

# An updated floristic map of the world

Received: 7 July 2022

Accepted: 28 April 2023

Published online: 30 May 2023

 Check for updates

Yunpeng Liu<sup>1,2,11</sup>, Xiaoting Xu<sup>1,3,11</sup>, Dimitar Dimitrov<sup>4,11</sup>, Loic Pellissier<sup>5,6</sup>, Michael K. Borregaard<sup>2</sup>, Nawal Shrestha<sup>1,7</sup>, Xiangyan Su<sup>1,8</sup>, Ao Luo<sup>1</sup>, Niklaus E. Zimmermann<sup>6</sup>, Carsten Rahbek<sup>1,2,9,10</sup> ✉ & Zhiheng Wang<sup>1</sup> ✉

Floristic regions reflect the geographic organization of floras and provide essential tools for biological studies. Previous global floristic regions are generally based on floristic endemism, lacking a phylogenetic consideration that captures floristic evolution. Moreover, the contribution of tectonic dynamics and historical and current climate to the division of floristic regions remains unknown. Here, by integrating global distributions and a phylogeny of 12,664 angiosperm genera, we update global floristic regions and explore their temporal changes. Eight floristic realms and 16 nested sub-realms are identified. The previously-defined Holarctic, Neotropical and Australian realms are recognized, but Paleotropical, Antarctic and Cape realms are not. Most realms have formed since Paleogene. Geographic isolation induced by plate tectonics dominates the formation of floristic realms, while current/historical climate has little contribution. Our study demonstrates the necessity of integrating distributions and phylogenies in regionalizing floristic realms and the interplay of macroevolutionary and paleogeographic processes in shaping regional floras.

Biogeographic regionalizations divide the world into regions based on the similarity between faunas or floras and have been one of the central topics in biogeography since the time of Charles Darwin and Alfred Russel Wallace<sup>1–4</sup>. Darwin's biogeographic observations on the similarities in faunas or floras across regions during the voyage of Beagle (1831–1836) led him to his theory on natural selection<sup>3,5</sup>. In 1876, Wallace published his global map of zoological realms based on compositional similarity and taxonomic relationships of animal families across regions. The early maps of biogeographic regions, including the Wallace zoological realms, have significantly improved our understanding of global biodiversity<sup>6</sup>, and provided a spatially explicit tool for conservation planning<sup>7,8</sup>.

To understand the evolution of plant diversity, several global floristic regionalization schemes were generated by early authors, including de Candolle (1820)<sup>9,10</sup>, Schouw (1823)<sup>4</sup>, Engler (1892) and others<sup>1,4,10–12</sup>. Later, Takhtajan (1969, 1970, 1974, 1978, and 1986) summarized the basic understanding on the distribution and origin of floras and developed the most widely used map of floristic biogeographic regionalization scheme up to now, which divided the world landmasses into six “kingdoms” and 35 “regions”<sup>4</sup>. These early floristic schemes identified biogeographic regions and their hierarchical relationships mainly based on endemism of floras at different taxonomic levels, combined with a generally descriptive understanding of paleoclimate and geological history<sup>1,4,10–12</sup>. The rapid accumulation of

<sup>1</sup>Institute of Ecology, College of Urban and Environmental Sciences, and Key Laboratory of Earth Surface Processes of Ministry of Education, Peking University, 100871 Beijing, China. <sup>2</sup>Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen Ø, Denmark. <sup>3</sup>Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, 610065 Chengdu, Sichuan, China. <sup>4</sup>Department of Natural History, University Museum of Bergen, University of Bergen, Postbox 7800, 5020 Bergen, Norway. <sup>5</sup>Landscape Ecology, Institute of Terrestrial Ecosystems, ETH Zurich, 8092 Zurich, Switzerland. <sup>6</sup>Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland. <sup>7</sup>State Key Laboratory of Grassland Agro-ecosystems, Institute of Innovation Ecology, Lanzhou University, 730000 Lanzhou, China. <sup>8</sup>Land Consolidations and Rehabilitation Center, Ministry of Natural Resources, 100035 Beijing, China. <sup>9</sup>Center for Global Mountain Biodiversity, GLOBE Institute, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen, Denmark. <sup>10</sup>Danish Institute for Advanced Study, University of Southern Denmark, 5230 Odense M, Denmark. <sup>11</sup>These authors contributed equally: Yunpeng Liu, Xiaoting Xu, Dimitar Dimitrov. ✉ e-mail: [crahbek@snm.ku.dk](mailto:crahbek@snm.ku.dk); [zhiheng.wang@pku.edu.cn](mailto:zhiheng.wang@pku.edu.cn)

phylogenetic and species distribution data has significantly improved the development of quantitative and repeatable biogeographic regionalizations, providing valuable insight on the historical relationships among floras or faunas<sup>2,13–19</sup>. For example, an updated global zoogeographic regionalization<sup>2</sup> was generated recently using quantitative phylogenetic relatedness of birds, mammals and amphibians showing the geographic variations in tetrapod evolution at the global scale<sup>20,21</sup>. A few regional floristic regionalizations using similar methods have been conducted in China<sup>17,22</sup>, Japan<sup>23</sup>, South Africa<sup>15</sup>, and the tropics<sup>14</sup>. However, the progress in building a global floristic regionalization using quantitative approaches has lagged behind state-of-the-art zoogeographic regionalizations due to the lack of phylogenetic and distribution data of plants at a global scale. Recently, Carta et al. proposed a global floristic regionalization with three floristic kingdoms based on phylogenetic beta diversity of a fraction (ca. 20%) of vascular plant species<sup>24</sup>. The impacts of incomplete sampling in both distribution and phylogeny on the regionalization of Carta et al.<sup>24</sup> remain unknown. Meanwhile, this regionalization was based on species rather than genera, limiting its comparison with previous floristic regionalizations based on genus and family endemism<sup>4,11,12</sup>.

The drivers shaping boundaries between different biogeographic regions are critical for understanding the macroevolution of floras and faunas in different regions<sup>20</sup>, yet they remain rarely explored<sup>20,25</sup>. The dynamics in geology and macroclimate over time have left profound effects on the speciation, extinction, and dispersal of plants and thus are considered as drivers shaping floristic boundaries<sup>20,21,26,27</sup>. The dynamics in plate tectonics throughout the earth's history, including plate collision, orogeny and the emergence and breakup of land bridges, have led to significant changes in geographic isolation and floristic exchanges among landmasses over time, subsequently influencing evolutionary processes such as speciation and extinction of floras in different regions<sup>20,28,29</sup>. Macroclimate, especially temperature and precipitation, represents a major dimension of the ecological niches of plants<sup>30,31</sup>. Geographic differences in paleoclimate during geological times led to significant changes in species compositions

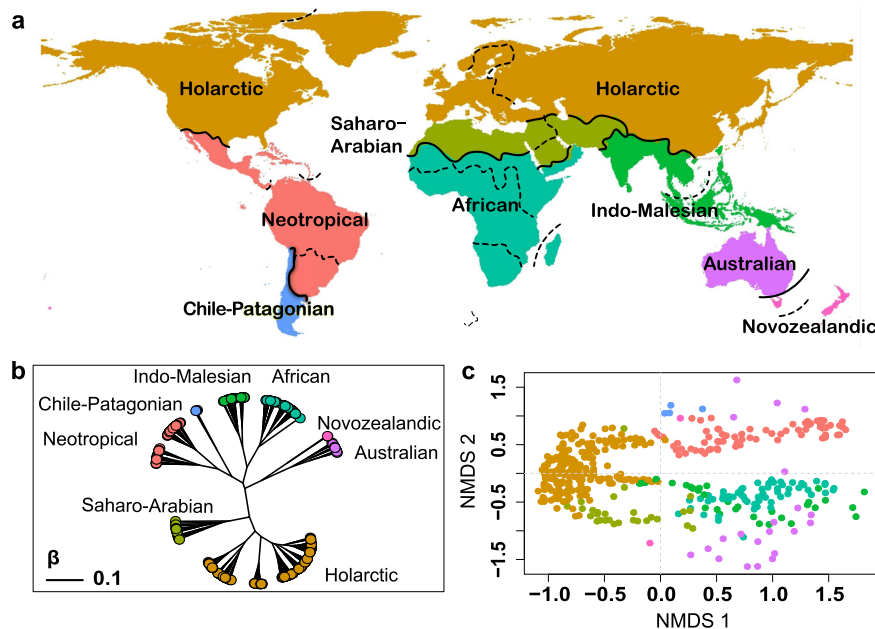
across space because of the effect of climatic filtering on species distributions<sup>30,32</sup>. A recent study indicated that current climate and tectonic movements contributed to the boundaries between zoological realms<sup>20</sup>. However, the spatiotemporal variation in the relative roles of these two drivers on the boundaries between global floristic realms remain unknown. Moreover, isolation-induced clade splitting, and the independent radiation of descendant lineages have led to abrupt floristic transitions across biogeographic boundaries<sup>33</sup>. In addition, different clades may contribute unequally to the division of different realms due to differences in clade evolutionary histories and geographic distributions<sup>34</sup>, which remains to be evaluated.

