New insights into the formation of biodiversity hotspots of the Kenyan flora

Qiang Zhang1,2 | Jian-Fei Ye3 | Chi-Toan Le4 | Dennis Mwithukia Njenga1,2 | Narinda Romer Rabariaona1,2 | Wyckliffe Omondi Omollo1,2 | Li-Min Lu1 | Bing Liu1,5 | Zhi-Duan Chen1,5

Abstract
Aim: This study aimed to investigate the distribution patterns of plant diversity in Kenya, how climatic fluctuations and orogeny shaped them, and the formation of its β-diversity.
Location: Kenya, East Africa.
Taxon: Angiosperms.
Methods: We quantified patterns of turnover and nestedness components of phylogenetic β-diversity for angiosperm species among neighbouring sites using a well-resolved phylogenetic tree and extensive distribution records from public databases and other published sources. We applied clustering methods to delineate biota based on pairwise similarities among multiple sites and used a random assembly null model to assess the effects of species abundance distribution on phylogenetic β-diversity.
Results: The phylogenetic turnover of the Kenyan flora, intersecting with the biodiversity hotspots Eastern Afromontane, Coastal Forests of Eastern Africa, and Horn of Africa, shows a non-monotonic pattern along a latitudinal gradient that is strongly structured into volcanic and coastal areas. The other areas are mainly dominated by phylogenetic nestedness, even in the eastern part of the equatorial region parallel to the volcanic area. Phylogenetic diversity and phylogenetic structure analyses explain the mechanism of the observed phylogenetic turnover and nestedness patterns. We identified five phytogeographical regions in Kenya: the Mandera, Turkana, Volcanic, Pan Coastal and West Highland Regions.
Conclusions: Variations in turnover gradient and coexistence are highly dependent on the regional biogeographical history resulting from climatic fluctuations and long-lasting orogeny, which jointly shaped the biodiversity patterns of the Kenyan flora. The nestedness component dominated climatically unstable regions and is presumed to have been caused by heavy local species extinction and recolonization from the
1 | INTRODUCTION

Kenya intersects three biodiversity hotspots, that is, Eastern Afromontane, Coastal Forests of Eastern Africa and Horn of Africa (Marchese, 2015). The complex topography is marked by large-scale extensional features, such as the East African Rift, widespread volcanic activity, and anomalously subsided basins and uplifted domes, which profoundly influenced the distribution and amount of rainfall under control by the African-Indian monsoon and westerly moisture sources, resulting in a richness of species (Couvreur et al., 2021; Griffiths, 1993; Wichura et al., 2010). These geological processes also led to stronger temperature and altitude gradients from the coastal plains to the uplands of the East African Plateau, which created an extensive mountainous area with cooler temperatures (Linder, 2017). Accompanied by the paleoclimatic fluctuations, such as the Last Glacial Maximum, the extensive mountainous area finally created a patchwork of closely spaced humid and arid sedimentary environments that support diverse vegetation and became centres of biodiversity and endemism (Dagallier et al., 2020; de Menocal, 1995; Linder, 2001; Maslin et al., 2014; Pierre et al., 2006).

It is well known that orogeny has promoted topographic heterogeneity and created new habitats for species to evolve and diversify (Antonelli et al., 2018). Many studies have argued that mountains are both cradles and museums and are sources of new evolutionary lineages that later colonized surrounding lowlands or nearby alpine mountains (Antonelli et al., 2009; Ding et al., 2020; Hutter et al., 2017). Irrespective of any focus on biological processes, the mountain-geobiodiversity hypothesis was proposed to explain the high levels of biodiversity in mountain regions (Muellner-Riehl et al., 2019). The mountain-geobiodiversity hypothesis is based on three conditions: (a) the presence of lowland, montane and alpine zones (areas with a high diversity of niches); (b) climatic fluctuations for a "species pump" effect (drivers of species divergence); and (c) high-relief terrain with full elevational zonation (forming candidate refugia regions and easily migrated regions; Muellner-Riehl et al., 2019). Biotic and abiotic factors may act in parallel to biodiversity patterns, even though we consider them separately according to different hypotheses.

It is challenging to account for both ecological and evolutionary processes when depicting biodiversity patterns (Graham & Fine, 2008; Saladin et al., 2019). In the face of this challenge, a widely recognized approach is to incorporate phylogenies into biodiversity analyses (Helms et al., 2007; Miles & Dunham, 1993; Pagel, 1999; Swenson, 2013). β-diversity interacts with α-diversity gradients, and both result in community assembly through local and regional filters. Thus, β-diversity may capture the dynamic nature of biodiversity patterns better than simply measuring α-diversity (Soininen et al., 2018). Turnover and nestedness are two components of β-diversity. The former refers to β-diversity attributable to species replacement, whereas the latter indicates species loss or gain and relates to richness differences between the communities (Baselga, 2010). Phylogenetic β-diversity is usually applied to measure the dissimilarity of phylogenetic composition between different assemblages (Anderson et al., 2011; Baselga, 2012). It considers phylogenetic relatedness to explain biota evolutionary history and biodiversity patterns across environmental gradients on regional or global scales (Dobrovolski et al., 2012). Understanding the processes that organize biodiversity in time and space helps design robust protected area networks (Socolar et al., 2016).

To date, there is no specific research on the delineation of the Kenyan flora. According to White (1983), most areas of Kenya between the highlands and the coastal belt are located in the Somali-Masai region; the remaining parts belong to the Afromontane archipelago and the Lake Victoria regions. Under the bioregionalization framework of the African continent, Linder et al. (2012) suggested that East Africa (from Kivu of the Democratic Republic of Congo, Uganda, Kenya, to northcentral Tanzania) is a large Regional Mosaic of intermingling vegetation types with distinct floras. This region falls into the transition zone between the dry north and south and the humid west (Fayolle et al., 2019; Linder et al., 2012). Other classifications roughly delineated Kenya into the East African montane region and the Coastal Region (Droissart et al., 2018; Fayolle et al., 2014). The African montane region was formed by a continental magmatic arc of hundreds of volcanic structures with a north–south orientation, from the southern margin of the Turkana basin to Kajado (Baker et al., 1971; Feibel, 2011; Global Volcanism Program, 2013; Hetzel & Strecke, 1994; McCall & Bristow, 1965; Williams, 1970; Williams et al., 1984). It is crucial to measure its spatial phylogeny and dissect its flora. Evolutionary biodiversity studies of the Trans-Mexican Volcanic Belt suggested that climatic fluctuations, together with recent volcanism, have driven the diversification and local persistence of biodiversity in the Mexican highlands (Mastretta-Yanes et al., 2015; Rodríguez et al., 2018; Ruiz-Sánchez et al., 2013; Sosa et al., 2018). We expect that there may be areas with strong turnover along the precipitation gradient extending from the volcanic to the coastal area, which is consistent with the delimitation of dry and wet habitats. The volcanic area is mainly distributed in western to central Kenya.
and coincides with the Eastern Afromontane biodiversity hotspot, while the coastal area in eastern Kenya overlaps with the Coastal Forests of Eastern Africa hotspot. As the environments are not functionally equivalent to determining community composition (Gomez et al., 2019), we hypothesized that high phylogenetic turnover was found in volcanic and coastal areas, whereas low phylogenetic turnover prevailed in dry areas.

