = -1 \times (MPD_{\text{observed}} - MPD_{\text{randomized}}) / sd(MPD_{\text{randomized}})$ , where  $MPD_{\text{observed}}$  is calculated from species occurring in the given cell, and  $MPD_{\text{randomized}}$  is the expected MPD distribution from the 1,000 null models. Positive values of NRI indicate that the species present in an assemblage are more closely related to each other than expected



**FIGURE 2** Example of a dated tree of Theaceae (170 species) from the set of 1,000 trees where species without genetic information have been added with random branch lengths in their responding sections, respectively. Black tip labels indicate sequenced species. Green tip labels indicate non-sequenced species. The ages of stratigraphic boundaries were obtained from the International Chronostratigraphic Chart (Cohen et al., 2013) (Pa, Paleocene; OI, Oligocene; P, Pliocene; Qu, Quaternary)

by chance (phylogenetic clustering), while negative values of NRI indicate that the species are less related to each other than expected by chance (phylogenetic overdispersion). Because the values are centered at zero and standardized by the standard deviation, values >1.96 indicate statistically significant phylogenetic clustering, while values <-1.96 indicate statistically significant phylogenetic overdispersion (Vamosi, Heard, Vamosi, & Webb, 2009). NRI was calculated in R 3.3.2 (R Core Team, 2016) with the “picante” package (Kembel et al., 2010).

### 2.3 | Statistical analyses

We first used Pearson correlations to assess pairwise relationships among the different variables. To account for spatial autocorrelation, Dutilleul’s (1993) modified *t* test was used to calculate statistical significance with package “SpatialPack 0.2-3” (Osorio & Vallejos, 2014) in R.

Spatial linear models (SLM) were applied to estimate the richness patterns of tea family species along environmental gradients (MINT, MAP, pH) across China. To account for nonlinear relationships



between the variables, the quadratic terms of the predictor variables were included in the regressions. To avoid multicollinearity problems, we divided environmental variables (i.e., MAP and soil pH, Tables 1 and 2) with high pairwise correlation ( $|r| > 0.7$ ) (Dormann et al., 2013) into different datasets for the modeling. Thus, we defined two groups of environmental variables that were only used separately: MINT+MAP and MINT+pH.

We sought to identify the drivers of species richness along environmental gradients by comparing patterns of phylogenetic relatedness among species in different species assemblages and relating these patterns to environmental factors (Algar, Kerr, & Currie, 2009; Qian et al., 2015). For example, if time is important for patterns of species richness along environmental gradients, in species-rich environments, species should be relatively distantly related to each other, indicating a long period of occupancy. Then for environments with rare and recent colonization, species should be relatively closely related (Qian et al., 2015, 2013). Therefore, a negative relationship between NRI and species richness, and a significant relationship between NRI and a certain environmental variable may support the time-for-speciation hypothesis. If higher richness in certain environments is much more likely caused by a higher rapid diversification rate, then species should be relatively more closely related to each other in high-richness environments. Thus, a positive relationship between NRI and species richness, and a significant relationship between NRI and a certain environmental variable should be found (Qian et al., 2015). If both time and diversification rate are important in explaining species richness patterns, the different relationships between species richness and NRI may cancel each other out, resulting in weak or absent species richness–NRI relationships. Here, we used SLM to analyze the relationship between richness and NRI, as well as the relationships between NRI and different groups of environmental variables.

We also fitted non-spatial ordinary least squares (OLS) regression models to complement the SLM results. Since significant spatial autocorrelation was found in the residuals of OLS models

(Supporting Information Tables S2, S3, and S4), we here emphasize the SLM results. Still, we also report the OLS results, as spatial autocorrelation has been argued to not seriously affect OLS estimation of regression coefficients (Hawkins, Diniz-Filho, Mauricio Bini, Marco, & Blackburn, 2007), and that controlling for spatial autocorrelation may shift the spatial scale of the analyses (Diniz-Filho, Bini, & Hawkins, 2003). The OLS model residuals were found to approximate a normal distribution (Supporting Information Figure S1).

Both SLM and OLS models were run in R. For SLM, spatial simultaneous autoregressive error (SAR) models were built using the “spdep” package (Bivand et al., 2015). The Moran's  $I$  values were used to quantify the presence of spatial autocorrelation in SAR or OLS models (Kissling & Carl, 2008). The best SAR or OLS models were identified based on the Akaike information criterion corrected for small sample size ( $AIC_c$ ) (Burnham & Anderson, 2002). To evaluate the relative importance of variables in the SAR or OLS modeling, the importance of each predictor was determined by Akaike weights, computed with the “MuMIn” package (Barton, 2015). Species richness and MAP were log-transformed to improve normality in our models, and all predictor variables were standardized to a mean of zero and standard deviation of one to allow for the direct comparison of regression coefficients.

### 3 | RESULTS

Within Theaceae, all three tribes (Stewartieae, Gordonieae, and Theeae) were found to be monophyletic with strong support values ( $BI = 1.0$ ; nodes 4, 5, and 6; Figure 1). Stewartieae was the first divergent clade ( $BI = 1.0$ ; node 2, Figure 1), while Gordonieae was sister to Theeae with strong support values ( $BI = 1.0$ ; node 3, Figure 1). The stem age of Theaceae was estimated at 89.8 Ma (95% highest posterior density (HPD): 68.1–97.3; node 1; Figure 1), and the crown age of the family was estimated to be 59.6 Ma (95% HPD: 48.8–71.8;

Variable	MINT	MAP	pH	NRI
MAP	<b>0.617 (0.046)</b>			
pH	–0.399 (0.141)	<b>–0.813 (0.035)</b>		
NRI	0.102 (0.263)	–0.034 (0.664)	0.145 (0.071)	
SR	<b>0.427 (0.031)</b>	0.424 (0.059)	<b>–0.512 (0.007)</b>	–0.067 (0.479)