Here, we present a global map of floristic realms by integrating distribution data and a phylogeny of 12,664 angiosperm genera (ca. 85% of all known angiosperm genera). Floristic realms are identified using hierarchical clustering methods based on phylogenetic beta diversity between regions at genus level. We then demonstrate the temporal dynamics of the identified floristic realms during the Cretaceous and the Cenozoic, and further compare our work with Takhtajan's floristic map<sup>4</sup> and Holt et al.'s zoological map<sup>2</sup>. To explore the mechanisms underlying the formation of the identified floristic realms, we evaluate the effects of contemporary climate, the dynamics of geographic isolation induced by long-term plate tectonics, and historical climate. We also evaluate the relative contributions of clade splitting events at different geological times to realm divisions.

## Results and discussion

### World's floristic regionalization

Hierarchical clustering analysis based on phylogenetic beta diversity (see Methods) indicates that the terrestrial world is divided into eight floristic realms, namely African, Australian, Novozealandic, Indo-Malesian, Neotropical, Chile-Patagonian, Holarctic and Saharo-Arabian realms (Fig. 1a). The African realm is closely related to the Indo-Malesian realm, the Australian realm is closely related to the Novozealandic realm, and the Neotropical realm is closely related to the Chile-Patagonian realm (Fig. 1b). The above realms are grouped into



**Fig. 1 | Floristic realms and sub-realms of the world. a** Boundaries of the eight floristic realms and 16 sub-realms are shown in solid and dashed lines, respectively. **b** The unrooted dendrogram depicts the relationships among floristic realms evaluated using UPGMA clustering method based on phylogenetic beta diversity between realms. The scale bar in the dendrogram shows the dissimilarity between realms. **c** The scatter plot shows the dissimilarities in the phylogenetic

compositions between different geographic standard units (GSU) generated using non-metric multidimensional scaling (NIMDS) ordination. Each tip in the dendrogram and each point in the scatter plot represents a geographic standard unit and the colors indicate the floristic realms that they belong to. For comparison, the floristic realms and sub-realms based on trees with alternative dating constraints are shown in Supplementary Fig. 2. Source data are provided as a Source Data file.

the Gondwanan super-realm. The Holarctic and Saharo-Arabian realms are grouped into the Laurasian super-realm. Within these eight realms, we further identified 16 floristic sub-realms (Fig. 1; also see Supplementary Table 1 for their names and Supplementary Fig. 1 for their relationships).

The boundaries between different realms generally have high confidence (i.e., “hardness”) in most cases as evaluated by silhouette analysis<sup>35</sup> (Supplementary Fig. 2, see Methods) and are robust to different assumptions on the crown age of angiosperms (Supplementary Fig. 3) and to variations in phylogenetic topology (Supplementary Figs. 4–8). We also evaluated the effect of incomplete sampling on realm boundaries by repeating our regionalization process using only the genera included in ref. 24 (9905 angiosperm genera), and we found that the boundaries were also robust (Supplementary Fig. 9, also see Supplementary Discussions for details). Interestingly, distribution ranges for ca. 53.6% genera do not extend beyond the identified boundaries (Supplementary Data 5), suggesting that the identified realm boundaries may reflect ecological or evolutionary barriers of plant distributions<sup>31</sup>. Although the boundaries identified by the UPGMA clustering are mostly consistent with those based on the fuzzy clustering method (i.e., fuzzy c-means), inconsistency exists in the identification of the Chile-Patagonian realm and the North American sub-realm, suggesting lower confidence in the identification of these realms/sub-realms than others (see Supplementary Fig. 17, Methods, and Supplementary Discussions for details).

Floristic realms and their relatedness based on taxonomic beta diversity are relatively consistent with phylogenetic-based floristic realms with three exceptions (Supplementary Fig. 10, also see Supplementary results and discussion for a detailed comparison). First, subtropical East Asia was grouped into the Indo-Malesian realm when taxonomic beta diversity was used but was grouped into the Holarctic realm when phylogenetic beta diversity was used. Second, Mexico was grouped into the Chile-Patagonian realm when taxonomic beta diversity was used but was grouped into the Neotropical realm when phylogenetic beta diversity was used. Third, (Australian, Chile-Patagonian) realms are grouped with (African, Indo-Malesian) realms, but was grouped with ((African, Indo-Malesian), (Neotropical, Chile-Patagonian)) realms when phylogenetic beta diversity was used. These differences between the taxonomic-based and phylogenetic-based results may suggest recent exchange through dispersal of lineages among these regions (see Supplementary Discussion for details).

### The divergent times between the identified realms

By cutting the dated phylogeny at different depths (i.e., geological times), we found that the identified floristic realms have not become distinct before the Cretaceous (160 Ma; Fig. 2 & Fig. 3). During the Early Cretaceous, the divergences in present-day floras in the Gondwanan and Laurasian super-realms had formed (Fig. 2c). The divergences between the floras of Neotropical and African realms were not clear until about 80 Ma ago. During the Cenozoic, the dissimilarity between the present-day floras of different realms significantly increased (Fig. 3).

The divergences among the present-day floras of most realms have formed since the early Cenozoic, and the boundaries between them remain largely unchanged towards the present with two exceptions (Fig. 2). First, the boundaries between the African and the Indo-Malesian realms disappeared when evaluated at the phylogenetic depth from the Eocene to the Pliocene, although their divergence was clear during the Paleogene (60 Ma; Fig. 2). The dissimilarity between the present-day floras of these two realms was much lower at the phylogenetic depth from the Eocene to the Pliocene than at the present (0 Ma), resulting in the disappearance of boundaries of these two realms at these times (Fig. 3). Fossil evidence on historical changes in woody assemblages during the Cenozoic<sup>36</sup> further supports this finding, which is possibly because the northward drifted of the Indian

subcontinent during the Cenozoic accumulated floristic exchanges between Eurasia and Africa<sup>37,38</sup>.

Second, the northern boundary of the Neotropics realm ends at the Greater Antilles and the Yucatan Peninsula in Mesoamerica during 60–40 Ma and further extends to Mexico afterwards (Fig. 2). These results are consistent with recent findings about the history of biotic interchange between South America and Mesoamerica<sup>39–42</sup>. Recent studies<sup>39,40</sup> found that dispersal of plant lineages from Amazonia to Mesoamerica and the Caribbean islands occurred continuously since the early Cenozoic and was much more frequent than dispersal from Amazonia to any adjacent regions in the south. Fossil and plate tectonic evidence suggest that the expansion of megathermal vegetation<sup>43</sup> and an emergent Aves Ridge during the Paleocene<sup>42,44</sup> may have facilitated the biotic interchange from South America to Mesoamerica and the Caribbean islands, which may have led to high phylogenetic similarities of flora in these regions.