Here, we investigated the patterns of phylogenetic turnover and identified the phytogeographical regions of Kenya based on the spatial phylogenetic analysis of angiosperms. The aims of our study were as follows: (a) to delineate the phytogeographical regions of Kenya; (b) to test whether high phylogenetic turnover component occurs in the volcanic area; and (c) to investigate the underlying mechanisms that shaped the β-diversity patterns in the biodiversity hotspots of East Africa.

2 | METHODS

2.1 | Distribution records

The comprehensive angiosperm checklist of Kenya was compiled based on the data from the Flora of Tropical East Africa (FTEA; Zhou et al., 2017). We gathered detailed distribution information of each species from FTEA, local checklists and other published sources (Agnew & Agnew, 1994; Fischer et al., 2011; FTEA editors, 1952–2012; Girma et al., 2015; Luke, 2005) and then transformed them into county-level distributions following the contemporary administrative divisions of Kenya. In addition, we collected the species occurrence data of angiosperms from RAINBIO (Dauby et al., 2016) and the Global Biodiversity Information Facility (GBIF, https://doi.org/10.15468/dl.njawff). We used the R package “speciesgeocodeR” (Töpel et al., 2017) to filter the species occurrence in the datasets, checking for obvious errors such as empty coordinates, terrestrial species reported in the sea and points located outside the country boundary of Kenya. Information on species growth forms were defined according to expert experience. The list of species growth forms is shown in Appendix S1. All the species names were standardized following The Plant List version 1.1 (available at http://www.theplantlist.org) by the R package “Taxonstand” (Cayuela et al., 2012). After excluding alien and non-angiosperm species and filtering out species not in the checklist of Kenya, there were 36,567 coordinate occurrence records including longitude and latitude from RAINBIO and GBIF, and 15,997 county-level records from FTEA and other consulted sources. All the distribution records included 5822 species belong to 191 families and 1465 genera. There were 1843 species uniquely from FTEA and other sources and only 29 species uniquely from GBIF and RAINBIO (Table S4.2). Since some small geographical gaps in specimen collection still exist in the northeast parts of Kenya, we reduced the sampling bias caused by unequal sampling areas by up-scaling the distribution records, testing both 50-km×50-km and 100-km×100-km grid cells. We assigned the county-level distribution data to the overlapping grid cells (Details for other conversion method were provided in Appendix S4, Figure S4.2 and Table S4.1. Species list completeness assessment was provided in Appendix S9, Table S9.5 and Figure S9.13), and occurrence data were projected into the grid cells directly. Grid cells were excluded if the land area occupied <50% of the entire cell (i.e. <1250km²). Because the distribution data consisted of presence-only data, we transferred the species into presence-absence data to extend their applicability (Appendix S4). Following these steps, we generated a species-level presence-absence matrix for the remaining 239 grid cells.

2.2 | Phylogenetic analyses and species tree

2.2.1 | Phylogenetic analyses

We sampled all available taxa for four plastid genes (atpB, matK, ndhF and rbcL) to reconstruct the phylogenetic tree of angiosperms. We downloaded all available sequences of 2583 species for the target DNA regions from GenBank (last download on January 11, 2019). If two or more sequences were available for the same locus of a species, we kept the longest sequence. Moreover, we generated 1910 new sequences representing 815 species for which the sequence was unavailable or of low quality in the public database. At least one copy of the specimen was deposited in the Herbarium of the Institute of Botany, Chinese Academy of Sciences (Herbarium Code: PE) for all the samples collected in the field. The methods for DNA extraction and amplification followed Chen's protocol (Chen, Yang, et al., 2016). The alignment was conducted with MAFFT (Kato et al., 2019), and the relatively fast-evolving markers matK and ndhF were further refined with GENEIOUS R9 v.9.0.5 (Biomatters Ltd., Auckland, New Zealand; Karse et al., 2012).

The matrix obtained by the above steps was used to reconstruct a preliminary phylogenetic tree using RaxML 8.2.12 (Stamatakis, 2014) on the CIPRES Science Gateway Web server (Miller et al., 2010) with 1000 bootstrap replicates under the GTR+CAT model. In total, 11 gymnosperm species were used as outgroups in the phylogenetic analyses (i.e. Afrocarpus gracilior, A. mannii, Encephalartos hildebrandtii, E. kisambii, E. villosus, Ginkgo biloba, Juniperus coxii, J. gamboana, Pinus leiophylla var. chihuahuana, P. ponderosa var. scopulorum and Podocarpus milanjianus). We checked our preliminary phylogeny by comparing the single gene trees and combining them with those of previous studies to identify species with probable inappropriate systematic positions. For species with unexpected positions, we determined whether any of the sequences were chimeric or incorrectly identified using the BLAST tool in GenBank (Altschul et al., 1990). We removed low-quality sequences, duplicates and ambiguous taxonomical samples, but kept the species at provisional status.

The final alignment included 3363 species, representing four gene regions and a total length of 8201 base pairs (bp). This combined matrix was divided into four partitions, and run on RaxML 8.2.12 (Stamatakis, 2014) with independent parameters on each
model. The analysis comprised 1000 bootstrap replicates performed using the GTR+GAMMA model of nucleotide substitution.

2.2.2 | Time calibration

We used a penalized likelihood method as implemented in treePL (Smith & O’Meara, 2012) to conduct the time calibration of our phylogeny. Node calibrations were estimated using 72 fossil calibration points scattered among the angiosperms (Lu et al., 2018). The hard maximum and minimum age for angiosperms was set at 140 and 125 Myr, respectively (Lu et al., 2018). At the beginning, we set the smooth parameter free as 1000, and the “prime” option was performed to determine the optimal run parameter for the “thorough” option of the cross-validation analysis (opt = 1, optd = 1 and optc-vad = 1). Then, we obtained the best optimization parameters for smooth (smooth = 0.1). Finally, the analysis was carried out again with the optimal smooth, and a time tree including 3363 angiosperm species native to Kenya was reconstructed (Appendix S10).