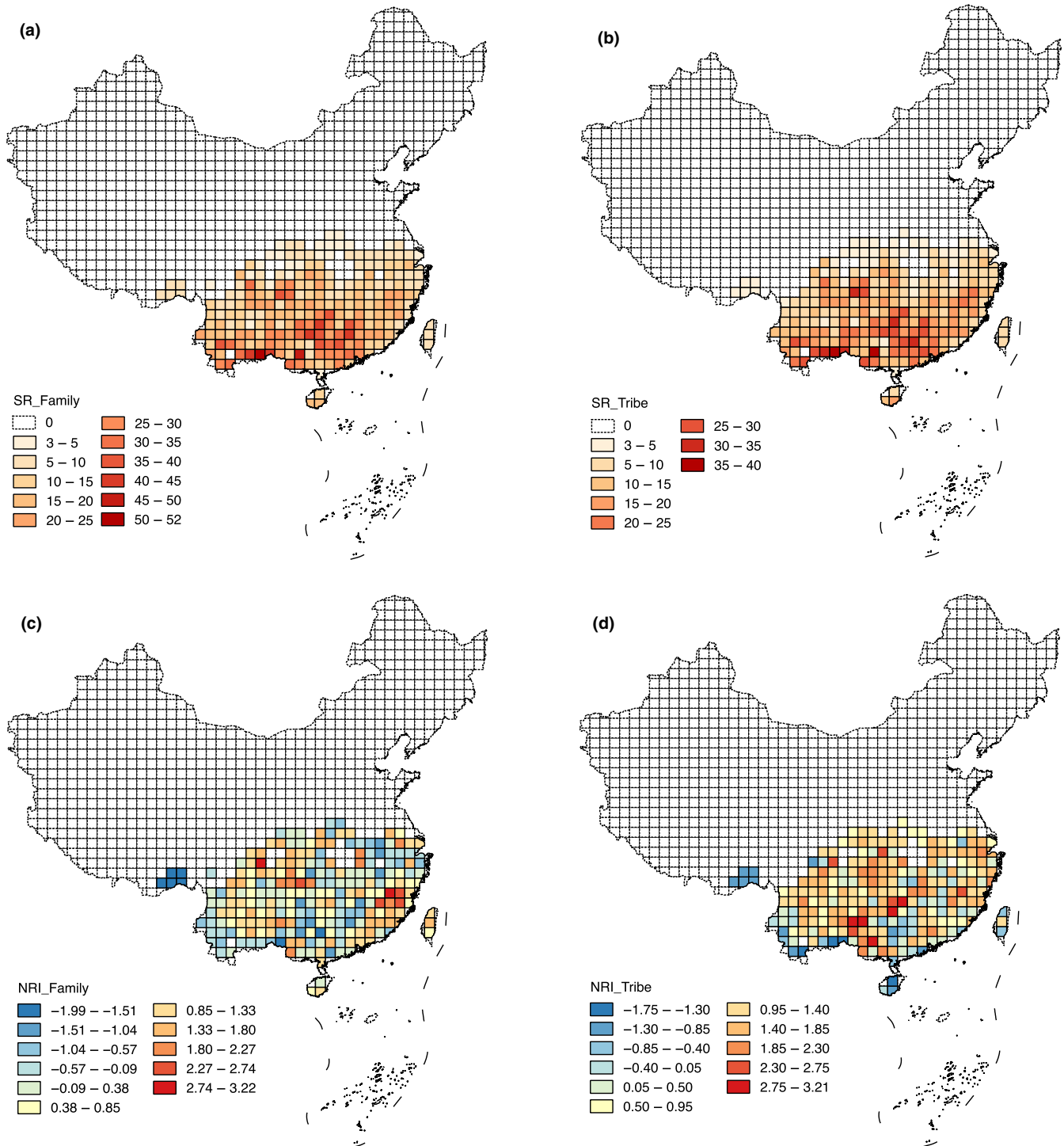
Note.  $P$  values were calculated after accounting for spatial autocorrelation in parentheses. Significant values ( $p < 0.05$ ) are marked in bold.

**TABLE 1** Pearson correlations among variables for Theaceae

Variable	MINT	MAP	pH	NRI
MAP	<b>0.587 (0.021)</b>			
pH	–0.364 (0.102)	<b>–0.804 (0.000)</b>		
NRI	–0.145 (0.221)	–0.119 (0.352)	0.199 (0.081)	
SR	<b>0.444 (0.019)</b>	0.366 (0.106)	<b>–0.397 (0.038)</b>	0.100 (0.316)

Note.  $P$  values were calculated after accounting for spatial autocorrelation in parentheses. Significant values ( $p < 0.05$ ) are marked in bold.

**TABLE 2** Pearson correlations among variables for Theeae



**FIGURE 3** Geographical patterns of (a) species richness for Theaceae (SR\_Family); (b) species richness for Theaceae (SR\_Tribe); (c) net relatedness index (NRI) for Theaceae (NRI\_Family); (d) NRI for Theaceae (NRI\_Tribe)

node 2; Figure 1). Gordonieae and Theae were estimated to have diverged at 50.2 Ma (95% HPD: 43.4–58.3; node 3; Figure 1).

The number of Theaceae species in the sampled quadrats generally decreased with latitude (Figure 3a), but also displayed more complex geographic variation (Figure 3b). Neither Theaceae nor Theae showed clear geographic gradients in phylogenetic structure (NRI) (Figure 3c,d).

At family level, species richness was significantly correlated with both MINT ( $r = 0.427$ ) and soil pH ( $r = -0.512$ ) after accounting for spatial autocorrelation ( $p < 0.05$ ). Similarly, species richness in the Theae tribe was positively correlated with MINT ( $r = 0.444$ ,  $p < 0.05$ ) and negatively correlated with soil pH ( $r = -0.397$ ,  $p < 0.001$ ). Precipitation (MAP) was not correlated with species richness at either taxonomic level (Tables 1 and 2). Species richness and

NRI were not related at either family level or tribe level (Tables 1 and 2). NRI was also not correlated with environmental variables at either level (Tables 1 and 2).

Among the SAR models for species richness, models including soil pH were much stronger than models including precipitation. Soil pH and species richness consistently had a strong negative relationship at both family level and tribe level (Table 3). MINT was negatively associated with species richness at family level, but the relationship was weak, and there was no effect of MINT on species richness at tribe level (Table 3).

According to the SAR models, species richness was independent of NRI at family level, but showed a U-shaped relationship at tribe level (Table 4), suggesting an evolutionary link at this lower scale.

Among the SAR models for NRI, the MINT+pH model was a better fit than the MINT+MAP model at family level (Table 5). Soil pH was much more important than temperature for phylogenetic structure, as shown by the strong positive NRI-pH relationship at family level. However, only MINT affected NRI at tribe level (Table 5), where MINT was unimodally related to NRI.