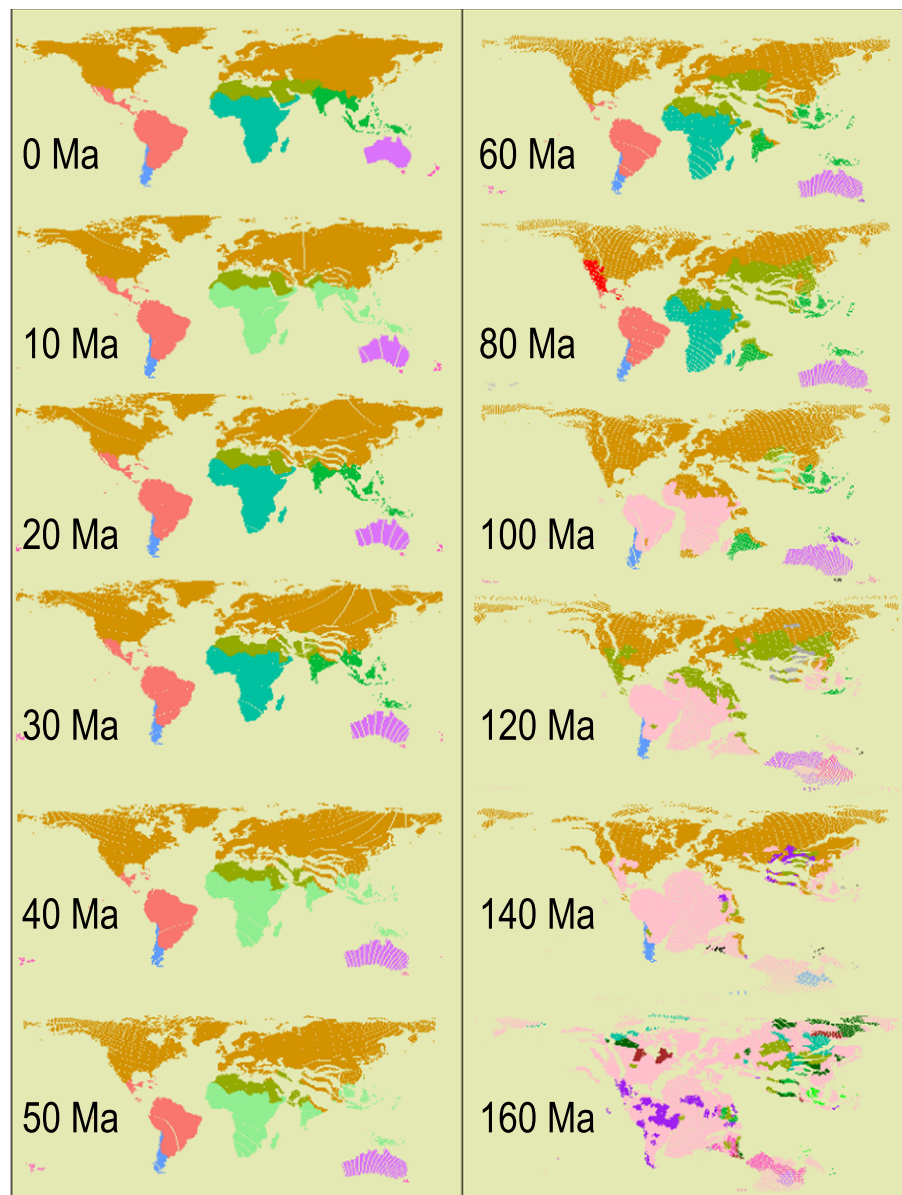
It is noteworthy that the inference of the divergence times between the identified realms was based on the phylogeny of present-day taxa. Although these results are relatively consistent with fossil evidence, how to reconstruct historical divergences between floristic realms by integrating distributions of current clades and fossils remains to be explored in future studies. Such studies may need a comprehensive framework that integrates analytical tools in paleoecology, systematics, paleoclimatology, and macroecology, in order to better explore the historical changes of floristic realms and the underlying drivers.

### Comparison with Takhtajan’s floristic regionalization and the updated Wallace realms

Several notable differences are recognized between the Takhtajan (1986) and our floristic maps (Supplementary Fig. 11). First, the “Paleotropical kingdom” in Takhtajan’s map is divided into the Indo-Malesian realm and the African realm. These two realms have been separated by the Indian Ocean since the late Jurassic<sup>31,37</sup>, which may have led to the division between them. The temporal changes in the floristic similarity between these two realms are also supported by woody angiosperm fossils in these regions<sup>36</sup>.

Second, the “Antarctic kingdom” in Takhtajan’s map is divided into the Chile-Patagonian and Novozealandic realms here, which is consistent with the view of Cox<sup>1</sup>. Our results indicate that the floras in these two regions are phylogenetically more similar to their adjacent realms than to each other throughout the geological history (Fig. 1b), which is also supported by the similarity in woody fossils during the Cenozoic<sup>36</sup>.

Third, we define the new Saharo-Arabian realm, which was treated as a subset of the “Holarctic kingdom” in Takhtajan’s map<sup>4</sup>. Our updated Holarctic realm mainly covers the ancient Laurasia landmasses<sup>19,31</sup>. The newly defined Saharo-Arabian realm covers northern Africa and the Arabian Peninsula, and has been connected with Africa and separated from the ancient Laurasia landmasses by the Tethys Sea since the Cretaceous<sup>37,45</sup>. Our analysis on realm dynamics shows that the flora in the Saharo-Arabian realm already differed from that in the Holarctic realm during the late Cretaceous (Fig. 2). Since the Early Miocene (23–16 Ma), the Saharo-Arabian realm experienced several waves of aridification, which may have further led to the evolutionary divergence of its flora from that of the Holarctic realm<sup>46,47</sup>. Notably, the separation of the Saharo-Arabian realm from the Holarctic realm is robust to different assumptions on the crown age of angiosperms (Supplementary Fig. 3), variations in phylogenetic topology (Supplementary Figs. 4–8), sampling biases (Supplementary Fig. 9), taxonomic beta diversity (Supplementary Fig. 10) and the chosen of different clustering methods (Supplementary Fig. 17). Even though, the boundary between the Saharo-Arabian and the Holarctic realms remains uncertain, as the fuzzy c-means clustering suggests that the boundary might encroach into Europe (Supplementary Fig. 17). The



**Fig. 2 | Chronology of the present-day floristic realms.** The phylogenetic tree is cut at successive phylogenetic depths and all descendent leaves are collapsed into the branches encountered at that depth. Then the realms at each phylogenetic depth are identified using the same clustering method as in Fig. 1. It is noteworthy that we do not intend to estimate the ancestral geographic ranges of phylogenetic branches. This chronological sequence of maps represents the divergence times of flora that survived to the present day, but it provides only limited information on the ancestral floristic relatedness, which should be evaluated by fossils. The floristic realms which can be matched to the present-day realms are shown in the same

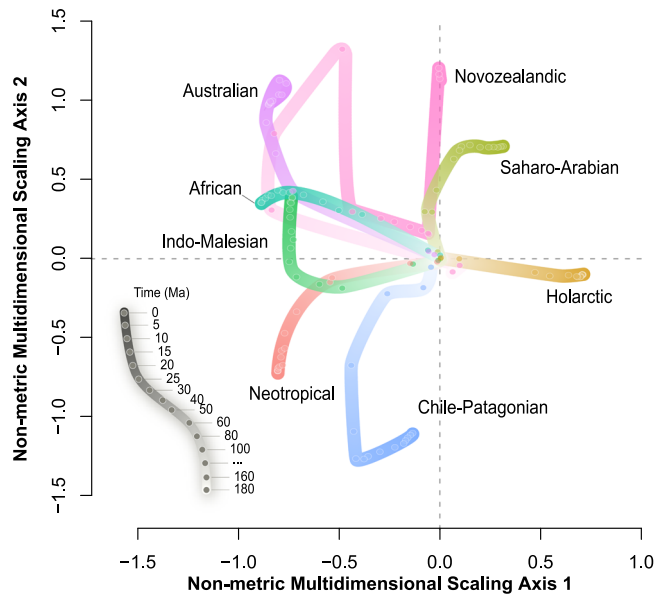
colors as shown in Fig. 1a. As the present-day floristic realms are not distinguishable in some historical periods, we used other colors to represent these ancestral floristic realms. Specifically, light green in maps of 10, 40, and 50 Ma represents the ancestral realm covering the geographic ranges of the present-day African and Indo-Malesian realms; pink appearing from 100 Ma to 140 Ma represents the ancestral realms covering the present-day Neotropical+African realms, the present-day Neotropical+African+Indo-Malesian realms, and the present-day Gondwanan super-realm, respectively. Notably, most present-day floristic realms are undistinguishable in 160 Ma.

boundary uncertainty may be induced by the overlap between the Saharo-Arabian and Holarctic floras (Supplementary Data 5, also see Supplementary Fig. 19), and therefore, future investigations at regional scales are needed to further clarify the northern boundary of the Saharo-Arabian realm.

