

10-9-2019

## Accumulation over evolutionary time as a major cause of biodiversity hotspots in conifers

Mekala Sundaram  
*Brown University*

Michael J. Donoghue  
*Yale University*

Aljos Farjon  
*Royal Botanic Gardens, Kew*

Denis Filer  
*University of Oxford*

Sarah Mathews  
*Commonwealth Scientific and Industrial Research Organization*

*See next page for additional authors*

Follow this and additional works at: [https://digitalcommons.lsu.edu/biosci\\_pubs](https://digitalcommons.lsu.edu/biosci_pubs)

---

### Recommended Citation

Sundaram, M., Donoghue, M., Farjon, A., Filer, D., Mathews, S., Jetz, W., & Leslie, A. (2019). Accumulation over evolutionary time as a major cause of biodiversity hotspots in conifers. *Proceedings of the Royal Society B: Biological Sciences*, 286 (1912) <https://doi.org/10.1098/rspb.2019.1887>

This Article is brought to you for free and open access by the Department of Biological Sciences at LSU Digital Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of LSU Digital Commons. For more information, please contact [ir@lsu.edu](mailto:ir@lsu.edu).

---

**Authors**

Mekala Sundaram, Michael J. Donoghue, Aljos Farjon, Denis Filer, Sarah Mathews, Walter Jetz, and Andrew B. Leslie

## Research



**Cite this article:** Sundaram M, Donoghue MJ, Farjon A, Filer D, Mathews S, Jetz W, Leslie AB. 2019 Accumulation over evolutionary time as a major cause of biodiversity hotspots in conifers. *Proc. R. Soc. B* **286**: 20191887. <http://dx.doi.org/10.1098/rspb.2019.1887>

Received: 13 August 2019

Accepted: 19 September 2019

**Subject Category:**

Evolution

**Subject Areas:**

ecology, evolution

**Keywords:**

biodiversity hotspot, richness, conifers, accumulation

**Author for correspondence:**

Mekala Sundaram

e-mail: msundara@stanford.edu

<sup>†</sup>Present address: Geological Sciences, Stanford University, Stanford, CA 94305, USA.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.4678910>.

# Accumulation over evolutionary time as a major cause of biodiversity hotspots in conifers

Mekala Sundaram<sup>1,†</sup>, Michael J. Donoghue<sup>2</sup>, Aljos Farjon<sup>3</sup>, Denis Filer<sup>4</sup>, Sarah Mathews<sup>5</sup>, Walter Jetz<sup>2</sup> and Andrew B. Leslie<sup>1,†</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, Brown University, Box G-W, 80 Waterman Street, Providence, RI 02912, USA

<sup>2</sup>Department of Ecology and Evolutionary Biology, Yale University, P.O. Box 208106, New Haven, CT 06520, USA

<sup>3</sup>Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AE, UK

<sup>4</sup>Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

<sup>5</sup>CSIRO National Research Collections Australia, Australian National Herbarium, Canberra, Australian Capital Territory 2601, Australia

MS, 0000-0003-4235-8783

Biodiversity hotspots are important for understanding how areas of high species richness form, but disentangling the processes that produce them is difficult. We combine geographical ranges, phylogenetic relationships and trait data for 606 conifer species in order to explore the mechanisms underlying richness hotspot formation. We identify eight richness hotspots that overlap known centres of plant endemism and diversity, and find that conifer richness hotspots occur in mountainous areas within broader regions of long-term climate stability. Conifer hotspots are not unique in their species composition, traits or phylogenetic structure; however, a large percentage of their species are not restricted to hotspots and they rarely show either a preponderance of new radiating lineages or old relictual lineages. We suggest that conifer hotspots have primarily formed as a result of lineages accumulating over evolutionary time scales in stable mountainous areas rather than through high origination, preferential retention of relictual lineages or radiation of species with unique traits, although such processes may contribute to nuanced differences among hotspots. Conifers suggest that a simple accumulation of regional diversity can generate high species richness without additional processes and that geography rather than biology may play a primary role in hotspot formation.