2.2.3 | Species tree reconstruction

With the above time tree as the backbone, a species-level tree including 5912 angiosperm species from Kenya was generated by inserting non-sampled congeners species using the R package “V.PhyloMaker” (Jin & Qian, 2019). Our species-level tree covered almost 98.2% of the native and currently accepted species of Kenya (Appendix S5, Figure S5.3). This full tree was used for the phylogenetic diversity and phylogenetic β-diversity analyses.

Polytomies are particularly acute when missing tips are added to phylogenies based on taxonomic information by V.PhyloMaker. To evaluate the potential influence of polytomies within these analyses, the underlying taxonomic turnover differences between grid cells. Null distributions were used to calculate standardized effect sizes of pβsim (SES.pβsim), where the expectations of the null distribution were subtracted from the observed pβsim and divided by the standard deviation of the null distribution. A positive SES indicates higher phylogenetic turnover than expected based on a random sampling of the regional flora; a negative SES indicated the opposite.

The analysis was performed in R 3.6.1 (R Core Team, 2019) with the R package “betapart” (Baselga & Orme, 2012). Under the Sørensen dissimilarity, we chose pβsim and pβjtu as the indices to delineate biogeographical regions because they are weakly affected by differences in species richness between localities (Fabien & Anthi, 2014; Kreft & Jetz, 2010). To evaluate the sensitivity of our results, we also compared the pβsim and pβjtu values with the turnover component obtained from the Jaccard’s dissimilarity index (pβjtu; Baselga, 2012).

2.3 | β-diversity analysis

2.3.1 | Metrics of taxonomic and phylogenetic β-diversity

Following Baselga (2012), we calculated three measures of taxonomic β-diversity, the Sørensen dissimilarity index (pβsim), the Simpson dissimilarity index (pβsim) and the nestedness index (pβjtu) for each pair of grid cells. The total β-diversity was represented by pβjtu where pβsim was the species turnover part of β-diversity and pβjtu was the nestedness component of β-diversity. pβsim and its two components are defined as follows (Baselga, 2012):

$$p_{β\text{sim}} = \frac{(b + c)}{(2a + b + c)}$$
$$p_{β\text{jtu}} = \frac{\min(b, c)}{(a + \min(b, c))}$$
$$p_{β\text{sim}} = p_{β\text{jtu}} - p_{β\text{sim}}$$

where “a” is the number of shared species between two grid cells, and “b” and “c” are the number of species unique to each grid cell. pβsim ranges from 0 (if species composition is identical between grid cells) to 1 (if there are no shared taxa). When applied to phylogenetic β-diversity, shared and unique species are replaced with shared and unique branch lengths, respectively (Qian et al., 2020). Phylogenetic turnover (pβsim), which is adapted from pβsim, can be used to quantify the phylogenetic distance among grid cells (i.e. pairwise compositional turnover) (Hardy et al., 2012). pβsim ranges from 0 (if species compositions are identical between grid cells and hence share the same branch lengths) to 1 (if grids’ species compositions are distinct that share no similar branch lengths).

As pβsim is strongly correlated with pβsim, we constructed null models to calculate the random expectations of pβsim based on a pairwise approach. 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. Null distributions were used to calculate standardized effect sizes of pβsim (SES.pβsim), where the expectations of the null distribution were subtracted from the observed pβsim and divided by the standard deviation of the null distribution. A positive SES indicated higher phylogenetic turnover than expected based on a random sampling of the regional flora; a negative SES indicated the opposite.

We applied two approaches to investigate patterns of β-diversity across Kenya: a pairwise approach and a multiple-site approach. To explicitly consider the scale dependence of the results (Soininen et al., 2007), we used the closest three, four, eight and twelve grid cells adjacent to the focal grid (Figure S5.4). For the pairwise approach (based on phylo.beta.pair function), we obtained the dissimilarity values between the focal grid and all adjacent grid cells and assigned the average value for the focal grid cell. In the multiple-site approach (based on phylo.beta.multi function), a single value was obtained from all compared samples (all grid cells in the window), and
we assigned the single value to the focal grid cell. Since each site was a repeated sample in the pairwise approach, it was not independent; also, the pairwise approach ignores information on species shared among more than two sites. The multiple-site approach avoids the repetition of each site and prevents information loss when species appear in more than two sites (Diserud & Ødegaard, 2007). Besides, when the sites are viewed as a random set of observations from a region, evaluating overall similarity from the average of the pairwise similarities could be misleading (Diserud & Ødegaard, 2007). We herein compared two methods.

Based on the two datasets, we produced maps showing the relative importance of the turnover-resultant $p_{\text{beta}}$ component according to the formula:

$$p_{\text{beta}} = p_{\text{sim}} / p_{\text{sne}}$$

Values less than one indicated that $p_{\text{beta}}$ was mainly determined by nestedness-resultant dissimilarity ($p_{\text{sim}}$). In contrast, values greater than one indicated that $p_{\text{beta}}$ was the result of a greater influence of the dissimilarity component due to species replacement between grids ($p_{\text{sne}}$). The $p_{\text{beta}}$ for woody and herbaceous genera was calculated separately to compare their spatial distribution patterns across Kenya. The map of $p_{\text{beta}}$ based on the multiple-site approach on eight grid cells was then intersected with the map of nature reserves in Kenya (UNEP-WCMC and IUCN, 2022).

### 2.4 Cluster analysis

Pairwise similarity among multiple sites is an appropriate approach for displaying species turnover along an environmental gradient or geodistance (Penone et al., 2016). Following Ye et al. (2019, 2020), the grid cells were clustered by applying eight hierarchical methods to $p_{\text{sim}}$ and $p_{\text{sim}}$ pairwise dissimilarity matrices and applied the cophenetic correlation coefficient and Gower’s distance to access the ideal cluster method. All the analyses were implemented by the R package “stats” (R Core Team, 2019). An elbow criterion was employed to determine the optimum number of clusters, which is based on the two-level hybrid clustering approach (Zhao et al., 2011). This function was implemented by the “GMD” package (Zhao & Sandelin, 2014) in R. The analysis results showed that the unweighted pair-group method with arithmetic means (UPGMA) was the reasonable method (Table S4.3).

### 2.5 Ordination analysis

Non-metric multidimensional scaling (NMDS) was performed to illustrate the spatial turnover and relationships between the clusters in an alternative way. The method is robust at simplifying the variables of a multidimensional space into a low-dimensional space (Minchin, 1987). Grid cells in the NMDS ordination plots were coloured according to their UPGMA cluster membership to visualize the relative distance between geographical regions. The degree of NMDS ordination was assessed by the stress value, which ranged from 0 to 1. Smaller values indicate that the NMDS results are more consolidated. The analysis was carried out using the “vegan” package (Dixon, 2003).