The OLS models provided consistent results with the SAR analyses for the relationships between species richness and soil pH (Tables 3 and Supporting Information Table S2, Figure 4), as well as the relationships between species richness and phylogenetic structure (NRI) (Tables 4 and Supporting Information Table S3, Figure 5). The main differences identified was that the effect of pH on NRI at tribe level was significant in the OLS modeling, but not supported in the SAR analyses (Tables 5 and Supporting Information Table S4, Figure 6). However, some differences for less supported environmental factors were also found (Tables 3, and 5, Supporting Information Tables S2, and S4, Figures S2 and S3).

## 4 | DISCUSSION

We here compiled species distribution data, environmental data, and phylogeny to assess the drivers of species richness in a key subtropical woody plant family (Theaceae) across China. Both the phylogenetic relationship and the divergence times estimated in our results were consistent with Yu et al. (2017). Our results indicate that both environmental and evolutionary factors play important roles in shaping species richness patterns of tea family species across China. The strong relationship between soil pH and species richness was found consistently at both family level and tribe level, supporting the idea that environmental factors explain much of the variation in the species richness of Theaceae. Although the species richness patterns along environmental gradients at family and tribe levels were similar, the processes for these patterns differed when considering evolutionary dynamics.

Species richness patterns for the whole tea family as well as the Theeae tribe in China exhibited significant relationships with environmental variables, of which soil pH was the strongest predictive variable, having a negative correlation with species richness in Theaceae and in Theeae (Table 3). Generally, climatic variables have been considered the main environmental factors that control large-scale patterns of species richness (Francis & Currie, 2003; McGill, 2010). At small scales, edaphic properties, like soil pH, have been shown to be key environmental factors influencing plant species richness patterns (Dubuis et al., 2013; Zellweger et al., 2016). However, our results indicate that soil is the most important environmental factor influencing the richness patterns of Theaceae, even on a biogeographic scale. The large areas of red soil in the subtropical region across south China used in our study, combined with the special adaptations of the tea family for low pH soils may

**TABLE 3** Multimodel inference from spatial simultaneous autoregressive error (SAR) models of species richness against environmental predictors for Theaceae and Theeae

Model parameters	Theaceae				Theeae			
	Coefficients	Akaike weight	Pseudo- $r^2$	Moran's $I$	Coefficients	Akaike weight	Pseudo- $r^2$	Moran's $I$
Group 1								
MINT	-0.084 <sup>†</sup>	0.680	0.515	-0.023 <i>ns</i>	-		0.478	-0.024 <i>ns</i>
MINT <sup>2</sup>	-				-			
MAP	0.167 <sup>***</sup>	1.000			0.116 <sup>***</sup>	1.000		
MAP <sup>2</sup>	-0.029 <sup>†</sup>	0.770			-			
Group 2								
MINT	-		0.558	-0.024 <i>ns</i>	-		0.514	-0.026 <i>ns</i>
MINT <sup>2</sup>	-				-			
pH	-0.169 <sup>***</sup>	1.000			-0.132 <sup>***</sup>	1.000		
pH <sup>2</sup>	-				-			

Note. Model sets involved all possible combinations of explanatory variables, for two groups of variables: Group 1: minimum temperature of the coldest month (MINT), mean annual precipitation (MAP). Group 2: MINT, soil pH (pH). Coefficients for the model with the lowest AICc for a given variable group are shown. The Akaike weight for each variable is based on the full model set per group. The superscript 2 indicates the quadratic form of the variables. Pseudo- $r^2$ , explained the variance of the SAR model. Moran's  $I$ , measure of residual spatial autocorrelation.

<sup>†</sup>Significance levels: \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; <sup>†</sup> $p < 0.05$ . *ns*, not significant.



**TABLE 4** Multimodel inference from spatial simultaneous autoregressive error (SAR) models of species richness against phylogenetic predictors for Theaceae and Theaeae

Model parameters	Theaceae				Theaeae			
	Coefficients	Akaike weight	Pseudo- $r^2$	Moran's $I$	Coefficients	Akaike weight	Pseudo- $r^2$	Moran's $I$
NRI	-0.017 <i>ns</i>	0.400	0.441	-0.020 <i>ns</i>	-		0.491	-0.014 <i>ns</i>
NRI <sup>2</sup>	-				0.037***	1.000		

Note. Coefficients for the model with the lowest AIC<sub>c</sub> are shown. The Akaike weight for each variable is based on the full model set. The superscript 2 indicates the quadratic form of the variable. Pseudo- $r^2$ , explained the variance of the SAR model. Moran's  $I$ , measure of residual spatial autocorrelation.

<sup>a</sup>Significance levels: \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$ . *ns*, not significant.

**TABLE 5** Multimodel inference from spatial simultaneous autoregressive error (SAR) models of phylogenetic structure (NRI) against environmental predictors for Theaceae and Theaeae

Model parameters	Theaceae				Theaeae			
	Coefficients	Akaike weight	Pseudo- $r^2$	Moran's $I$	Coefficients	Akaike weight	Pseudo- $r^2$	Moran's $I$
Group 1								
MINT	0.264*	0.550	0.210	0.003 <i>ns</i>	-0.230*	0.630	0.216	-0.004 <i>ns</i>
MINT <sup>2</sup>	-				-0.153**	0.900		
MAP	-0.186 <i>ns</i>	0.440			-			
MAP <sup>2</sup>	-				-			
Group 2								
MINT	0.212 <i>ns</i>	0.590	0.226	0.002 <i>ns</i>	-0.230*	0.660	0.216	-0.004 <i>ns</i>
MINT <sup>2</sup>	-				-0.153**	0.910		
pH	0.242**	0.810			-			
pH <sup>2</sup>	-				-			