## 1. Introduction

One of the most striking features of biodiversity is its uneven distribution across the planet, from global patterns like the latitudinal species gradient to regional 'hotspots' of richness [1–4]. Understanding why spatial variability exists has been a major focus of research, because it may provide keys for understanding diversification in a broader sense [3–5] and because many high diversity regions are important targets for conservation [6–8]. Regions of high species richness are generally thought to form through high diversification rates, low extinction rates, high net immigration rates or some combination of these processes [4,9,10], but disentangling them is difficult due to the complex histories of lineages and of landscapes.

Nevertheless, a large body of literature has developed exploring how these processes might contribute to spatial variability in biodiversity, particularly with regard to lineage preservation or origination [10–14]. In plants, for example, studies have implicated long-term climatic stability as a major driver of endemism and increased richness; mature radiations are thought to occur in climatically and geologically stable environments due to low extinction rates, while more recent and rapid radiations may also occur within such broadly stable areas if new

ecological opportunities arise within them, such as newly exposed land [10,13,15–19]. If climatic stability is an important factor for the origination or persistence of richness hotspots, then hotspot species might also be expected to exhibit unique suites of traits associated with stability, such as increased habitat specialization, poor dispersal and mechanisms preventing inbreeding [10,20–23]. Teasing apart potential associations among traits, geography, climatic stability, endemism and high richness is challenging because it requires integrating a large number of disparate datasets.

Focused studies of individual lineages provide one way to combine detailed geographical, phylogenetic and trait data [20–23]. Conifers are an especially promising group in this regard, because they are diverse (approx. 615 living species [24]), globally distributed, show high spatial variability in species richness [25], and their phylogenetic relationships are generally well resolved (e.g. refs. [26,27]). A wide variety of conifer trait data is also available, including those that have been associated with endemism and diversification [24,28,29]. In this study, we take advantage of these datasets to explore how richness hotspots form in conifers, and, by extension, in woody plants. We combine detailed geographical range data for nearly all extant conifer species (90%) with a comprehensive time-calibrated molecular phylogeny, and ask if the spatial and phylogenetic structure of hotspot diversity is best explained by the preservation of relict lineages, the recent diversification of young lineages, or whether hotspot richness forms instead through a broad sampling of regional diversity without unique diversification processes. Using the trait data available for conifers, we also test if hotspots preferentially contain species with characteristics associated with persistence in stable areas.

## 2. Material and methods

Conifer species geographical ranges were based on point occurrences from carefully vetted herbarium specimens [25] combined with georeferenced records from Global Biodiversity Information Facility (GBIF, [www.gbif.org](http://www.gbif.org)). Conifers in this study include groups traditionally recognized as Pinales and exclude Gnetales. We estimated species ranges using  $\alpha$ -hulls [30–32] and by digitizing based on previously published range maps (a detailed explanation of methodology can be found in the electronic supplementary material). After generating species ranges, we summarized conifer assemblages in 100 km  $\times$  100 km grid cells across the globe using a cylindrical equal area projection and by tallying conifer ranges that overlapped each grid cell. Richness hotspots were identified by computing the local Getis-Ord  $G_i^*$  statistic on grid cell species richness in ‘hotspot analysis’ in ARCMAP 10.4.1, using an inverse squared distance weighting process and after controlling for false discovery rate using the Benjamini–Hochberg correction [33–35].

To test whether hotspots were distinct from non-hotspots, we first compared their basic physical features. For these analyses, we divided the world into separate biogeographic regions, each with unique species pools, and then compared hotspots and non-hotspot grid cells within these. We computed these regions using the R package ‘betapart’ [36] by calculating pairwise grid cell taxonomic and phylogenetic  $\beta$  diversity measures with Simpson’s dissimilarity metric, which accounts for taxonomic and phylogenetic turnover [36–39]. We then used UPGMA and McQuitty methods of hierarchical agglomerative clustering to group grid cells based on taxonomic dissimilarities, as the latter allows for unequal cluster sizes [40]. To group grid cells based on phylogenetic dissimilarities, we used a Ward’s D method [40]. Ultimately, we defined a total of 50 global taxonomic regions