Fourth, the Cape region is ranked as one of six “kingdoms” in Takhtajan’s map<sup>4</sup>, but as a sub-realm of the African realm here. Our result is consistent with Cox<sup>1</sup>. The Cape region has not been geographically separated from Africa and shares similar tectonic history with the African continent<sup>31</sup>. The endemism in the Cape flora is primarily observed at species level, while the number of endemic taxa at higher levels, e.g., genus and family, is much lower than in other

realms<sup>1,48</sup>. Many genera with high proportions of endemic species in the Cape region are also widely distributed in Africa, such as *Erica*, *Protea*, *Helichrysum*<sup>49</sup>. However, most endemic genera in the Cape region contain only very few species<sup>1</sup>. These findings suggest that the Cape flora may not have higher evolutionary distinctiveness at higher taxonomic (e.g., the genus or family) levels compared with floras in other African regions<sup>1</sup>.

A comparison between our floristic map and the recently updated zoogeographic map (9) indicates several consistencies, suggesting that there are common drivers of terrestrial plant and animal biogeographic patterns. Specifically, the Saharo-Arabian and Indo-Malesian realms are identified in both maps and the boundaries of these realms



**Fig. 3 | The sequential appearance of the present floristic dissimilarities between different realms evaluated using non-metric multidimensional scaling (NMDS) ordination.** The Euclidean distance between dots is positively associated with the dissimilarity in phylogenetic composition of the flora between them: the larger the distance, the higher is the dissimilarity. Each line represents the distance of present phylogenetic composition of a realm to other realms at different phylogenetic depths. The colors of the lines are consistent with the colors of the realms shown in Fig. 1a, and the color gradient of each thick line represents evolutionary time. Source data are provided as a Source Data file.

are largely consistent (Supplementary Fig. 11). Moreover, the Cape region is not recognized as a realm in both maps. Notably, there are also interesting differences between our floristic and the zoogeographic maps. Our floristic map supports: (1) a Holarctic realm rather than the separated Palearctic, Nearctic and Sino-Japanese realms in the zoogeographic map; (2) an African realm rather than separated Afro-tropical and Madagascan realms; (3) an Indo-Malesian realm rather than separated Oriental and Oceanian realms. These differences may to some extent reflect the effects of different dispersal abilities of plants and vertebrates on biotic interchanges between regions<sup>1</sup>. Fossil and molecular evidence suggests that dispersal across land bridges separated by water is relatively common and has occurred in many plant clades<sup>39,40</sup>. In contrast, long-distance dispersal across sea water in vertebrates has been found to be biased to clades with specific locomotion, e.g., flight in birds<sup>50</sup>.

### The drivers on the division between floristic realms

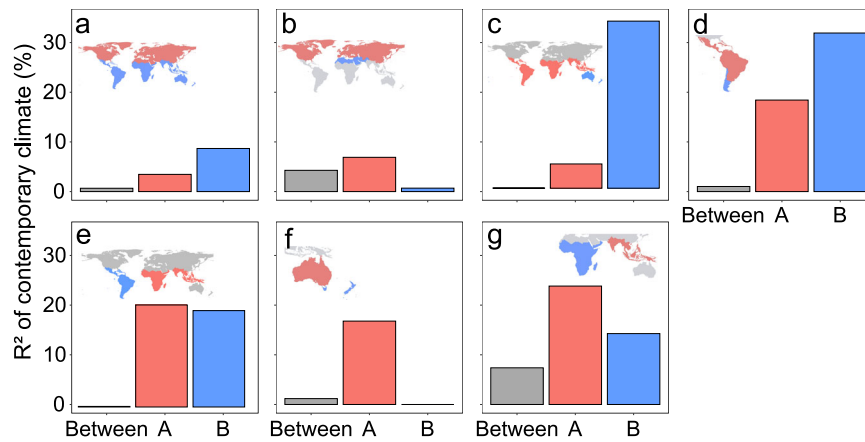
Contemporary climate explains considerable variations in phylogenetic beta diversity within the Neotropical ( $R^2 = 17.8\%$ ), Chile-Patagonian ( $R^2 = 32.9\%$ ), Indo-Malesian ( $R^2 = 23.8\%$ ), African ( $R^2 = 14.7\%$ ) and Australian realms ( $R^2 = 17.6\%$ ), but not within the Holarctic ( $R^2 = 6.6\%$ ) and Saharo-Arabian realms ( $R^2 = 1.3\%$ ). In contrast, contemporary climate has consistently extremely low explanatory power on the phylogenetic beta diversity between realms ( $R^2 < 8\%$  for all realm pairs). These results suggest that, although contemporary climate influences floristic variations within some realms, it is not a dominant driver for realm division (Fig. 4).

We then evaluated the relative effects of the other two factors on the division of realms, i.e., the historical climatic differences across space during geological times (historical climate hereafter) and historical geographic isolation induced by plate tectonics (geographic isolation hereafter). It is noteworthy that geographic isolation and historical climate may, to some extent, interlink with each other as plate tectonics may lead to shifts in landmasses and their climates. Our results indicate

that the correlations between temporal dynamics in historical climate and geographic isolation are generally low in most cases (Supplementary Fig. 12, Pearson  $r < 0.3$ ) except in the comparison between the Australian and Neovozealandic realms (Supplementary Fig. 12f, Pearson  $r = 0.56 \pm 0.01$ ). To better compare the effects of historical climate and geographic isolation on realm division, we conducted hierarchical partitioning analysis to estimate their independent effects (Fig. 5). We find that historical climate has weak effects on the division between most realms. For the division of the temperate realms (i.e., Saharo-Arabian, Chile-Patagonia, and Novozealandic realms, Fig. 5b, d, f), the effect of historical climate increased from the Oligocene to the present. This may be because the global climate started to become cooler and drier since the late Eocene and this trend intensified after the mid-Miocene<sup>31</sup>. Compared with historical climate, geographic isolation has stronger effects on the division between the Gondwanan and Laurasian super-realms (Fig. 5a), between the Saharo-Arabian and Holarctic realms (Fig. 5b), between the (Neotropical, Chile-Patagonian) and the (African, Indo-Malesian) realms (Fig. 5e), and between the (Australian, Novozealandic) realms and other realms of the Gondwanan super-realm (Fig. 5c). These results suggest that geographic isolation induced by plate tectonics has played a dominant role in the division of these super-realms and realms.

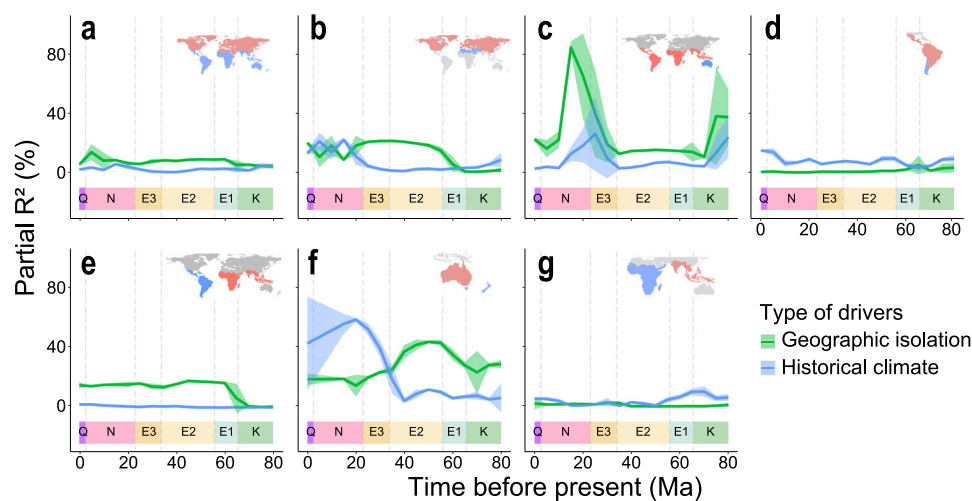
Geological evidence indicates that the ancient Tethys Seaway separated the Gondwana and Laurasia landmasses before the Cenozoic<sup>31,37</sup>, which may have led to the dominant effects of geographic isolation on the division between the Laurasian and Gondwanan super-realms (Fig. 5a) and on the division between the Holarctic and Saharo-Arabian realms (Fig. 5b). The breakup of Gondwana and the opening of the Atlantic Ocean may have enhanced the effects of geographic isolation on the division between the realms within the Gondwanan super-realm. (Fig. 5c, e, f). The present-day floras in South America and Africa cannot be distinguished from each other in the Middle Cretaceous (100 Ma, Fig. 2), but increasingly diverged from each other since 80 Ma (Fig. 3), which is possibly due to the reduced flora exchange caused by the expansion of the Atlantic Ocean as shown by the fossil evidence<sup>29,31</sup>. Hence the effect of geographic isolation induced by the expansion of the Atlantic Ocean on the division of the Neotropical and the Chile-Patagonian realms from African and Indo-Malesian realms within the Gondwanan super-realm also increased over time (Fig. 5e). The northward drift of the Australian plate and the southward drift of the Antarctic plate starting from the late Cretaceous combined with the onset of the Antarctic glaciation in the Oligocene (30–28 Ma) cut off the floristic exchange between Australia and other Gondwana landmasses (i.e., South America and Africa)<sup>31</sup>. This may have led to an increased effect of geographic isolation from the Oligocene to the mid-Miocene on the separation of the Australian and Novozealandic from the African and Indo-Malesian realms (Fig. 5c). New Zealand drifted away from the ancient Gondwana in the Late Cretaceous (80 Ma)<sup>29</sup>, likely leading to a higher contribution of geographic isolation on the evolution of its flora than historical climate before the Eocene (Fig. 5f).