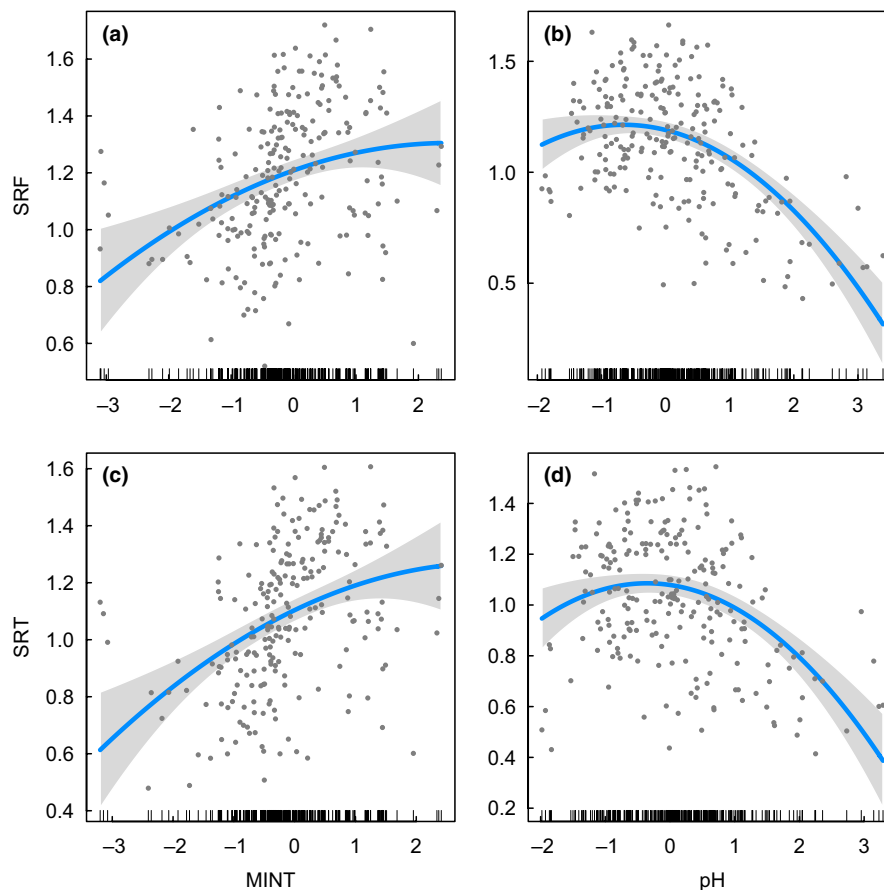
Note. Model sets involved all possible combinations of explanatory variables, for two groups of variables: Group 1: minimum temperature of the coldest month (MINT), mean annual precipitation (MAP). Group 2: MINT, soil pH (pH). Coefficients for the model with the lowest AIC<sub>c</sub> for a given variable group are shown. The Akaike weight for each variable is based on the full model set per group. The superscript 2 indicates the quadratic form of the variables. Pseudo- $r^2$ , explained the variance of the SAR model. Moran's  $I$ , measure of residual spatial autocorrelation.

<sup>a</sup>Significance levels: \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$ . *ns*, not significant.

explain this pattern. Theaceae's tolerance of low pH has been illustrated through cultivation studies of tea plants (*Camellia sinensis*), which have shown them to grow well in soils of low pH (appropriate 4.0–5.5) and a high Al concentration (Fu, Wang, & Ding, 2013). Notably, a high Al concentration has been found to stimulate tea plant growth (Konishi, Miyamoto, & Taki, 1985), but the same treatment would be toxic for most other plants (Fung & Wong, 2002).

The analysis of the whole family showed a strong positive relationship between NRI and pH (Table 5), which combined with the negative relationship between species richness and pH, provides support to the time and niche conservatism effect. However, no relationship was found between NRI and species richness (Table 4). This lacking species richness–NRI relationship may have resulted from the canceling out of different underlying relationships, caused by the following different processes: The processes of the time-for-speciation hypothesis may lead to an assemblage with many species and low NRI (Qian et al., 2015; Stephens & Wiens, 2003), whereas the processes of the diversification rates hypothesis could also lead to many species, but with high NRI (Qian et al., 2015).

Our results showed that there is a significant relationship between species richness and NRI for the Theaeae tribe (Table 4), but the U-shaped relationship is not consistent with either the pure time-for-speciation hypothesis or the diversification rate hypothesis. Some species-rich areas (species richness  $\geq 25$ ; Figure 3b) exhibited phylogenetically clustered patterns, while others tended to be phylogenetically overdispersed (Figure 3d). The species compositions of these two types of species-rich areas mainly included the *Camellia*, the *Pyrenaria* or the *Polyspora*. However, the proportions of *Pyrenaria* and/or *Polyspora* species ( $>19\%$ ) in areas with overdispersed phylogenetic structure were higher than in areas where phylogenetic structures were clustering ( $<17\%$ ). According to the locations of these species-rich areas, most of them were situated in one of three areas: the southwestern region bordering Vietnam, the Nanling Mountains and surrounding land in southcentral China, and the northeastern edge of the Yungui Plateau. In previous studies of endemic Chinese plants, these areas have been hypothesized to be former glacial refuges (Huang et al., 2011, 2015; López-Pujol, Zhang, Sun, Ying, & Ge, 2011). Hence, we suggest that the long-term



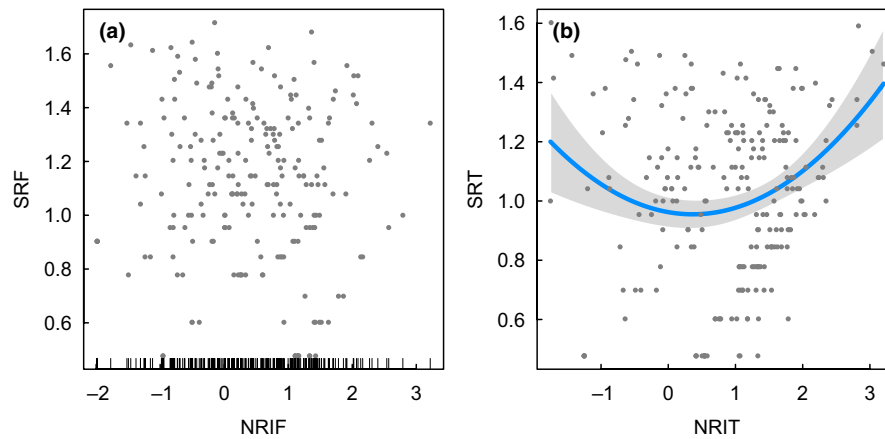
**FIGURE 4** Relationship between species richness and minimum temperature of the coldest month (MINT) or soil pH (pH) in the ordinary least squares (OLS) regression models. SRF indicates species richness at family level (i.e., Theaceae), and SRT indicates species richness at tribe level (i.e., Theaeae). Species richness responses to each predictor in multiple models are shown one at a time, holding all other predictors constant. The regression line is given in blue, and the 95% confidence interval is given in gray. Multimodel inference results are given in Supporting Information

stable environmental conditions in these refuge areas may have enabled them to harbor more species from clades that are distantly related via relictual survival, while others may primarily have harbored one clade, due to in situ diversification within the area. The presence of the above processes in species-rich areas should lead to U-shaped relationship.