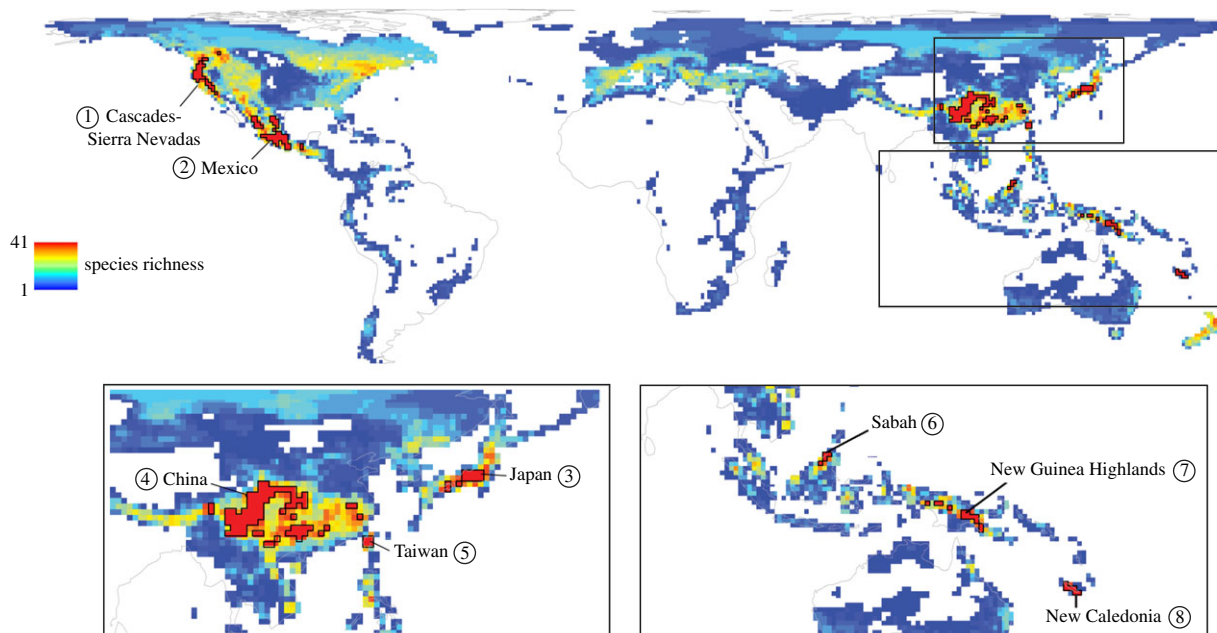
using UPGMA, 30 taxonomic regions using McQuitty and 30 phylogenetic regions. We present results from the taxonomic clustering approach in this paper because these regions were closer to traditional conifer biogeographic provinces and results from the phylogenetic clustering approach were qualitatively similar.

Within biogeographic regions, we compared physical and climatic characteristics of hotspot grid cells with those non-hotspot grid cells. We obtained climate, soil, lithologic and topographic data from WorldClim [41], global lithological map database [42], Oak ridge national laboratory [43] and USGS GTOPO30 digital elevation model. The climate data include mean annual temperature, temperature seasonality, annual precipitation and precipitation seasonality for a grid cell, and topographic data consist of grid cell topographic heterogeneity (maximum elevation–minimum elevation). We also used geologic data to test if conifer hotspots were associated with high soil or bedrock variability or with an abundance of low-nutrient substrates, which may favour conifer abundance [44]; these include ultramafic plutonic and metamorphic rocks such as serpentinites and soils such as podzols (see electronic supplementary material for full list of soil and lithology types). After summarizing these variables in each grid cell and for each hotspot, we performed a spatial simultaneous autoregressive lag regression [45–47] explaining log-transformed climatic and physical variables as a function of an indicator variable (1 = grid cell is in a hotspot or 0 = grid cell is not; see electronic supplementary material for full details). Residuals of all regressions were tested for residual spatial autocorrelation as measured by Moran’s  $I$  [47,48] and for departures from normality. To ensure results were robust to model choice, we also performed regressions with a spatial simultaneous autoregressive error model [45–47].