Geographic isolation has weaker effects on the division between the Neotropical and the Chile-Patagonian realms than historical climate through geological times (Fig. 5d). These two realms have been geographically connected during most of the Cenozoic period, which may have led to the weak effects of geographic isolation. Interestingly, neither geographic isolation nor historical climate well explain the division between the African and Indo-Malesian realms (Fig. 5g). This may be due to biotic interchange between the floras of African and the Indo-Malesian realms. The northward drift of the Indian subcontinent since the early Cenozoic and its final collision with Eurasia brought a large number of floristic elements that are closely related to the African flora to the Indo-Malesian realm<sup>37,38</sup>, which may have led to the low explanatory power of both geographic isolation and historical climate on the division between these two realms.



**Fig. 4 | Effects of contemporary climate on the phylogenetic beta diversity within and between floristic realms (or clusters of realms).** The bars show the explained variance in the phylogenetic beta diversity within (red and blue) and between (gray) different realms (or clusters of realms) as shown by the inset maps. **a** Gondwanan super-realm vs. Laurasian super-realm; **b** Holarctic realm vs. Saharo-

Arabian realm; **c** (Australian, Novozealandic) realms vs. ((African, Indo-Malesian), (Neotropical, Chile-Patagonian)) realms; **d** Neotropical realm vs. Chile-Patagonian realm; **e** (African, Indo-Malesian) realms vs. (Neotropical, Chile-Patagonian) realms; **f** Australian realm vs. Novozealandic realm and **g** African realm vs. Indo-Malesian realm. Source data are provided as a Source Data file.



**Fig. 5 | Temporal changes in the relative effects of geographic isolation and historical climate on the phylogenetic beta diversity between floristic realms (or clusters of realms).** At each time interval of 1 Ma during the last 80 Ma, the partial  $R^2$  of geographic isolation and historical climate are evaluated using hierarchical partitioning model as  $\ln(\text{phylogenetic beta diversity}) - \ln(\text{Climate isolation}) + \ln(\text{Geographic isolation})$ . Colors of lines represent the independent  $R^2$  of geographic isolation (green) and historical climate (blue) on phylogenetic beta diversity, respectively. The lines and the shaded areas represent the mean  $\pm$  SE of

the  $R^2$  summarized every 5 Ma. **a** Gondwanan super-realm vs. Laurasian super-realm; **b** Holarctic realm vs. Saharo-Arabian realm; **c** (Australian, Novozealandic) realms vs. ((African, Indo-Malesian), (Neotropical, Chile-Patagonian)) realms; **d** Neotropical realm vs. Chile-Patagonian realm; **e** (African, Indo-Malesian) realms vs. (Neotropical, Chile-Patagonian) realms; **f** Australian realm vs. Novozealandic realm and **g** African realm vs. Indo-Malesian realm. Source data are provided as a Source Data file.

### The effects of clade evolution on the division of floristic realms

The clades with high contribution to realm divisions vary among realms. Specifically, we find that younger clades generally have higher contribution than older ones to the division of the (Australian, Novozealandic) realms from the other realms of the Gondwana super-realm (Supplementary Fig. 13c), the division between the (Neotropical, Chile-Patagonian) and the (African, Indo-Malesian) realms (Supplementary Fig. 13e), and the division between the African and Indo-Malesian realms (Supplementary Fig. 13g). This result corroborates the enhanced effect of geographic isolation caused by the breakup of Gondwana on realm divisions (Fig. 5). Among major angiosperm clades, malvids, campanulids and lamiids have high contributions to the divisions between African and Indo-Malesian realms ( $R^2 > 89.8\%$ ), between Neotropical and Chile-Patagonian realms ( $R^2 > 38.0\%$ ) and between Holarctic and Saharo-Arabian realms ( $R^2 > 54.7\%$ ), respectively (see Supplementary Data 3 for details). Notably, our study only

indicates the contribution of present-day clades to global realm divisions. Further investigations could compare the relative contribution of extinct and present-day clades by integrating contemporary and fossil evidence.

We present the global map of floristic regionalization quantitatively delineated using the distributions and phylogeny of global angiosperm genera. Global lands are divided into two super-realms and eight realms. The boundaries and the hierarchical divisions between the identified realms mostly reflect the effect of geographic isolation induced by plate tectonics over geological times rather than the effect of contemporary and historical climate. These findings together with the high consistency between the boundaries in floristic and zoogeographic realms suggest that geographic isolation during the geological history is likely a common driver for the formation of both floristic and zoological realms. Our global map of floristic realms provides a geographic framework for a wide variety of comparative

studies in historical and ecological biogeography, macroecology, and systematics.

## Methods

A full description of the methods, such as phylogenetic reconstruction, compilation of distribution data, identification of floristic realms and sub-realms, and evaluation on the sensitivity of realm boundaries, is included in the Supporting Methods.

### The phylogeny of global angiosperm genera

A genus-level phylogeny for seed plants was constructed using molecular data for 8 gene markers obtained from GenBank (May 19, 2018), including 18S rDNA, ITS (i.e., ITS1, 5.8S ribosomal DNA and ITS2), and 26S rDNA from the nuclear genome; *atpB*, *matK*, *ndhF* and *rbcL* from the chloroplast genome; and *matR* from the mitochondrial genome. Sequences from hybrids and taxa with dubious identification were excluded. To construct the genus-level dataset, we first assessed the monophyly of each genus following<sup>51</sup>. For monophyletic genera, one representative sequence per marker per genus (generally the longest one) was selected. For a non-monophyletic genus (totally 593 genera, 4.7%), we only selected species from its core or the largest monophyletic clade. This procedure ensured that we only combined sequences from species belonging to the same monophyletic lineage. Accession number of the sequences that were used in our molecular analyses are available in Supplementary Data 6.

The genus-level sequences were aligned within each order separately and then merged using MAFFT v7.4 with the most accurate L-INS-i strategy<sup>52</sup>. Phylogenetic analysis were partitioned by RAXML v8.0.26<sup>53</sup> with GTRGAMMA model. We constrained the phylogenetic analyses in RAXML v8.0.26 using the APG IV relationships among angiosperm orders and among eudicots, monocots and magnoliids. The tree was dated with treePL v1.0<sup>54</sup> using fossil calibrations from ref. 55. As the crown age of angiosperms is still debated<sup>56</sup>, we conducted three dating analyses with different constraints on the age of angiosperm crown: (1) between 149 Ma and 256 Ma following<sup>57</sup>; (2) between 140 Ma and 210 Ma following<sup>58</sup>; and (3) between 140 and 150 Ma following<sup>59</sup>. Seed plant genera without sequence data were added to the dated phylogenies as polytomies based on current taxonomy, and then were resolved using the polytomy resolver following<sup>60</sup>. The final molecular and full phylogenies contain 12,539 and 14,244 seed plant genera, respectively. As all results are consistent across the phylogenies, we reported results based on the phylogeny with a constraint of 140–210 Ma for angiosperm crown age, and others in the supplementary materials.

To further explore the potential influence of the fast-evolving genes (particularly ITS1 and ITS2) on phylogenetic topology, we reconstructed the molecular phylogeny using sequence data without ITS and dated it in the same way as previously described. Then we compared it with the phylogeny based on the full sequence dataset and found both the topologies and phylogenetic distances among genera to be highly consistent between the two phylogenies (see supplementary method for details).