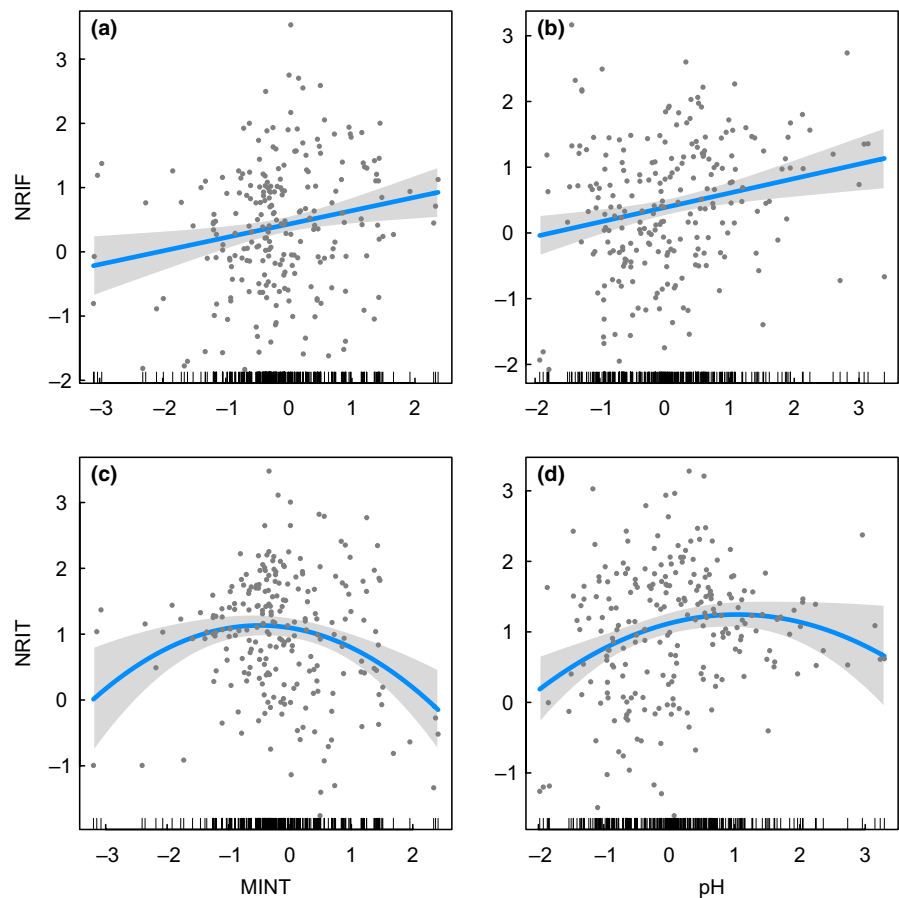
Our results show that processes underlying species richness patterns for tea family species in China differ between taxonomic levels. We suggest this might be because richness patterns at the family level are the sum of processes occurring within different clades. This would even out clade-specific idiosyncratic evolutionary patterns, whereas tribe-level patterns would reflect a tribe's specific evolutionary history and response to environmental factors (Bregovic & Zigmajster, 2016; Terribile, Olalla-Tarraga, Diniz, & Rodriguez, 2009). For example, *Stewartia*, which is the sole genus in Stewartieae, has both deciduous and evergreen species in China, whereas all species across China in Gordonieae and Theaeae are evergreen (Li, 2011; Li, Li, Tredici, Corajod, & Fu, 2013). Because Gordonieae and Stewartieae contributed considerably to the overall phylogenetic structure pattern found in our study, the phylogenetic structure's responses to environmental factors changed after removing these two

clades (Table 5). However, given their low richness, they contributed little to the overall pattern of tea family species richness. Therefore, the species richness–environmental relationships were similar at the family-wide and the tribe level (Table 3).

In conclusion, soil pH provides the strongest explanatory predictor for the geographic variation in species richness of Theaceae across China. This is likely linked to the Theaceae family's specific adaptations to acidic soil. This pattern contrasts the general assumption that soil only influences ecological patterns at small scales (Palpurina et al., 2017; Pärtel, 2002). Furthermore, the relationships between species richness and phylogenetic structure caused by different processes (time-for-speciation vs. diversification rate) may be canceling each other out at the family level, leading to no species richness–NRI relation at this taxonomic level. At the tribe level, the relationship between species richness and phylogenetic structure was significant, but more complex than predicted by time-for-speciation and diversification rate hypotheses. Some species-rich areas tended to host relatively distantly related species of Theaeae, while others exhibited phylogenetically clustered patterns. This is likely due to different refuges in southern China (López-Pujol et al., 2011) having played different roles in tea species' diversity. Some likely



**FIGURE 5** Relationship between species richness and net relatedness index (NRI) based on ordinary least squares (OLS) regression models. SRF and NRIF indicate species richness and NRI at family level (i.e., Theaceae), SRT and NRIT indicate species richness and NRI at tribe level (i.e., Theeae). The regression line is given in blue, and the 95% confidence interval is given in gray. The scatter plot, with no significant relationship, is also shown. Multimodel inference results are given in Supporting Information



**FIGURE 6** Relationship between net relatedness index (NRI) and minimum temperature of the coldest month (MINT) or soil pH (pH) based on ordinary least squares (OLS) regression models. NRIF indicates NRI at family level (i.e., Theaceae), and NRIT indicates NRI at tribe level (i.e., Theeae). The NRI responses to each predictor in multiple models are shown one at a time, holding all other predictors constant. The regression line is given in blue, and the 95% confidence interval is given in gray. Multimodel inference results are given in Supporting Information

allowed several deep lineages to survive, while others promoted the diversification of a single lineage. In addition, our results imply that the forces shaping species richness patterns vary among different groups and with taxonomic scale, even within the same family. Overall, our findings show that environmental and evolutionary processes interact in complex ways to shape species richness patterns within the subtropical forest biome.