### 2.6 Spatial phylogenetic analysis

#### 2.6.1 Phylogenetic diversity

Faith’s (1992) phylogenetic diversity (PD) is the sum of all phylogenetic branch lengths that connect species in a community. Considering that Faith’s PD is affected by species richness, we calculated the standardized effect size of PD (SES.PD) based on the taxa labels model (Lu et al., 2018). Both PD and SES.PD were calculated by species tree by the R package “picante” (Kembel et al., 2010).

#### 2.6.2 Phylogenetic structure

The net relatedness index (NRI) and the nearest taxon index (NTI) were calculated to investigate the phylogenetic structure (clustering or overdispersion) of angiosperms across Kenya (Webb et al., 2002). NRI is based on the mean phylogenetic distance (MPD), an estimate of the average phylogenetic relatedness between all possible pairs of taxa within a grid cell, and primarily reflects the structure in deeper parts of the phylogeny. NTI is based on the mean nearest taxon distance (MNTD), an estimate of the mean phylogenetic relatedness of the closest relative for each taxon within a grid cell, and reflects the terminal parts of the phylogeny. Positive values of NRI and NTI indicate phylogenetic clustering, whereas negative values indicate phylogenetic overdispersion in a grid cell. Both SES.MPD and SES.MNTD were calculated based on the full species tree by the R package “picante” (Kembel et al., 2010), then transformed to NRI and NTI by multiplying by −1.

### 2.7 Surrogate variable for volcanic influence

We investigated all occurrences, distribution and volume information of volcanos in Kenya (Guth, 2013). To investigate whether the $p_{\text{beta}}$ pattern is related to the distance between volcanoes and grid cells, we first computed the Euclidean distances between the focal grid cell and all volcanoes in the grid cells group (here, we used one focal and eight neighbouring grid cells group as research unit); second, we computed the distances between the focal grid cell and the nearest volcano (volcano not shown in the nine grid cells group); third, for the grid cell groups with no volcano, we computed the distances between the focal grid cell and the nearest volcano (volcano not shown in the nine grid cells group); finally, the reciprocal sum of the Euclidean distances for a focal grid cell was used as the independent variable representing volcanic influence on $p_{\text{beta}}$. The relationship between the reciprocal distance and $p_{\text{beta}}$ was modelled using a generalized linear model. We assumed that the closer the grid is to the volcano, the greater the impact of

\[ p_{\text{beta}} = \frac{p_{\text{sim}}}{p_{\text{sne}}} \]
the volcano on $p_{\beta_{\text{ratio}}}$. After calculating the inverse distance, we gave more weight to volcanoes closer to the focal grid. When the grid groups contain many volcanos, this processing method can effectively prevent a large distance sum causing a violation of our hypothesis.

3  |  RESULTS

3.1  |  Phytogeographical regionalization

The UPGMA clustering of grid cell assemblages based on both $p_{\beta_{\text{sim}}}$ and $\beta_{\text{sim}}$ dissimilarity matrices yielded five major phytogeographical regions in Kenya (Figures 1b, S6.5b and Appendix S6). We named the regions considering their environmental characteristics: (1) Mandera Region, (2) West Highland Region, (3) Pan Coastal Region, (4) Turkana Region and (5) Volcanic Region. The representative plants of each phytogeographical region can be seen in Figure 2.

The two regionalization schemes derived from taxonomic and phylogenetic approaches in this study are broadly similar, with the exception of a noticeable difference in the size of the Turkana Region and the Mandera Region. According to the phylogenetic approaches, the Turkana Region has a large area, extending from west to east in northern Kenya and broadly including Turkana County, Marsabit County and Wajir County. In the dendrogram derived from the $\beta_{\text{sim}}$ dissimilarity matrix, however, Wajir County was belonged to the Mandera Region. In addition, the relationships among the five regions based on $p_{\beta_{\text{sim}}}$ and $\beta_{\text{sim}}$ matrix were totally different (Figures 1a and S6.4a).

The clustered regions were also separated in the NMDS plot (Figures 1c and S6.4c). The stress values were relatively low ($p_{\beta_{\text{sim}}}$: 0.2155; $p_{\beta_{\text{sim}}}$: 0.265), which indicated a relatively good projection of the dissimilarity matrices. The change in floristic composition in northern Kenya was discontinuous: the Mandera and Turkana regions were distinctly separated from each other. Conversely, there was a gradual transition in floristic composition in southern Kenya. In particular, grid cells belonging to the West Highland Region overlapped with those belonging to the Volcanic Region.

3.2  |  Patterns of phylogenetic $\beta$-diversity

The Jaccard and Sørensen indices produced similar results. They all showed significant correlations ($p<0.001$), with correlation coefficients all above .91 (Table S4.4). We present the results for pairwise and multiple-site approaches with the Sørensen index,
although both approaches resulted in a similar pattern. Patterns of $p_{\beta\text{ratio}}$ and $p_{\beta\text{sim}}$ produced by different window scale definitions were similar (Appendix S7, Figures S7.6 and S7.7). To avoid bias in grid selection and unequal connection opportunities between focal and nearby grid cells, we chose a moderate-scale (the closest eight grid cells adjacent to the focal grid, Figure S5.4c) map in this study. Analysing $p_{\beta\text{ratio}}$ and $p_{\beta\text{sim}}$ with a map for other scales showed similar patterns (Figure S7.8). The standardized effect size (SES) values of $p_{\beta\text{sim}}$ were positive in 227 grid cells and negative in three grid cells (Figure S7.9a). The null model test revealed that the volcanic, coastal and Mandera regions exhibited significantly higher phylogenetic turnover than expected (SES values > 1.96). The moderate phylogenetic turnover detected by the null model test occurred in the rest dry regions (0 < SES values < 1.96).

The $p_{\beta\text{ratio}}$ was found to be significantly associated with the distance between a grid and a volcano, irrespective of growth forms (Figure 3). The effect of the presence of a volcano on phylogenetic $\beta$-diversity decreased with the increase in the distance between grid...
and volcano. The observed $p_β_{ratio}$ showed a general monotonic decline with increasing latitude in the north of the equator on pairwise and multiple-site approaches and growth forms (woody and herbaceous; Figure S7.10). The phylogenetic turnover peaked at the equator and at latitude 4°S, which corresponds to the volcanic and coastal area. The lowest phylogenetic turnover was represented by the Turkana Basin (Figure S7.10a–c). In Kenya, the main volcano distribution area (the Western part of Kenya between 2°S and 2°N) has a distinctly different topography than the eastern part (Figure S7.10a–c). These results from the V.PhyloMaker tree were consistent with those based on the TACT tree (Figure S7.11), suggesting that polytomies did not significantly bias our findings. The proportion of protected areas in the volcanic region was higher than in other regions (Figure 5). These nature reserves are fragmented and the larger contiguous areas with nature reserves are distributed in coastal areas (Taita Hills).