We next tested for biological or ecological differences in hotspots. We first compared grid cell level endemism index, defined as the median of the inverse of species range sizes in a grid cell [49], for hotspot and non-hotspot regions using spatial simultaneous autoregressive lag regressions. We then compared several traits in hotspot versus non-hotspot species, including genome size, seed size, breeding system (monoecy versus dioecy) and dispersal syndrome (biotic versus abiotic), all of which have been associated with restricted or specialized endemic plant species in previous studies [10,20,21]. Seed size is associated with dispersal abilities and germination probability [50–53]; all things being equal, species with smaller abiotically dispersed and/or larger biotically dispersed seeds are more likely to move and less likely to be found in highly restricted areas [20,21]. Genome size has also been related to habitat restriction, because very large genomes can result in low diversification rates and high extinction risks [20,54–58]. We collected genome sizes (C-values) from the Kew Royal Botanic Gardens database for gymnosperms [59] (includes 251 conifer species) and the other traits from published compilations [24,28,29].

To analyse hotspot traits, we compiled the total number of species in a given hotspot and its taxonomic cluster and then used a phylogenetic generalized least-squares regression, assuming an OU model, to predict seed volume and genome size as a function of the proportion of the species range that occurs in the hotspot. We calculated hotspot range proportion using the presence of a species in hotspot grid cells or in a less restrictive two-cell wide buffer around hotspot grid cells. We also used phylogenetic generalized linear models in the R package ‘phylolm’ [60,61] to predict binary variables dioecy (1 = dioecious tree, 0 = monoecious tree) and biotic dispersal (1 = specialized for a biotic agent, 0 = unspecialized for biotic agent) as a function of proportion of hotspot range. We also performed a phylogenetic PCA to compare multivariate trait space of hotspot and non-hotspot species using R package ‘phytools’ [62].

To compare the phylogenetic structure of hotspots, we used several approaches that tested for differences between hotspots and their broader species pools. We first characterized grid cell



**Figure 1.** Global map of conifer species richness at the scale of  $100 \times 100$  km grid cells. Grid cells constituting the eight identified richness hotspots are outlined in black and labelled 1–8. Inset maps of East Asia and the Southeast Asian tropics are also shown for clarity. (Online version in colour.)

phylogenetic structure using the median and variance in evolutionary distinctiveness (ED) of all the species in a grid cell [63], calculated with the fair ‘proportion metric’ [64–66] using a recently published time-calibrated conifer phylogeny [67] based on three genes (*rbcL*, *matK*, 18S) that include 544 species (88% of extant diversity using the taxonomy of Farjon [24]). To test for significantly high or low median ED and variance ED, we generated rarefaction curves and 95% confidence bands of median ED and variance ED from random samples of species in three continent-level species pools, including North America, Eurasia and the Southern Hemisphere (see electronic supplementary material for full explanation). We also used the three continental-scale species pools as the basis of the CANAPE analysis, which computes phylogenetic endemism (phylogenetic diversity of a grid cell weighted by the global range associated with the branch) using the given phylogeny and a tree of equal branch lengths with the same topology. We tested for significance (in BIODIVERSE 2.0 [68]) for each metric by generating null distributions from random communities drawn from the continental-scale pool for each hotspot while maintaining constant species richness. For CANAPE, centres of ‘palaeo-endemism’ are defined as regions of significantly high phylogenetic endemism whose branch lengths are also significantly longer than a tree of equal branch lengths. By contrast, centres of ‘neo-endemism’ are areas of significantly high phylogenetic endemism whose branch lengths are significantly smaller than a tree of equal branch lengths [68]. The final CANAPE results here represent the three continental-scale analyses joined together in one map.