## ACKNOWLEDGMENTS

This study was financially supported by National Key Technologies Research and Development Program (2017YFC0503803), the National Natural Science Foundation of China (31170494 and 31670212), the Strategic Priority Research Program of Chinese Academy of Sciences (XDPB020306), and the Chinese Scholarship Council (201506040162). Jens-Christian Svenning was supported

by the European Research Council (ERC-2012-StG-310886-HISTFUNC) and also considers this work a contribution to his VILLUM Investigator project "Biodiversity Dynamics in a Changing World" funded by VILLUM FONDEN (grant 16549) and the TREECHANGE project, funded by Danish Council for Independent Research | Natural Sciences (grant 6108-00078B). Manuel J. Steinbauer was supported by the Carlsberg Foundation (grant CF14-0148 to J.-C. S.) as well as the EU's Marie Skłodowska-Curie action (grant 707491).

## CONFLICT OF INTEREST

None declared.

## AUTHOR CONTRIBUTIONS

XCM, JTZ, JCS, and MDR conceived the ideas. MDR, JCS, and MJS designed the methodology. KPM and MGZ collected the data. MDR and XGX analyzed the data. All authors contributed to the manuscript text.

## DATA ACCESSIBILITY

Data available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.s526b27>.

## ORCID

Mide Rao  <http://orcid.org/0000-0002-1573-1207>

Xiangcheng Mi  <http://orcid.org/0000-0002-2971-5881>

Keping Ma  <http://orcid.org/0000-0001-9112-5340>

Jens-Christian Svenning  <http://orcid.org/0000-0002-3415-0862>

## REFERENCES

- Algar, A. C., Kerr, J. T., & Currie, D. J. (2009). Evolutionary constraints on regional faunas: Whom, but not how many. *Ecology Letters*, *12*, 57–65. <https://doi.org/10.1111/j.1461-0248.2008.01260.x>
- Azevedo, L. B., van Zelm, R., Hendriks, A. J., Bobbink, R., & Huijbregts, M. A. (2013). Global assessment of the effects of terrestrial acidification on plant species richness. *Environmental Pollution*, *174*, 10–15. <https://doi.org/10.1016/j.envpol.2012.11.001>
- Barton, K. (2015). *MuMIn: Multi-model inference*. R package version 1.15.1. Retrieved from <https://cran.r-project.org/web/packages/MuMIn>
- Bivand, R., Altman, M., Anselin, L., Assunção, R., Berke, O., Bernat, A., ... Yu, D. L. (2015). *spdep: Spatial dependence: Weighting schemes, statistics and models*. R package version 0.5-92. Retrieved from <http://CRAN.R-project.org/package=spdep>.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., & Suchard, M. A. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*, e1003537.
- Bregovic, P., & Zgajster, M. (2016). Understanding hotspots within a global hotspot—Identifying the drivers of regional species richness patterns in terrestrial subterranean habitats. *Insect Conservation and Diversity*, *9*, 268–281. <https://doi.org/10.1111/icad.12164>
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference—A practical information-theoretic approach*. New York, NY: Springer.
- Cardillo, M., Orme, C. D. L., & Owens, I. P. F. (2005). Testing for latitudinal bias in diversification rates: An example using New World birds. *Ecology*, *86*, 2278–2287. <https://doi.org/10.1890/05-0112>
- Cohen, K. M., Finney, S. C., Gibbard, P. L., & Fan, J. X. (2013). The ICS international chronostratigraphic chart. *Episodes*, *36*, 199–204.
- Diniz-Filho, J. A. F., Bini, L. M., & Hawkins, B. A. (2003). Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, *12*, 53–64. <https://doi.org/10.1046/j.1466-822X.2003.00322.x>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carre, G., ... Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Dubuis, A., Giovanettina, S., Pellissier, L., Pottier, J., Vittoz, P., & Guisan, A. (2013). Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. *Journal of Vegetation Science*, *24*, 593–606. <https://doi.org/10.1111/jvs.12002>
- Dutilleul, P. (1993). Modifying the T-test for assessing the correlation between 2 spatial processes. *Biometrics*, *49*, 305–314.
- Fang, J. Y., & Yoda, K. (1991). Climate and vegetation in China V. Effect of climatic factors on the upper limit of distribution of evergreen broad-leaf forest. *Ecological Research*, *6*, 113–125. <https://doi.org/10.1007/BF02353874>
- Farris, J. S., Källersjö, M., Kluge, A. G., & Bult, C. (1994). Testing significance of incongruence. *Cladistics*, *10*, 315–319. <https://doi.org/10.1111/j.1096-0031.1994.tb00181.x>
- Field, R., Hawkins, B. A., Cornell, H. V., Currie, D. J., Diniz-Filho, J. A. F., Guegan, J. F., ... Turner, J. R. G. (2009). Spatial species-richness gradients across scales: A meta-analysis. *Journal of Biogeography*, *36*, 132–147. <https://doi.org/10.1111/j.1365-2699.2008.01963.x>
- Francis, A. P., & Currie, D. J. (2003). A globally consistent richness-climate relationship for angiosperms. *The American Naturalist*, *161*, 523–536. <https://doi.org/10.1086/368223>
- Fu, D. L., Wang, Y., & Ding, Z. T. (2013). Effects of ferrous sulfate with sulfuric acid on tea (*Camellia sinensis* (L.) O. Kuntz) soil environment. *Soil Science and Plant Nutrition*, *59*, 783–789.
- Fung, K. F., & Wong, M. H. (2002). Effects of soil pH on the uptake of Al, F and other elements by tea plants. *Journal of the Science of Food and Agriculture*, *82*, 146–152. <https://doi.org/10.1002/jsfa.1007>
- Grote, P. J., & Dilcher, D. L. (1989). Investigations of angiosperms from the eocene of north america a new genus of Theaceae based on fruit and seed remains. *Botanical Gazette*, *150*, 190–206. <https://doi.org/10.1086/337764>
- Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/ 98/NT. *Nucleic Acids Symposium Series*, *41*, 95–98.
- Hawkins, B. A., Diniz-Filho, J. A. F., Mauricio Bini, L., De Marco, P., & Blackburn, T. M. (2007). Red herrings revisited: Spatial autocorrelation and parameter estimation in geographical ecology. *Ecography*, *30*, 375–384. <https://doi.org/10.1111/j.0906-7590.2007.05117.x>
- Hawkins, B. A., Field, R., Cornell, H. V., Currie, D. J., Guegan, J. F., Kaufman, D. M., ... Turner, J. R. G. (2003). Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, *84*, 3105–3117. <https://doi.org/10.1890/03-8006>
- Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., ... Gonzalez, M. R. (2014). SoilGrids1km—Global soil information based on automated mapping. *PLoS One*, *9*, e105992.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978. <https://doi.org/10.1002/joc.1276>