### 3.3 Spatial phylogenetic pattern

We obtained a series of spatial phylogenetic patterns, including species richness (SR), phylogenetic diversity (PD), standard effective size of phylogenetic diversity (SES.PD), NRI and NTI (Figure 4). The patterns of SR and PD were similar, presenting higher values in the volcanic and coastal areas (Figure 4a,b). The higher values of SES.PD were in the volcanic area surrounding Mount Kenya and the coastal area around Taita Hills and scattered in the west of Kenya (Figure 4c,d). There was an apparent difference between the spatial patterns of NRI and NTI. The negative values of NRI were mainly in the southern part near the sea, whereas those of NTI were mainly in the volcanic area (Figure 4d,e). The remaining parts of Kenya mainly showed positive values for NRI and NTI. Spatial analyses based on the TACT tree, except the NTI, obtained consistent results with those on the V.PhyloMaker tree.
Based on the TACT tree, a more significant phylogenetic overdispersion of NTI was detected in the volcanic area (Figure S8.12e).

4 | DISCUSSION

Our study is the first quantitative synthesis of $\beta$-diversity components in a volcano-dominated landscape. It allows starting to understand the causes underlying turnover and nestedness and the formation history of the Kenyan flora.

4.1 | Formation of Kenyan flora

In this study, both phylogenetic dissimilarity and taxonomic dissimilarity approaches yielded five floristic regions in Kenya. The shape and boundary of each region are broadly similar. The circumscription of the Manda Region and Turkana Region varies between the two approaches, which might be caused by how evolutionary history is accounted for in the $p_{\beta\text{sim}}$ method. The $p_{\beta\text{sim}}$ method considers the phylogenetic distance of lineages from distinct clades segregated along grid cells in the Manda and Turkana regions, which reflects the long compositional history.
within each grid cell. However, it is hard for the $\beta_{\text{sim}}$ method to detect the compositional history. This situation exacerbates when old lineage extinction and new radiation lead to consistent nestedness patterns along the homogeneity habitat. As both areas were covered by a homogeneous habitat and presented high nestedness, lineage exchange was more important than species exchange (Peixoto et al., 2017). Referring to $\beta_{\text{sim}}$, the Mandera Region may be more similar to the Horn of Africa, representing a unique composition of floristic history (Wang et al., 2019). However, results from $\beta_{\text{sim}}$ suggested that the Mandera Region broadly included the eastern part of Kenya, which was defined as the Somali Region before (Linder et al., 2012). The delineation of the Mandera Region based on $\beta_{\text{sim}}$ might be more reasonable, considering the sampling coverage and evolutionary history in shaping the floristic regions. The smallest West Highland Region is surrounded by the Volcanic Region. The largest continuous tropical forests of Kenya (Kakamega Forest and Nandi Forest) are distributed in this region and are more similar to the western extensive rainforests ranging from West Africa to Uganda, making this region unique in Kenya. The circumscription of the West Highland Region varies between the results of $\beta_{\text{sim}}$ and $\beta_{\text{sim}}$. We followed the results of $\beta_{\text{sim}}$ as in the Mandera Region. Therefore, because of the variable boundary of the small regions shown here, it is somewhat difficult to explain which is more reasonable and needs to be considered in parallel with the nature of $\beta_{\text{sim}}$ and $\beta_{\text{sim}}$. On the contrary, some species near the northern and western boundaries of Kenya may be distributed in only a few localities within Kenya (restricted to the Mandera and West Highland Regions) but have a much wider range extending beyond these regions. This may also lead to bias in the cluster analysis results. The boundary of the Volcanic and Pan Coastal regions was broadly similar to the high turnover area of Kenya. The phytegeographical regionalization results implied potential distinct evolutionary histories of these two regions, even though they are both high turnover areas.

The nestedness and turnover components of phylogenetic $\beta$-diversity are higher than the random expectation in climatically and geologically unstable regions, implying that the flora of Kenya has been impacted by climate change and orogeny (Figure 57.9a,b). Within a region like Turkana under the influence of past climatic change, phylogenetic diversity would be lower than the randomized phylogenetic diversity expectation from whole phylogenetic tree when the species are clustering in specific clades (Figure 4c). This means that the extant species from specific clades that adapted to a similar environment have suffered from strong environmental filtering and show strong phylogenetical signals. Diverse habitats are currently distributed in the Volcanic Region, in which the Miocene paleobotanical records also document a great variety of environments and habitats, ranging from evergreen to deciduous forests and from wet to dry woodlands (Bonnefille, 2010). Changeable habitats may have already existed in Kenya for a long time under several climate and inter-climate optimal phases (Westerhold et al., 2020). The interaction between climate and orogeny driving the formation of unique montane forest types, savanna or alpine vegetation could serve as a source of new species to neighbouring areas. The high turnover components observed in climatically and geologically stable regions such as the coastal area could be caused by the preservation of ancient lineages.

4.2 | Macroevolutionary mechanisms underlying phylogenetic nestedness

The $\beta$-diversity pattern in low latitude areas has been reported to be dominated by turnover (Dobrovolski et al., 2012; Soininen et al., 2018). However, we found several areas that were dominated by nestedness at low latitudes, which may have been shaped by different evolutionary histories compared with turnover-dominated areas. Under the floristic evolutionary history, long-term climatic fluctuations that potentially affect biodiversity need to be considered (Muellner-Riehl et al., 2019). Environmental filtering plays a strong role in determining species’ composition assemblage at a regional/local scale (Swenson et al., 2012); the assembly only maintains the species that can tolerate the limits of the environment (Cardillo, 2011). The phylogenetic niche conservatism hypothesis indicates that closely related species tend to adapt to similar environmental conditions; thus, an assemblage may become phylogenetically clustered.
(Cardillo, 2011; Wiens & Donoghue, 2004). Typically, the amount and seasonal distribution of rainfall in East Africa is mainly controlled by three factors, that is, the Intertropical Convergence Zone (ITCZ), the interaction of moist airflows from the Atlantic Ocean and the Indian Ocean at the continental scale, and the regional scale orogeny (Linder, 2017). These factors could be the best predictors of the distribution of modern East African patchy vegetation. Pliocene Pleistocene aridification events in East Africa deeply reshaped the vegetation from closed canopy to open savanna (Bobe, 2006; Cane & Molnar, 2001; de Menocal, 1995). Gradually, the progression towards a harsh environment took place at the eastern equator and the northern parts of Kenya. Besides, global scale aridification events and topographic change at the regional scale induced by rifting played a dominant force in the late Neogene climatic formation in East Africa rather than a background role (Pierre et al., 2006).