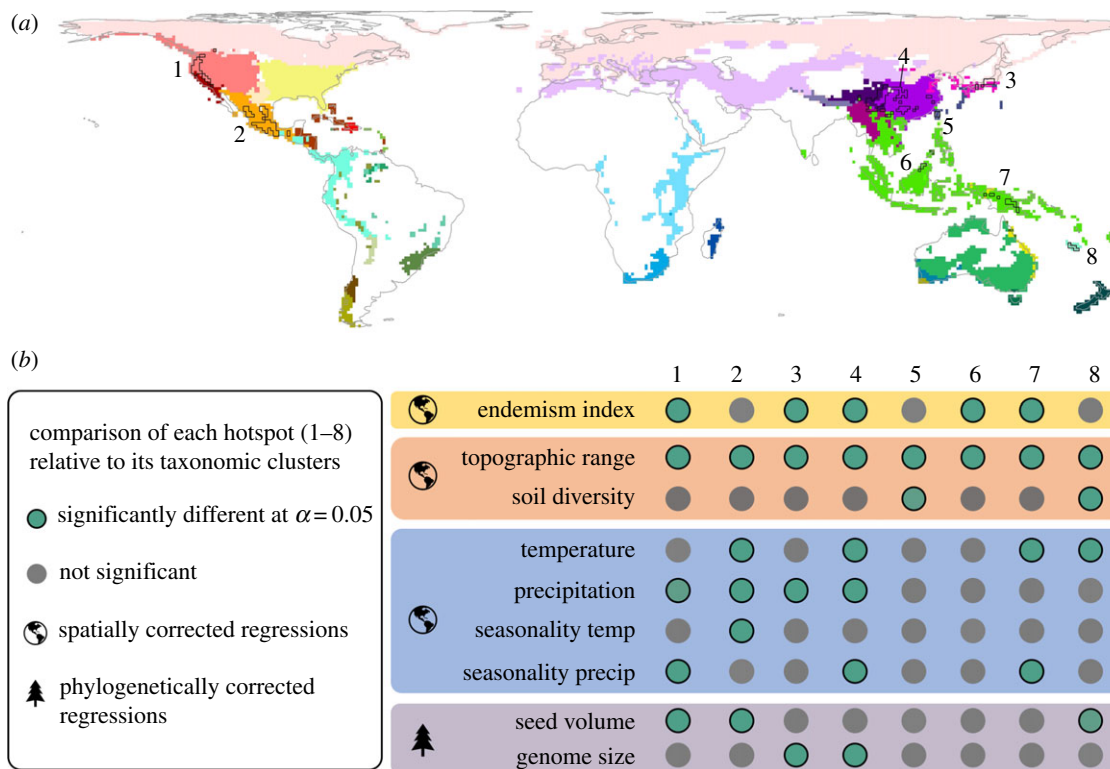
Finally, we explored how hotspot diversity may be influenced by the ecology and history of their broader biogeographic regions. Previous work shows that regional area, net primary productivity (NPP) and the integration of area and productivity through time (which would be impacted by major changes in climate such as glaciation) are important drivers of local species diversity [19,69–72]. We therefore tested if any of these variables explains maximum local conifer diversity among the biogeographic regions and their hotspots, by fitting negative binomial regressions and comparing models using AIC in R with library ‘MASS’. We chose maximum local diversity as the response variable in order to ask what factors produce high local richness areas and are likely to generate ‘hotspots’. For each biogeographic region, we determined maximum species richness in a grid cell, and then ranked models predicting this value as a function of regional area (number of grid cells within the region), average NPP of the region, total NPP of the

region (interaction of area and average NPP), proportion of grid cells not glaciated at the last glacial maximum (LGM), total area not glaciated (interaction of area and proportion of grid cells not glaciated at LGM), NPP of non-glaciated area (interaction of area, proportion of grid cells not glaciated, and average NPP) and total richness of biogeographic region. Data for NPP measures came from the Numerical Terradynamic Simulation Group [73,74] and data for glaciation from Ehlers *et al.* [75].