- Huang, J., Chen, B., Liu, C., Lai, J., Zhang, J., & Ma, K. (2011). Identifying hotspots of endemic woody seed plant diversity in China. *Diversity and Distributions*, *18*, 673–688. <https://doi.org/10.1111/j.1472-4642.2011.00845.x>
- Huang, Y. J., Jacques, F. M. B., Su, T., Ferguson, D. K., Tang, H., Chen, W. Y., & Zhou, Z. K. (2015). Distribution of Cenozoic plant relicts in China explained by drought in dry season. *Scientific Reports*, *5*, 14212. <https://doi.org/10.1038/srep14212>
- Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., ... Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, *26*, 1463–1464. <https://doi.org/10.1093/bioinformatics/btq166>
- 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. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 8125–8130. <https://doi.org/10.1073/pnas.1308932111>
- Kissling, W. D., & Carl, G. (2008). Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, *17*, 59–71.
- Konishi, S., Miyamoto, S., & Taki, T. (1985). Stimulatory effects of aluminum on tea plants grown under low and high phosphorus supply. *Soil Science and Plant Nutrition*, *31*, 361–368. <https://doi.org/10.1080/00380768.1985.10557443>
- Li, J., He, Q. X., Hua, X., Zhou, J., Xu, H. D., Chen, J. K., & Fu, C. Z. (2009). Climate and history explain the species richness peak at mid-elevation for Schizothorax fishes (Cypriniformes: Cyprinidae) distributed in the Tibetan Plateau and its adjacent regions. *Global Ecology and Biogeography*, *18*, 264–272.
- Li, M. M., Li, J. H., Tredici, P. D., Corajod, J., & Fu, C. X. (2013). Phylogenetics and biogeography of Theaceae based on sequences of plastid genes. *Journal of Systematics and Evolution*, *51*, 396–404. <https://doi.org/10.1111/jse.12017>
- Li, M. M. (2011). Phylogenetic and biogeographic studies of Theaceae with a special reference to Stewartia. (PhD thesis) Zhejiang University, Hangzhou, China.
- Lieb, A. M., Darrouzet-Nardi, A., & Bowman, W. D. (2011). Nitrogen deposition decreases acid buffering capacity of alpine soils in the southern Rocky Mountains. *Geoderma*, *164*, 220–224. <https://doi.org/10.1016/j.geoderma.2011.06.013>
- López-Pujol, J., Zhang, F. M., Sun, H. Q., Ying, T. S., & Ge, S. (2011). Centres of plant endemism in China: Places for survival or for speciation? *Journal of Biogeography*, *38*, 1267–1280. <https://doi.org/10.1111/j.1365-2699.2011.02504.x>
- Luna, I., & Ochoterena, H. (2004). Phylogenetic relationships of the genera of Theaceae based on morphology. *Cladistics*, *20*, 223–270. <https://doi.org/10.1111/j.1096-0031.2004.00024.x>
- Marin, J., & Hedges, S. B. (2016). Time best explains global variation in species richness of amphibians, birds and mammals. *Journal of Biogeography*, *43*, 1069–1079. <https://doi.org/10.1111/jbi.12709>
- McGill, B. J. (2010). Matters of scale. *Science*, *328*, 575–576. <https://doi.org/10.1126/science.1188528>
- Ming, T. L. (2000). *Monograph of the genus Camellia*. Kunming: Yunnan Science and Technology Press.
- Ming, T. L., & Bartholomew, B. (2007). Theaceae. In Z. Y. Wu, P. H. Raven, & D. Y. Hong (Eds.), *Flora of China* (pp. 366–429). Beijing: Science Press. Beijing.
- Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., ... McCain, C. M. (2007). Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. *Ecology Letters*, *10*, 315–331. <https://doi.org/10.1111/j.1461-0248.2007.01020.x>
- Nishii, K., Hughes, M., Briggs, M., Haston, E., Christie, F., DeVilliers, M. J., ... Möller, M. (2015). Streptocarpus redefined to include all Afro-Malagasy Gesneriaceae: Molecular phylogenies prove congruent with geography and cytology and uncovers remarkable morphological homoplasies. *Taxon*, *64*, 1243–1274.
- Osorio, F., & Vallejos, R. (2014). *SpatialPack: Package for analysis of spatial data*. R package version 0.2-3. Retrieved from <http://cran.r-project.org/web/packages/SpatialPack/index.html>.
- Palpurina, S., Wagner, V., von Wehrden, H., Hájek, M., Horsák, M., Brinkert, A., ... Danihelka, J. (2017). The relationship between plant species richness and soil pH vanishes with increasing aridity across Eurasian dry grasslands. *Global Ecology and Biogeography*, *26*, 425–434. <https://doi.org/10.1111/geb.12549>
- Pärtel, M. (2002). Local plant diversity patterns and evolutionary history at the regional scale. *Ecology*, *83*, 2361–2366. <https://doi.org/10.2307/3071796>
- Pausas, J. G., & Austin, M. P. (2001). Patterns of plant species richness in relation to different environments: An appraisal. *Journal of Vegetation Science*, *12*, 153–166. <https://doi.org/10.2307/3236601>
- Posada, D., & Crandall, K. A. (1998). Modeltest: Testing the model of DNA substitution. *Bioinformatics*, *14*, 817–818. <https://doi.org/10.1093/bioinformatics/14.9.817>
- Prince, L. M. (2009). Theaceae. In Flora of North America Editorial Committee (Ed.) *Flora of North America North of Mexico* (pp. 322–328). New York: Oxford University Press.
- Pyron, R. A., & Wiens, J. J. (2013). Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proceedings of the Royal Society B: Biological Sciences*, *280*, 20131622. <https://doi.org/10.1098/rspb.2013.1622>
- Qian, H., Wiens, J. J., Zhang, J., & Zhang, Y. J. (2015). Evolutionary and ecological causes of species richness patterns in North American angiosperm trees. *Ecography*, *38*, 241–250. <https://doi.org/10.1111/ecog.00952>
- Qian, H., Zhang, Y. J., Zhang, J., & Wang, X. L. (2013). Latitudinal gradients in phylogenetic relatedness of angiosperm trees in North America. *Global Ecology and Biogeography*, *22*, 1183–1191. <https://doi.org/10.1111/geb.12069>
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>.
- Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. (2014). *Tracer v1.6*. Retrieved from <http://beast.bio.ed.ac.uk/Tracer>.
- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*, 217–223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>
- Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, *7*, 1–15. <https://doi.org/10.1046/j.1461-0248.2003.00554.x>
- Sakai, A., & Weiser, C. J. (1973). Freezing resistance of trees in North America with reference to tree regions. *Ecology*, *54*, 118–126. <https://doi.org/10.2307/1934380>
- Song, Y. C., Chen, X. Y., & Wang, X. H. (2005). Studies on evergreen broad-leaved forests of China: A retrospect and prospect (in Chinese with English abstract). *Journal of East China Normal University (Nature Sciences)*, *1*, 1–8.
- Stephens, P. R., & Wiens, J. J. (2003). Explaining species richness from continents to communities: The time-for-speciation effect in emydid turtles. *The American Naturalist*, *161*, 112–128.
- Svenning, J. C., Borchsenius, F., Bjorholm, S., & Balslev, H. (2008). High tropical net diversification drives the New World latitudinal gradient in palm (Arecaceae) species richness. *Journal of Biogeography*, *35*, 394–406. <https://doi.org/10.1111/j.1365-2699.2007.01841.x>
- Swofford, D. L. (2003). *PAUP\*: Phylogenetic analysis using parsimony, version 4.0 b10*. Sunderland, MA: Sinauer.
- Terribile, L. C., Olalla-Tarraga, M. A., Diniz, J. A. F., & Rodriguez, M. A. (2009). Ecological and evolutionary components of body size: Geographic variation of venomous snakes at the global scale.