Today, the topography of the East African Plateau influences the distribution and amount of rainfall associated with the African-Indian monsoon and westerly moisture sources at the continental scale (Wichura et al., 2010). Associated hydrological modifications induce shifts in vegetation. There may have been a strong selection of species caused by an adaptation to dry biomes. Eastern and northern parts of Kenya have experienced greater biome instability, leading to limited speciation and increased extinction compared with western and southern Kenya. In turn, environmental filtering has been primarily responsible for the nestedness and phylogenetic clustering pattern in the Turkana and Mandera regions, where landscapes are homogeneous and the climate is unstable. Similar to the studies that focused on the extremely low temperature of the Last Glacial Maximum phase, which shaped nestedness patterns (Dobrovolski et al., 2012; Qian et al., 2020), our results also support that high extinction risks in a harsh environment might have been critical to shaping the nestedness dominant phylogenetic β-diversity pattern of Kenya.

4.3 | Macroevolutionary mechanisms underlying phylogenetic turnover

The highest turnover was found in volcanic and coastal areas (Figure 5S.10), which represent the Eastern Afromontane hotspot and Coastal Forests of Eastern Africa hotspot, respectively. The turnover shows a non-monotonic relationship between high and low latitudes, a pattern found in North America as well (Xing & He, 2019). In the vicinity of the Rift, escarpments, mountains and volcanoes present a succession of different vegetation belts, that is, savanna and woodland at low elevation, mountain forest at mid-elevation, and afro-alpine heathland at high elevations (above 3000m), following increasing rainfall and decreasing temperature altitudinal gradients (Bonnefille, 2010). Thus, a higher phylogenetic turnover was well predicted by the greater ecological opportunity in the heterogeneous landscape of the Volcanic Region. Because of the highly heterogeneous topography, the volcanic area is known as a “sky island” and is considered a natural laboratory for testing the progress of species diversification and colonization (Cox et al., 2014). Significant phylogenetic Neo- and paleo-endemism was detected in the volcanic area (Dagallier et al., 2020). The NRI in the volcanic area showed phylogenetic clustering at a deep level, meanwhile, the NTI showed phylogenetic overdispersion at a shallow level (Figures 4 and S8.12). The higher proportion of clustered assemblages for NRI compared with NTI parallels the high species richness and high PD in the volcanic area. This pattern suggests that if there were a niche filled by ancient species, the interspecific competition would increase at a local scale as the niche becomes saturated. The higher proportion of overdispersion assemblages for NTI compared with NRI indicates the possibility that species are rapidly radiating in a suitable habitat of the Volcanic Region. This pattern suggests that both opportunities and challenges coexist in the volcanic regions. In terms of the opportunities, geological events contributed greatly to the ecological opportunities for speciation and colonization; in terms of the challenges, frequent geological events were a disadvantage to old lineages accumulation. The accumulation of young lineages caused by in situ differentiation and colonization promoted the floristic diversification in the Southeast Asia and North America volcanic areas (Mercckx et al., 2015; Rodriguez et al., 2018; Sosa et al., 2018), and the African volcanic areas also have the similar routes (Blackburn & Measey, 2009; Chen, Wang, & Renner, 2016; Gizaw et al., 2016; Kandziora et al., 2016).

A larger proportion of overdispersion for NRI was detected in the coastal area compared with NTI in the volcanic area. The coastal area environment is stable under the influence of aridification, including part of the East Arc Mountain (Lovett et al., 2005), where the flora has strong affinities to that of the western Guineo-Congolian Region (Lovett, 1993). This is a relic of tropical forests continuously distributed since the early Tertiary (Couvreur et al., 2008). Ancient lineages such as Platypterocarpus, Neohemsleya and Brochoneura can be observed in this area. The coastal area was once a refuge for a high percentage of tropical elements that were widely distributed in Tropical East Africa. Thanks to environmental stability, the ancient lineages persisted and generated habitat specialists in coastal areas. These relic elements, in turn, mainly contributed to the high turnover component pattern in this area.

4.4 | Biological conservation within hotspots

The distribution of protected areas in Kenya roughly coincides with hotspots of phylogenetic turnover (p$_{\text{fratio}}$ > 1) that are evenly distributed along coastal-volcanic regions. Protected area networks like those in Kenya could capture species turnover along spatial or environmental gradients and the inestimable potential risk of species loss (Tuomisto et al., 2003). Small protected areas are scattered between larger ones and serve as corridors (especially in the Eastern Afromontane Region) and could facilitate species range-shifts in response to climate change to decrease the likelihood of local and regional extinction (Fagan, 2002; Pearson, 2006). Given that conservation is underfunded, protected area selection should be optimized
The creation of few large and several small reserves could reduce conflicts between agricultural development and biodiversity conservation.

4.5 | Limitations

According to the species–area relationship (SAR) model constructed by Qian et al. (2021), we found that the average completeness of species in the grid cell for the data from GBIF and RAINBIO is only 11% in Kenya. The current GBIF and RAINBIO data along are not sufficient in addressing biogeographic and macroecological questions for the flora of Africa (Qian et al., 2021). While, after combining with data from FTEA and other consulted sources, the average completeness of species in the grid cell reached 70.3% in Kenya (Appendix S9, Table S9.5). Data from FTEA and other consulted sources not only improve the completeness of Kenya angiosperms list (Tables S4.1 and S4.2), but also are more evenly distributed and can reduce the geographic bias of GBIF and RAINBIO data (Figure S9.13b). In addition, species list completeness estimation by SAR model may be higher than the actual richness in eastern and northern Kenya or lower than actual species richness in volcanic area of Kenya. Therefore, the grid cells distributed in the volcanic area of Kenya with species completeness between 51% and 100% should be the focal area of plant investigation in the future (Figure S9.13c). The species richness of these grid cells may be much higher than estimated value. Based on this fact, our current analysis may underestimate the level of species/phyllogenetic diversity in this area. However, a more complete set of data will supply to identify local diversity characteristics underneath regional patterns. With the further improvement of distribution data, the understanding of the formation mechanism of plant diversity pattern in Kenya will become more specific and refined.