### 3. Results

We identify eight broad hotspots of conifer species richness, including southern and central Mexico, the Cascades-Sierra Nevada (CSN) mountain ranges, southern and central Japan, Taiwan, two regions of China (western and southern China), Sabah in Borneo, highlands in New Guinea and New Caledonia (figure 1). Hotspots were consistently associated with greater topographic heterogeneity than their surrounding regions but were not consistently associated with differences in climate, soil types or bedrock geology across multiple methods of defining regional pools (figure 2; electronic supplementary material, figure S1 and table S2). Most hotspots were associated with a greater proportion of endemic species (figure 2), but many hotspot species are also found outside them; species with 90% of their range within the hotspot grid cells never constitute more than 20% of the total pool of hotspot species (electronic supplementary material, table S3). Even if the hotspot grid cells are expanded by including a buffer of two grid cells around them, species with 90% of their range within the buffered hotspot grid cells never comprise more than 55% of the total species in a given hotspot (electronic supplementary material, table S3).

The biological characteristics of conifer hotspots are likewise not consistently different from non-hotspots, at least with regard to the traits studied here (figure 2; electronic supplementary material, tables S4–S11). Significantly larger genome sizes were found in hotspot-restricted species in only two hotspots (figure 2), while seed volume was significantly different in just three (figure 2; where each was associated with smaller



**Figure 2.** Physical and biological distinctiveness of conifer richness hotspots. (a) Map of 50 biogeographic regions based on a UPGMA hierarchical clustering analysis of grid cell taxonomic dissimilarities, with hotspots outlined in black and identified by numbers as in figure 1. (b) Significant differences between hotspot grid cells and the biogeographic region(s) in which they occur are shown for endemism, physical and climatic variables, and biological differences (seed size, genome size). Endemism index is defined as the median of the inverse of the geographical range sizes in a grid cell, topographic range is the difference between the highest and lowest elevations in a grid cell, soil diversity is diversity of zoller soil types in a grid cell, temperature ( $^{\circ}\text{C}$ ) is the median temperature in a grid cell, precipitation (mm) is the median precipitation in a grid cell, 'seasonality temp' is temperature seasonality or standard deviation of temperature  $\times 100$  and summarized as the median value for every grid cell, 'seasonality precip' is precipitation seasonality or coefficient of variation in precipitation and summarized as the median value for every grid cell. Significance for grid cell level attributes was tested using spatially corrected regressions. Genome size ( $C$ -value) and seed volume (cubic mm) were compared for hotspot species and non-hotspot species using phylogenetically corrected regressions. (Online version in colour.)