### Global distributions of angiosperm genera

Geographic distributions of angiosperm species were compiled from >1100 sources, including published regional and local floras, floristic investigations, specimen records and online databases (see Supplementary Data 1 for the full list of data sources). The geographic standard units (hereafter GSUs) used for the compilation of species distributions were generated following<sup>61</sup>, and the average size of GSUs was ca. 4° longitude × 4° latitude. After removing small islands (<25,000 km<sup>2</sup>) and Antarctica, the earth's landmass was divided into 420 GSUs (see Supplementary Data 2). We classified the raw distributional data into four types: coordinates, range maps, gridded

distributions, and recorded localities. Depending on the types of the raw data, we applied different methods to reduce spatial conflicts between the original records and the boundaries of the GSUs used in our study and to improve the accuracy of species distributions in the final dataset (see Supplementary Methods for details of these methods). To improve the quality of species distribution data, we set a threshold for the number of data sources to keep an occurrence record of a species in a given GSU. For geographical units in Europe, Australia, China, South Africa, Madagascar and North America, an occurrence record of species in a geographical unit corroborated by at least 3 data sources was retained, leading to high confidence of the data quality in these regions. For the geographical units in Central America, Greenland, Amazon and Turkey, an occurrence record of species in a geographical unit corroborated by at least 2 data sources was retained, leading to medium confidence of the data quality in these regions. The entire data was retained for India, North and Central Africa, and Patagonia because of data deficiency in these regions, leading to relatively low confidence of the data quality in these regions.

The taxonomic status and the accepted names of species from all data sources were standardized following the World Flora Online (WFO, <http://www.worldfloraonline.org/>, accessed: December, 2022), Catalog of Life (COL, <https://www.catalogueoflife.org/>, accessed: May, 2018), ThePlantList (TPL, <http://www.theplantlist.org/>, accessed: Jan 3, 2015) and POWO (accessed: December, 2022). Synonyms are replaced with the accepted names. We kept accepted names with the highest confidence level. Taxonomic names that were identified as 'unresolved' in both COL and POWO were removed. The misspelt taxonomic names were corrected using the Taxonomic Name Resolution Service 4.0 (TNRS, <https://tnrs.biendata.org/>), which has been widely used in plant studies.

During data compilation, we also collected the status of species (i.e., being native, cultivated, introduced, invasive and hybrid) from regional data sources as much as we could, and non-native and hybrid species in different regions with clear records were not included in the database. After the compilation of distribution data at species level, we further checked the distribution maps and removed cultivated records from the database following the Plants of the World Online (POWO, <https://powo.science.kew.org/>, accessed: May, 2019) and efloras (<http://www.efloras.org/>, accessed: May, 2019).

Finally, we compiled the distributional data for each genus by aggregating distribution data of all its species. The distribution maps of all genera were carefully verified manually to improve data quality. We then integrated the phylogeny and distribution data, and the final distributional database contains 384,771 records for 12,664 angiosperm genera (see Supplementary Data 2), representing 90.63% of the total 13,974 accepted genera in POWO<sup>62</sup>.

### Taxonomic and phylogenetic beta diversity

Simpson beta diversity<sup>63</sup> was used to evaluate the dissimilarity between species assemblages of GSUs:

$$\text{Beta diversity} = 1 - \frac{a}{a + \min(b, c)} \quad (1)$$

where *a* represents species shared between two GSUs, *b* and *c* represent species unique to each GSU. We used this metric because it is not affected by the number of species and provides unbiased estimate of compositional turnover across space<sup>2,13</sup>. Using Equ. (1), pairwise matrices of taxonomic beta diversity (calculated using the number of shared (*a*) and unique (*b*, *c*) species between two GSUs) and phylogenetic beta diversity (calculated based on the length of shared (*a*) and unique (*b*, *c*) branches of angiosperm phylogenies between two GSUs) were generated separately.

## Identification of floristic realms and sensitivity analysis of realm boundaries

Hierarchical clustering analyses were conducted using the “hclust” function in stats (version 3.6.2) package in R 3.6.1<sup>64</sup> to group GSUs into floristic super-realms, realms, and sub-realms. With this method, GSUs with highest similarity (i.e., lowest distance) were first grouped together and then the most similar groups were grouped into clusters. This process was repeated until all GSUs were all grouped into a single cluster. The hierarchical clustering analyses were conducted for taxonomic and phylogenetic beta diversity between the GSUs, separately.

Following Holt et al.<sup>2</sup> we compared the performance and accuracy of different clustering methods. Performance evaluation aims to choose the clustering algorithm that can best represent the floristic divergences with the lowest number of clusters. For each clustering method, we calculated the proportion of beta diversity explained by identified clusters at a certain dendrogram height ( $P_{\text{beta}}$ ) as the sum of between-cluster beta diversity divided by the sum of beta diversity between all GSUs<sup>2</sup>. The best performing method was identified as the one returning the minimum number of clusters when  $P_{\text{beta}}$  reached 99%. Accuracy evaluation aims to choose the clustering algorithm that can represent the floristic divergences with the lowest biases. To do this, the co-phenetic correlation coefficients were calculated for each clustering algorithm using the “cophenetic” function<sup>65</sup> in stats (version 3.6.2) package in R 3.6.1<sup>64</sup>, and the clustering method with the highest accuracy has the highest co-phenetic coefficients.

We found that the “average” method, also known as the Unweighted Pair Group Means Algorithm (UPGMA) performed the best (Supplementary Fig. 14). Using the UPGMA, an unrooted dendrogram was generated. Floristic realms and sub-realms were identified by cutting the dendrogram at different heights, which were determined following the approach of<sup>2</sup>. Floristic realms and sub-realms were identified as the clusters required to reach  $P_{\text{beta}} = 80\%$  and  $P_{\text{beta}} = 95\%$ , respectively<sup>2</sup>. Then we used the Non-metric Multidimensional Scaling (NMDS hereafter) to illustrate the relationships between floristic realms in a two-dimension non-hierarchical space.

The GSUs on realm boundaries may contain mixed floristic components, which may lead to soft boundaries<sup>19,31</sup>. For comparison with the hierarchical clustering analysis, we redefined the floristic realms using the fuzzy c-means clustering method. Unlike the hierarchical method, which assigns GSUs exclusively to one cluster, the fuzzy c-means clustering method estimates the likelihood of each GSU belonging to a certain cluster, under a given degree of fuzziness<sup>66</sup>. We then conducted silhouette analysis for both results of UPGMA and fuzzy c-means clustering to evaluate the uncertainties in assigning GSUs to a single floristic realm<sup>67</sup>. The identification of floristic realms may also be influenced by uncertainties in phylogenetic topology and incomplete sampling across lineages<sup>24</sup>. Therefore, we repeated the hierarchical clustering analysis using: (1) the single maximum credibility tree; (2) the trees containing only genera with molecular data, (3) randomly sampled post burn-in posterior trees from the polytomy resolver, and (4) trees containing only the genera used in a recent study<sup>24</sup>.

## Changes in floristic realms through time

To explore the phylogenetic depth at which the spatial divergence of realms appears, we cut the phylogenetic trees at different geological times and generated the maps of floristic realms at each time following the approach used in refs. 34, 68–71. This approach collapsed all the descendent leaves of each branch encountered at a given time and then identified the floristic realms using the UPGMA clustering method as described above. We then illustrated the changes in the dissimilarity between the floras of different realms by overlaying the NMDS ordinations conducted at different geological times. Note that this analysis

did not intend to estimate the ancestral geographic ranges of phylogenetic branches or the ancestral floristic assemblages.

## Effects of contemporary climate, historical climate, and geographic isolation on realms divisions

If the contemporary climate dominates the division between floristic realms, contemporary climate should have higher explanatory power on the beta diversity between than within realms. The climate difference of a given pair of GSUs was defined as the Euclidean distance between them in a two-dimensional (i.e., annual mean temperature and precipitation) climatic space. Then we evaluated the effect of contemporary climate on beta diversity between and within realms using Ordinary Least Square (OLS) regressions with phylogenetic beta diversity as the response variable and contemporary climate differences as the predictor.