- Biological Journal of the Linnean Society*, 98, 94–109. <https://doi.org/10.1111/j.1095-8312.2009.01262.x>
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The CLUSTAL\_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 25, 4876–4882. <https://doi.org/10.1093/nar/25.24.4876>
- 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. <https://doi.org/10.1111/j.1365-294X.2008.04001.x>
- Wang, Z. H., Fang, J. Y., Tang, Z. Y., & Lin, X. (2011). Patterns, determinants and models of woody plant diversity in China. *Proceedings of the Royal Society B: Biological Sciences*, 278, 2122–2132.
- Wang, X. H., Kent, M., & Fang, X. F. (2007). Evergreen broad-leaved forest in Eastern China: Its ecology and conservation and the importance of resprouting in forest restoration. *Forest Ecology and Management*, 245, 76–87. <https://doi.org/10.1016/j.foreco.2007.03.043>
- Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. *The American Naturalist*, 156, 11. <https://doi.org/10.1086/303378>
- Wiens, J. J. (2011). The causes of species richness patterns across space, time, and clades and the role of "ecological limits". *Quarterly Review of Biology*, 86, 75–96. <https://doi.org/10.1086/659883>
- Wiens, J. J., Ackerly, D. D., Allen, A. P., Anacker, B. L., Buckley, L. B., Cornell, H. V., ... Stephens, P. R. (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, 13, 1310–1324. <https://doi.org/10.1111/j.1461-0248.2010.01515.x>
- Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, 19, 639–644. <https://doi.org/10.1016/j.tree.2004.09.011>
- 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. *The American Naturalist*, 168, 579–596. <https://doi.org/10.1086/507882>
- Wiens, J. J., Parra-Olea, G., Garcia-Paris, M., & Wake, D. B. (2007). Phylogenetic history underlies elevational biodiversity patterns in tropical salamanders. *Proceedings of the Royal Society B: Biological Sciences*, 274, 919–928. <https://doi.org/10.1098/rspb.2006.0301>
- Wikstrom, N., Kainulainen, K., Razafimandimbison, S. G., Smedmark, J. E. E., & Bremer, B. (2015). A revised time tree of the Asterids: Establishing a temporal framework for evolutionary studies of the coffee family (Rubiaceae). *PLoS One*, 10, e0126690.
- Wilson, M., He, Z. L., & Yang, X. E. (2004). *The red soils of China: Their nature, management and utilization*. Dordrecht, the Netherlands: Springer Press.
- Wu, C. Y. (1995). *Vegetation of China*. Beijing, China: Science Press.
- Wu, Z. Y., & Wu, S. (1998). *A proposal for a new floristic kingdom (realm): the E. Asiatic Kingdom, its delineation and characteristics. Floristic characteristics and diversity of East Asian plants: proceedings of the first international symposium of floristic characteristics and diversity of East Asian plants*. Beijing: Springer Verlag, China Higher Education Press.
- Xiong, Y., & Li, Q. (1987). *Soils in China*. Beijing: Science Press.
- Yu, X. Q., Gao, L. M., Soltis, D. E., Soltis, P. S., Yang, J. B., Fang, L., ... Li, D. Z. (2017). Insights into the historical assembly of East Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family. *New Phytologist*, 215, 1235–1248. <https://doi.org/10.1111/nph.14683>
- Zellweger, F., Baltensweiler, A., Ginzler, C., Roth, T., Braunisch, V., Bugmann, H., & Bollmann, K. (2016). Environmental predictors of species richness in forest landscapes: Abiotic factors versus vegetation structure. *Journal of Biogeography*, 43, 1080–1090. <https://doi.org/10.1111/jbi.12696>
- Zhang, W., Kan, S. L., Zhao, H., Li, Z. Y., & Wang, X. Q. (2014). Molecular phylogeny of tribe Theeae (Theaceae s.s.) and its implications for generic delimitation. *PLoS One*, 9, e98133.
- Zhu, H. (2013). Geographical elements of seed plants suggest the boundary of the tropical zone in China. *Palaeogeography Palaeoclimatology Palaeoecology*, 386, 16–22. <https://doi.org/10.1016/j.palaeo.2013.04.007>
- Zhu, Y., Zhao, G. F., Zhang, L. W., Shen, G. C., Mi, X. C., Ren, H. B., ... Ma, K. P. (2008). Community composition and structure of Gutianshan forest dynamic plot in a mid-subtropical evergreen broad-leaved forest, east China. *Chinese Journal of Plant Ecology*, 32, 262–273. (in Chinese).

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Rao M, Steinbauer MJ, Xiang X, et al. Environmental and evolutionary drivers of diversity patterns in the tea family (Theaceae s.s.) across China. *Ecol Evol*. 2018;8:11663–11676. <https://doi.org/10.1002/ece3.4619>