Assigning each county-level distribution record to 50-km×50-km grid cells can lead to some false occurrences. The greater the difference between county and grid area, the more likely there will be such erroneous distribution data and bias may be introduced in the biogeographic analysis. Comparing the results from 50-km×50-km grid cells (small scale) to those from a 100-km×100-km grid cell (large scale), we found that PD and \( p_{\text{beta}} \) were not sensitive to the grid scale and the patterns were broadly similar (Figures 3 and 4b; Appendix S9, Figure S9.14a,b), while the boundaries of the phytogeographical regions varied (Figures 1b and S9.14c). The West highland and Volcanic regions produced by small-scale grid cells matched the vegetation map (Figure S3), while the larger-scale grid cells could not distinguish the boundaries between rainforest and montane forest. This may be due to the small county size and the highly heterogeneous habitats. While, the Turkana and Mandera regions, both of which have homogenous habitats (mainly the Acacia–Commiphora bushland or the semi-desert scrub), were less sensitive to the way we processed distributed data. For the Pan Coastal Region, neither the small nor the larger scale could identify reasonable boundaries that could separate the inland bushland and thicket from the coastal vegetation. This is because there are fewer, larger and highly heterogeneous counties distributed here. In the future, more coordinated data could help delineate the boundaries more accurately. Despite limitations in the way we approached the distribution records, we still believe that the patterns on 50-km×50-km grid cells reasonably explained the biogeographic formation of the flora.

5 | CONCLUSIONS

Our results showed that climatic fluctuations and orogeny influenced the phylogenetic \( p \)-diversity pattern of the Kenyan flora. The turnover and nestedness components of \( p \)-diversity display distinct geographic patterns, reflecting contrasting legacies of past climate change and orogeny. We suggest that the complex topography of the Volcanic Region generated high climatic complexity with environmental gradients along elevation and latitude, promoting a high turnover component that may explain why the Volcanic Region is the most diverse floristic region of Kenya. Areas with long-term climate stability are prone to keeping ancient lineages, which are the main elements that contribute to the turnover dominant region. This study contributes to elucidating the evolutionary history and formation mechanism of the Kenyan flora and improves the understanding of the key factors responsible for the biodiversity patterns of the hotspots of East Africa.

ACKNOWLEDGEMENTS

We would like to thank Yang-Jun Lai, Qian Zhang, Yu-Chang Yang and Yi-Shuo Liang for providing helpful suggestions. We thank Hai-Hua Hu, Yun Liu and Yan-Ting Niu for assistance with data analysis. We thank two anonymous reviewers and the editor for constructive feedback on earlier versions of this manuscript. This study was supported by the International Partnership Program of Chinese Academy of Sciences (grant number: 151853KYSB20190027), Sino-Africa Joint Research Center, Chinese Academy of Sciences, CAS International Research and Education Development Program (SAJC202101), National Natural Science Foundation of China (NNSF 31970212 and 31800178), the Youth Innovation Promotion Association of the Chinese Academy of Sciences (2021077) and Vietnam National Foundation for Science and Technology Development (NAFOSTED, grant number: 106.03-2019.12).

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The datasets that support the findings of this study are available on Dryad, and DNA sequences were deposited in the GenBank with accession numbers offered in Dryad at https://doi.org/10.5061/dryad.18931zd0p.
REFERENCES

Agnew, A. D., & Agnew, S. (1994). Upland Kenya wild flowers: A flora of the ferns and herbaceous flowering plants of upland Kenya (2nd ed.). East Africa Natural Society.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403–410. https://doi.org/10.1016/0022-2836(90)90320-2

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., Sanders, N. J., Cornell, H. V., Comita, L. S., Davies, K. F., Harrison, S. P., Kraft, N. J. B., Stegen, J. C., & Swenson, N. G. (2011). Navigating the multiple meanings of beta diversity: A roadmap for the practicing ecologist. Ecology Letters, 14(1), 19–28. https://doi.org/10.1111/j.1466-0265.2010.01552.x

Antonelli, A., Kissling, W. D., Flantua, S. G. A., Bermúdez, M. A., Mulch, A., Muller-R_intentz, A. N., Kreft, H., Linder, H. P., Badgley, C., Fjeldså, J., Fritz, S. A., Rahbek, C., Herman, F., Hooghiemstra, H., & Hoorn, C. (2018). Geological and climatic influences on mountain biodiversity. Nature Geoscience, 11(10), 718–725. https://doi.org/10.1038/s41567-018-0236-z

Antonelli, A., Nylander, J. A. A., Persson, C., & Sanmartín, I. (2009). Tracing the impact of the Andean uplift on Neotropical plant evolution. Proceedings of the National Academy of Sciences of the United States of America, 106(24), 9749–9754. https://doi.org/10.1073/pnas.0811411106

Baker, B. H., Williams, L. A. J., Miller, J. A., & Fitch, F. J. (1971). Sequence and geochronology of the Kenya rift volcanics. Tectonophysics, 11(3), 191–215. https://doi.org/10.1016/0040-1951(71)90030-8

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography, 19(1), 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. Global Ecology and Biogeography, 21(12), 1223–1232. https://doi.org/10.1111/j.1466-8238.2011.00756.x

Baselga, A., & Orme, C. D. L. (2012). betapart: An R package for the study of beta diversity. Methods in Ecology and Evolution, 3(5), 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x

Blackburn, D. C., & Measey, G. J. (2009). Dispersal to or from an African biodiversity hotspot? Molecular Ecology, 18(9), 1904–1915. https://doi.org/10.1111/j.1365-294X.2009.04156.x

Bobe, R. (2006). The evolution of arid ecosystems in eastern Africa. Journal of Arid Environments, 66(3), 564–584. https://doi.org/10.1016/j.jaridenv.2006.01.010

Bonnefille, R. (2010). Cenozoic vegetation, climate changes and hominid evolution in tropical Africa. Global and Planetary Change, 72(4), 390–411. https://doi.org/10.1016/j.gloplacha.2010.01.015

Cane, M. A., & Molnar, P. (2001). Closing of the Indonesian seaway as a precursor to east African afidization around 3–4 million years ago. Nature, 411(6834), 157–162. https://doi.org/10.1038/35075500

Cardillo, M. (2011). Phylogenetic structure of mammal assemblages at large geographical scales: Linking phylogenetic community ecology with macroecology. Philosophical Transactions of the Royal Society B: Biological Sciences, 366(1577), 2545–2553. https://doi.org/10.1111/j.2041-210X.2012.00224.x

Cayuela, L., Granzow-de la Cerda, I., Albuquerque, F. S., & Golicher, D. J. (2012). taxonstack: An R package for species names standardisation in vegetation databases. Methods in Ecology and Evolution, 3(6), 1078–1083. https://doi.org/10.1111/j.2041-210X.2012.00232.x

Chang, J., Rabosky, D. L., & Alfaro, M. E. (2020). Estimating diversification rates on incompletely sampled phylogenies: Theoretical concerns and practical solutions. Systematic Biology, 69(3), 602–611. https://doi.org/10.1093/sysbio/syz081

Chen, L. Y., Wang, Q. F., & Renner, S. S. (2016). East Asian Lobelioideae and ancient divergence of a giant rosette Lobelia in Himalayan Bhutan. Taxon, 65(2), 293–304. https://doi.org/10.12705/652.6