Following the method in ref. 72 and the reconstructed map in ref. 73, we produced paleogeographic maps with a spatial resolution of  $1 \times 1$  degree from 80 Ma up to the present in a 1 Ma step. Then, at every geological time, we calculated the geographic isolation between each two grid cells as the minimum total route cost between them, which was a function of geographic distance, separation by ocean, and elevation of the grid cells on the route. The geographic isolation was estimated using the ‘gdistance’ R package (version 1.6)<sup>74</sup>. Paleotemperature and paleo-precipitation at 1 Ma step from 80 Ma to the present were reconstructed by ref. 72. We calculated the Pearson correlation of geographic isolation and historical climate differences for each pair of floristic realms from 80 Ma to the present, and then hierarchical partitioning analysis was used to compare the relative effects of geographic isolation and historical climate at each geological time on the division of floristic realms. Specifically, phylogenetic beta diversity between GSUs of two realms was used as the response variable, and geographic isolation and historical climate distances between GSUs were used as the predictors.

## Contribution of different clades to floristic realms division

We first detected angiosperm clades whose descendant lineages have little overlap in their geographic distributions, and then measured their contribution to the division of floristic realms as the degree of consistency between the realm boundaries and the geographic divergences of their descendant lineages. Specifically, the geographic divergences of descendant lineages of a clade was evaluated using node-based analysis conducted in the “nodiv” R package (version 1.4.0)<sup>34</sup>. Specifically, this method calculates the degree of mismatch in the geographic distribution of two sister lineages diverging at a given node, i.e., the ‘geographic node divergence’ (GND) score, and the specific overrepresentation scores (SOS) for each geographic unit occupied by the two sister lineages. GND values over 0.65 indicate significant distributional divergence between the two sister lineages. Positive (or negative) SOS values indicate predominance of one of the two descendant lineages in a geographic unit (see Supplementary Fig. 15). Using node-based analysis, we detected the clades with GND scores over 0.65 and extracted the SOS scores of all GSUs occupied by these clades. The contribution of each of these clades on the division of two floristic realms was measured by the  $R^2$  of the one-way analysis of variance (ANOVA) with SOS scores as the response variable, and the two realms as the predictor.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The distribution data can be found in Supplementary Data 2. The phylogeny and the



shapefiles of the floristic realms identified in this paper are available at <https://en.geodata.pku.edu.cn/index.php?c=content&a=list&catid=199> (User: flowertree; Password: flowertree). DNA sequence data were downloaded using from GenBank (as of May 19, 2018) and the Accession numbers for DNA sequences can be found in Supplementary Data 6. Global Administrative Areas boundaries were downloaded from <http://www.gadm.org> (accessed: May 2016). Distribution data was obtained from both on-line databases and directly from the literature and the complete list of distributional data sources is provided as supplementary data. Species distribution data recorded as locality names were searched in the global geographical names service <http://www.geonames.org>. The taxonomic status and the accepted names of species from all data sources were standardized following the *World Flora Online* (WFO, <http://www.worldfloraonline.org/>, accessed: December, 2022), *Catalog of Life* (COL, <https://www.catalogueoflife.org/>, accessed: May, 2018), *ThePlantList* (TPL) available at <http://www.theplantlist.org/> (accessed: Jan 3, 2015) and POWO (accessed: December, 2022). Climate data was downloaded from the WorldClim database v2.0 (<https://www.worldclim.org/>, accessed: December, 2022). Paleo-digital elevation models are obtained from Scotese's paleoatlas<sup>73</sup>. Paleoclimate data are obtained from<sup>72</sup>. Source data are provided with this paper.

### Code availability

All code needed to evaluate the conclusions in the paper can be found in <https://github.com/yunpengliu1994/regionalization> (<https://doi.org/10.5281/zenodo.7758185>)<sup>75</sup>.

### References

- Cox, B. The biogeographic regions reconsidered. *J. Biogeogr.* **28**, 511–523 (2001).
- Holt, B. G. et al. An update of Wallace's Zoogeographic regions of the world. *Science* **339**, 74–78 (2013).
- Briggs, J. C. Darwin's biogeography. *J. Biogeogr.* **36**, 1011–1017 (2009).
- Takhtajan, A., Crovello, T. J. & Cronquist, A. *Floristic Regions of The World*. 544 (University of California Press, 1986).
- Desmond, A., Moore, J. & Browne, J. *Charles Darwin* (Oxford University Press, 2007).
- Morrone, J. J. Biogeographical regions under track and cladistic scrutiny. *J. Biogeogr.* **29**, 149–152 (2002).
- Grenyer, R. et al. Global distribution and conservation of rare and threatened vertebrates. *Nature* **444**, 93–96 (2006).
- Wilson, K. A., McBride, M. F., Bode, M. & Possingham, H. P. Prioritizing global conservation efforts. *Nature* **440**, 337–340 (2006).
- Nelson, G. From Candolle to croizat: comments on the history of biogeography. *J. Hist. Biol.* **11**, 269–305 (1978).
- De Candolle, A. P. *Essai élémentaire de géographie botanique*. (éditeur non identifié, 1820).
- Good, R. *Geography of the Flowering Plants*. (Longmans, 1964).
- Engler, A. & Prantl, K. A. E. *Die Natürlichen Pflanzenfamilien nebst ihren Gattungen und wichtigeren Arten: insbesondere den Nutzpflanzen* (Wilhelm Engelmann, 1899).
- Kreft, H. & Jetz, W. A framework for delineating biogeographical regions based on species distributions. *J. Biogeogr.* **37**, 2029–2053 (2010).
- Slik, J. W. F. et al. Phylogenetic classification of the world's tropical forests. *Proc. Natl Acad. Sci. USA* **115**, 1837–1842 (2018).
- Daru, B. et al. A novel phylogenetic regionalization of phytogeographical zones of southern Africa reveals their hidden evolutionary affinities. *J. Biogeogr.* **43**, 155–166 (2016).
- Daru, B. H., Karunarathne, P. & Schliep, K. phyloregion: R package for biogeographical regionalization and macroecology. *Method. Ecol. Evol.* **11**, 1483–1491 (2020).
- Ye, J. et al. Phylogenetic delineation of regional biota: a case study of the Chinese flora. *Mol. Phylogenet. Evol.* **135**, 222–229 (2019).
- Vilhena, D. A. & Antonelli, A. A network approach for identifying and delimiting biogeographical regions. *Nat. Commun.* **6**, 1–9 (2014).
- Morrone, J. J. Biogeographical regionalisation of the world: a reappraisal. *Aust. Syst. Botany* **28**, 81–90 (2015).
- Ficetola, G. F., Mazel, F. & Thuiller, W. Global determinants of zoogeographical boundaries. *Nat. Ecol. Evol.* **1**, 1–7 (2017).
- Daru, B. H., Elliott, T. L., Park, D. S. & Davies, T. J. Understanding the processes underpinning patterns of phylogenetic regionalization. *Trend. Ecol. Evol.* **32**, 845–860 (2017).
- Li, R., Kraft, N. J., Yang, J. & Wang, Y. A phylogenetically informed delineation of floristic regions within a biodiversity hotspot in Yunnan, China. *Sci. Rep.* **5**, 1–7 (2015).
- Kubota, Y., Hirao, T., Fujii, S. J., Shiono, T. & Kusumoto, B. Beta diversity of woody plants in the Japanese archipelago: the roles of geohistorical and ecological processes. *J. Biogeogr.* **41**, 1267–1276 (2014).
- Carta, A., Peruzzi, L. & Ramirez-Barahona, S. A global phylogenetic regionalization of vascular plants reveals a deep split between Gondwanan and Laurasian biotas. *New Phytol.* **233**, 1494–1504 (2022).
- Antonelli, A. Biogeography: drivers of bioregionalization. *Nat. Ecol. Evol.* **1**, 1–2 (2017).
- Graham, C. H. & Fine, P. V. Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecol. Lett.* **11**, 1265–1277 (2008).
- Ricklefs, R. E. History and diversity: explorations at the intersection of ecology and evolution. *Am. Nat.* **170**, S56–S70 (2007).
- Descobes, P., Leprieur, F., Albouy, C., Heine, C. & Pellissier, L. Spatial imprints of plate tectonics on extant richness of terrestrial vertebrates. *J. Biogeogr.* **44**, 1185–1197 (2017).
- Sanmartin, I. & Ronquist, F. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Syst. Biol.* **53**, 216–243 (2004).
- Brown, J. H. *Macroecology*. (University of Chicago Press, 1995).
- Lomolino, M. V., Riddle, B. R. & Whittaker, R. J. *Biogeography* (Sinauer Associates, 2016).
- Currie, D. J. & Paquin, V. Large-scale biogeographical patterns of species richness of trees. *Nature* **329**, 326–327 (1987).
- Futuyma, D. J. & Kirkpatrick, M. (Sinauer Associates, Sunderland, 2005).
- Borregaard, M. K. et al. Node-based analysis of species distributions. *Method. Ecol. Evol.* **5**, 1225–1235 (2015).
- Holt, B. G. et al. Response to comment on "An update of Wallace's zoogeographic regions of the world". *Science* **341**, 343–343 (2013).
- Xing, Y., Gandolfo, M. A. & Linder, H. P. The Cenozoic biogeographical evolution of woody angiosperms inferred from fossil distributions. *Glob. Ecol. Biogeogr.* **24**, 1290–1301 (2015).
- Scotese, C. R. *Atlas of Earth History* (PALEOMAP project, 2001).
- Li, J. T. et al. Diversification of rhacophorid frogs provides evidence for accelerated faunal exchange between India and Eurasia during the Oligocene. *Proc. Natl Acad. Sci. USA* **110**, 3441–3446 (2013).
- Antonelli, A. et al. Amazonia is the primary source of Neotropical biodiversity. *Proc. Natl Acad. Sci. USA* **115**, 6034–6039 (2018).
- Roncal, J., Nieto-Blázquez, M. E., Cardona, A. & Bacon, C. D. in *Neotropical Diversification: Patterns and Processes*. 521–546 (Springer, 2020).
- Bacon, C. D. et al. Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proc. Natl Acad. Sci. USA* **112**, 6110–6115 (2015).
- Ali, J. R. & Hedges, S. B. Colonizing the Caribbean: new geological data and an updated land-vertebrate colonization record challenge the GAARlandia land-bridge hypothesis. *J. Biogeogr.* **48**, 2699–2707 (2021).