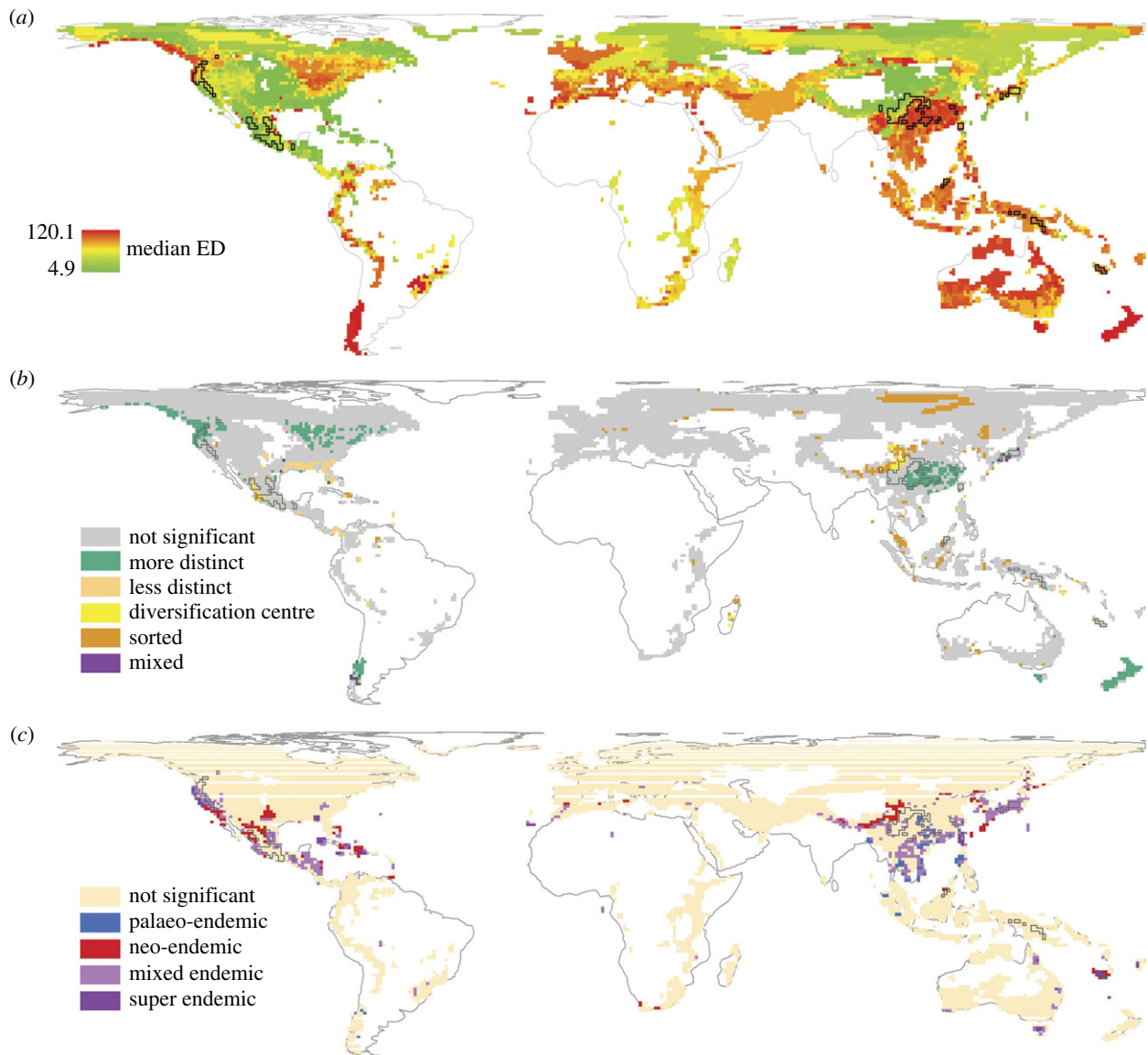
seed sizes). These relationships were not significant, however, when the proportion of each species range in the hotspot was calculated using a two grid cell buffer (electronic supplementary material, table S6) or when hotspots were broken up into their component biogeographic regions (electronic supplementary material, tables S8 and S11). Breeding strategy and dispersal syndrome also generally showed no significant relationships with the degree of hotspot restriction (electronic supplementary material, tables S5 and S7). Additionally, hotspot and non-hotspot species in aggregate did not occupy different areas of a phylogenetic PCA space (electronic supplementary material, figure S2).

The phylogenetic structure of conifer richness hotspots is also generally not unique relative to surrounding regions (figure 3). Among species that occur in hotspots, we find no relationship between the degree to which species are restricted to hotspots and their ED (electronic supplementary material, figure S3). Hotspots do not then preferentially contain relict lineages (high ED) or lineages nested within recent radiations (low ED), and species that are geographically restricted to hotspots tend to exhibit both high and low ED. At the assemblage level, the global distribution of grid cell median ED suggests that hotspots are not qualitatively different from their immediate surroundings (figure 3a). A rarefaction analysis of median ED is consistent with this impression, as 68% of hotspot grid cells in aggregate are not statistically different from a random sampling of their respective continental species pools (figure 3b; electronic supplementary material, figure S4

and table S12). On the other hand, the rarefaction analysis does reveal localized differences in structure (figure 3b); for example, grid cells in the Cascades, Japan and southern China have higher than expected median ED values, and some hotspot grid cells in western China, Borneo and Mexico show lower than expected median ED or lower than expected variance in ED (figure 3b; electronic supplementary material, figure S4 and table S12). A few grid cells in western China and in northern and western portions of the Mexican hotspot also exhibit both low ED and low variance in ED, consistent with the presence of a few diversifying lineages.