Chen, Z. D., Yang, T., Lin, L., Lu, L. M., Li, H. L., Sun, M., Liu, B., Chen, M., Niu, Y. T., Ye, J. F., Cao, Z. Y., Liu, H. M., Wang, X. M., Wang, W., Zhang, J. B., Meng, Z., Cao, W., Li, J. H., Wu, S. D., ... Lu, A. M. (2016). Tree of life for the genera of Chinese vascular plants. Journal of Systematics and Evolution, 54(4), 277–306. https://doi.org/10.1111/jse.12219

Couvreur, T. L., Chatrou, L. W., Sosef, M. S., & Richardson, J. E. (2008). Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. BMC Biology, 6(1), 54. https://doi.org/10.1186/1741-7007-6-54

Couvreur, T. L., Dauby, G., Blach-Overgaard, A., Deblauwe, V., Dessein, S., Droissart, V., Hardy, O. J., Harris, D. J., Janssens, S. B., Ley, A. C., Mackinder, B. A., Sonké, B., Sosef, M. S., Stévart, T., Svenning, J.-C., Wieringa, J. J., Faye, A., Missoup, A. D., Tolley, K. A., ... Sepulchre, P. (2021). Tectonics, climate and the diversification of the tropical African terrestrial flora and fauna. Biological Reviews, 96(1), 16–51. https://doi.org/10.1111/brv.12644

Cox, S. C., Prys-Jones, R. P., Habel, J. C., Amakobe, B. A., & Day, J. J. (2014). Niche divergence promotes rapid diversification of East African sky Island white-eyes (Aves: Zosteropidae). Molecular Biology, 23(6), 4103–4118. https://doi.org/10.1186/s12980-014-0180-3

Dagallier, L.-P., Janssens, S. B., Haile, O. J., Harris, D. J., Janssens, S. B., Ley, A. C., Mackinder, B. A., Sonké, B., Sosef, M. S., Stévart, T., Svenning, J.-C., Wieringa, J. J., Faye, A., Missoup, A. D., Tolley, K. A., ... Sepulchre, P. (2021). Tectonics, climate and the diversification of the tropical African terrestrial flora and fauna. Biological Reviews, 96(1), 16–51. https://doi.org/10.1111/brv.12644

Dauby, G., Zaisz, R., Blach-Overgaard, A., Catarino, L., Damen, T., Deblauwe, V., Duarte, M. C., Engledow, H., Fadeur, G., Figueira, R., Gereau, R. E., Hardy, O. J., Harris, D. J., de Heij, J., Janssens, S., Klomberg, Y., Ley, A. C., Mackinder, B. A., Meert, P., ... Covrre, T. L. P. (2016). RAINBIO: A mega-database of tropical African vascular plants distributions. PhytoKeys, 74, 1–18. https://doi.org/10.3897/phytokeys.74.9723

de Menocal, P. B. (1995). Plio-Pleistocene African Climate. Science, 270(5233), 53–59. https://doi.org/10.1126/science.aao2200

Ding, W. N., Nee, R. H., Spicer, R. A., & Xing, Y. W. (2020). Ancient orogenic and monsoon-driven assembly of the world's richest temperate alpine flora. Science, 369(6503), 578–581. https://doi.org/10.1126/science.abb4484

Diserud, O. H., & Ødegaard, F. (2007). A multiple-site similarity measure. Biology Letters, 3(1), 20–22. https://doi.org/10.1098/rsbl.2006.0553
Tuomisto, H., Ruokolainen, K., & Yli-Halla, M. (2003). Dispersal, environment, and floristic variation of western Amazonian forests. Science, 299(5604), 241–244. https://doi.org/10.1126/science.1078037

UNEP-WCMC and IUCN. (2022). Protected planet: The World Database on Protected Areas (WDPA) and World Database on Other Effective Area-based Conservation Measures (WD-OCM). UNEP-WCMC and IUCN. www.protectedplanet.net

Venter, O., Fuller, R. A., Segan, D. B., Carwardine, J., Brooks, T., Butchart, S. H. M., MacDonald, R., & Chapman, G. R. (1984). Late Quaternary caldera volcanoes of the Kenya Rift Valley. Journal of Geophysical Research, 89(B10), 8553–8570. https://doi.org/10.1029/JB089iB10p08553

Xing, D., & He, F. (2019). Environmental filtering explains a U-shape latitudinal pattern in regional beta-diversity for eastern North American trees. Ecology Letters, 22(2), 284–291. https://doi.org/10.1111/ele.13188

Ye, J. F., Liu, Y., & Chen, Z. D. (2020). Dramatic impact of metric choice on biogeographical regionalization. Plant Diversity, 42(2), 67–73. https://doi.org/10.1016/j.pld.2019.12.003

Zhao, X., & Sandelin, A. (2014). GMD: Generalized Minimum Distance of distributions. R package version 0.3.3. https://doi.org/10.1371/journal.pone.0023409

Zhou, Y. D., Liu, B., Mbuni, Y., Yan, X., Mwachala, G., Hu, G. W., & Wang, Q. F. (2017). Vascular flora of Kenya, based on the Flora of Tropical East Africa. PhytoKeys, 90, 113–126. https://doi.org/10.3897/phytokeys.90.20531

**BIOSKETCHES**

Qiang Zhang is pursuing PhD in the Institute of Botany, Chinese Academy of Sciences. His research interests include biodiversity, biogeography and conservation biology.

Jian-Fei Ye is senior engineer at Beijing Botanical Garden, Institute of Botany, Chinese Academy of Sciences. His main research interests are phylofloristics, biogeography and plant systematics.

Bing Liu is associate professor at State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences. He is specialized in taxonomy, biodiversity and biogeography. All other authors have interests in evolution, biogeography, phylogeny and biodiversity of regional and global plants.

Authors contributions: B.L., Z.D.C. and J.F.Y. conceived the study; Q.Z., C.T.L., D.M.N. and N.R.R. undertook the analyses; Q.Z., B.L., J.F.Y. and Z.D.C. wrote the manuscript; B.L., Z.D.C., J.F.Y., Q.Z., C.T.L., D.M.N., W.O.O. and L.M.L. contributed data; all authors contributed to and approved the final manuscript.

**SUPPORTING INFORMATION**

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

**How to cite this article:** Zhang, Q., Ye, J-F, Le, C-T, Njenga, D. M., Rabarijaona, N. R., Omollo, W. O., Lu, L-M, Liu, B., & Chen, Z-D (2022). New insights into the formation of biodiversity hotspots of the Kenyan flora. Diversity and Distributions, 28, 2696–2711. https://doi.org/10.1111/ddi.13624