43. Morley, R. J. Interplate dispersal paths for megathermal angiosperms. *Perspect. Plant Ecol. Evol. System.* **6**, 5–20 (2003).
44. Garrocc, C. et al. Genetic relations between the Aves Ridge and the Grenada back-arc Basin, East Caribbean Sea. *J. Geophys. Res.* **126**, e2020JB020466 (2021).
45. Popov, S. V. et al. Lithological-paleogeographic maps of Paratethys-10 maps late Eocene to pliocene. *Courier Forschungsinstitut Senckenberg* **250**, 1–46 (2004).
46. Le Houérou, H. Outline of the biological history of the Sahara. *J. Arid Environ.* **22**, 3–30 (1992).
47. Le Houérou, H. N. Climate, flora and fauna changes in the Sahara over the past 500 million years. *J. Arid Environ.* **37**, 619–647 (1997).
48. Goldblatt, P. & Manning, J. *Cape Plants: a Conspectus of the Cape flora of South Africa* (National Botanical Institute, 2000).
49. Linder, H. P. The evolution of African plant diversity. *Frontiers in Ecology and Evolution* **2**, <https://doi.org/10.3389/fevo.2014.00038> (2014).
50. Medina, I., Cooke, G. M. & Ord, T. J. Walk, swim or fly? Locomotor mode predicts genetic differentiation in vertebrates. *Ecol. Lett.* **21**, 638–645 (2018).
51. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Botany* **105**, 302–314 (2018).
52. Katoh, K., Kuma, K.-i., Miyata, T. & Toh, H. Improvement in the accuracy of multiple sequence alignment program MAFFT. *Genome Informat.* **16**, 22–33 (2005).
53. Silvestro, D. & Michalak, I. raxmlGUI: a graphical front-end for RAxML. *Org. Divers. Evol.* **12**, 335–337 (2012).
54. Smith, S. A. & O’Meara, B. C. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**, 2689–2690 (2012).
55. Zanne, A. E. et al. Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).
56. Foster, C. S. P. et al. Evaluating the impact of genomic data and priors on Bayesian estimates of the angiosperm evolutionary time-scale. *Syst. Biol.* **66**, 338–351 (2017).
57. Barba-Montoya, J., dos Reis, M., Schneider, H., Donoghue, P. C. J. & Yang, Z. Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. *New Phytol.* **218**, 819–834 (2018).
58. Stevens, P. *Angiosperm phylogeny website. Version 13*, <http://www.mobot.org/MOBOT/research/APweb/> (2015).
59. Silvestro, D., Cascales-Miñana, B., Bacon, C. D. & Antonelli, A. Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. *New Phytol.* **207**, 425–436 (2015).
60. Kuhn, T. S., Mooers, A. Ø. & Thomas, G. H. A simple polytomy resolver for dated phylogenies. *Meth. Ecol. Evol.* **2**, 427–436 (2011).
61. Shrestha, N. et al. Global patterns of Rhododendron diversity: The role of evolutionary time and diversification rates. *Global Ecol. Biogeogr.* **27**, 913–924 (2018).
62. Powo. *Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew.*, <http://www.plantsoftheworldonline.org/> (Powo, 2023).
63. Lennon, J. J., Koleff, P., Greenwood, J. J. D. & Gaston, K. J. The geographical structure of British bird distributions: diversity, spatial turnover and scale. *J. Anim. Ecol.* **70**, 966–979 (2001).
64. Team, R. C. (R Foundation for Statistical Computing, Vienna, Austria, 2021).
65. Sneath, P. H. A. & Sokal, R. R. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. 278 (Freeman, 1973).
66. Bezdek, J. C. *Pattern Recognition with Fuzzy Objective Function Algorithms*. (Springer Science & Business Media, 2013).
67. Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).
68. Groussin, M. et al. Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nat. Commun.* **8**, 1–12 (2017).
69. Mazel, F. et al. Global patterns of  $\beta$ -diversity along the phylogenetic time-scale: the role of climate and plate tectonics. *Glob. Ecol. Biogeogr.* **26**, 1211–1221 (2017).
70. Daru, B. H., Michelle, V. D. B., Davies, T. J. & Franklin, J. Unravelling the evolutionary origins of biogeographic assemblages. *Divers. Distrib.* **24**, 313–324 (2018).
71. He, J., Lin, S., Li, J., Yu, J. & Jiang, H. Evolutionary history of zoogeographical regions surrounding the Tibetan Plateau. *Commun. Biol.* **3**, 1–9 (2020).
72. Hagen, O. et al. Mountain building, climate cooling and the richness of cold-adapted plants in the Northern Hemisphere. *J. Biogeogr.* **46**, 1792–1807 (2019).
73. Scotese, C. R. *PALEOMAP paleodigital elevation models (PaleoDEMS) for the Phanerozoic.*, <https://www.earthbyte.org/paleodem-resource-scotese-and-wright-2018> (2018).
74. van Etten, J. R. Package gdistance: distances and routes on geographical grids. *J. Stat. Softw.* **76**, 21 (2017).
75. Liu, Y. et al. An updated floristic map of the world. *GitHub*. <https://doi.org/10.5281/zenodo.7758185>. (2023).

## Acknowledgements

This work was supported by the National Key Research Development Program of China (#2022YFF0802300), National Natural Science Foundation of China (#31988102, #32125026, #31770566), and the Strategic Priority Research Program of Chinese Academy of Sciences (#XDB31000000). D.D received additional support by the Norwegian Metacenter for Computational Science (NOTUR; project NN9601K). MKB, CR and YL acknowledge the Danish National Research Foundation (DNRF96) and VILLUM FONDEN (25925) for support of the Center for Macroecology, Evolution and Climate.

## Author contributions

Z.W., C.R., and Y.L. conceived the idea; Z.W. Y.L., X.X., X.S., and A.L. constructed the distribution dataset; Z.W., X.X., and D.D. conducted the phylogeny; L.P. and N.Z. conducted the palaeogeographical and paleoclimate data; Y.L. and M.B. conducted the node-based analysis; Y.L. and N.S. generated the figures in the manuscript; Y.L. led the analysis and writing, and all authors contributed to the writing.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-38375-y>.

**Correspondence** and requests for materials should be addressed to Carsten Rahbek or Zhiheng Wang.

**Peer review information** *Nature Communications* thanks Tiina Sarkinen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023