Using CANAPE, we found that statistically significant phylogenetic endemism among conifer assemblages occurs across the globe and is also found in parts of many hotspots (figure 3c). Five of the eight hotspots (Borneo, China, CSN, Mexico and New Guinea), however, show a high proportion of grid cells that lack significant PE (electronic supplementary material, table S12). Where significant PE does occur, communities most often contain mixtures of palaeo- and neo-endemics, and it is rare that communities exhibit pronounced palaeo-endemism or neo-endemism alone (figure 3c). Three hotspots on smaller islands (Japan, New Caledonia and Taiwan) do show significant PE in all constituent grid cells (reflecting small species ranges in these hotspots), but these grid cells also contain mixtures of palaeo- and neo-endemics (figure 3c).

Finally, maximum species richness at a grid cell level was best explained by the richness of the biogeographic region alone, rather than models that included NPP and/or glacial



**Figure 3.** (a) Global map of grid cell median ED for conifers, with hotspot grid cells outlined in black. (b) Results of rarefaction analysis of grid cell ED. ED rarefaction curves were generated by sampling species at random from a continental-scale species pool (Eurasia, North America, Tropics and Southern Hemisphere) and generating null distributions of median species ED and variance in species ED for 1000 random conifer communities. Grid cells showing no significant departures in median ED or variance ED compared with null distributions are represented as 'not significant'. Significantly high median ED grid cells are interpreted as 'More distinct' communities, significantly low median ED grid cells are interpreted as 'less distinct' communities, grid cells with both low median ED and low variance ED are interpreted as 'diversification centres', grid cells with low variance ED were interpreted as 'sorted' communities and finally grid cells with high variance ED were interpreted as 'mixed' communities. (c) Results for CANAPE analysis that identifies centres of 'palaeo-endemism', 'neo-endemism', mixed age endemics ('mixed endemic' and 'super endemic'), or no significant phylogenetic endemism or 'not significant' relative to continental species pools. Final CANAPE map was created by joining the results from individual continental pools. (Online version in colour.)

history (electronic supplementary material, table S13). All other models had an AIC weight  $\ll 0.001$  with no consistency in second or third best predictors across the different biogeographic region defining algorithms. The high relative importance of total richness of the region (AIC weight greater than 0.9) in explaining maximum local richness was unchanged regardless of the method used to define clusters (electronic supplementary material, table S13).

## 4. Discussion

Across all of our datasets, the most consistent feature of conifer hotspots is lack of consistent differences with their surrounding regions in terms of trait and phylogenetic structure. This suggests that conifer hotspots do not necessarily

form through the operation of unique diversification processes. For example, if hotspots generally resulted from high origination rates and the radiation of lineages, we would expect to see a preponderance of recently diverging, low ED and/or neo-endemic species compared with surrounding regions [3,10,18,76]. Likewise, if hotspot-specific ecological processes were responsible for hotspot formation, we might expect to consistently find species with unique traits [10,20]. Because such patterns are found only within certain regions of some hotspots (e.g. western Mexico, western China), we suggest instead that most hotspot diversity forms through a simpler process; they are areas that accumulate lineages found throughout their broader region over evolutionary time, with a more limited fraction of their total diversity resulting from *in situ* processes like diversification. The most unique feature of conifer hotspots then

appears to be the places in which they occur rather than any particular aspects of their biology or ecology.

Hotspot regions are likely to accumulate taxa because they are topographically heterogeneous and contain a high density of vertically stacked habitats. As conifer lineages have moved across landscapes in response to climatic fluctuations [77–79], current hotspots may simply represent those areas with the most niche space and therefore where the greatest number of taxa are likely to be found. This idea is consistent with the substantial fraction of species in hotspots whose ranges extend into broader surrounding regions and the relationship between high species richness and large species pools among biogeographic regions generally (electronic supplementary material, table S13). Recent simulations also suggest that high richness may result from a similar sampling of regional species pools, especially when combined with adaptation rate to heterogeneous microclimates [18]. Although we see little phylogenetic evidence of *in situ* diversification in most hotspots, some degree of local adaptation is likely to occur within them. For example, conifers do generally show evidence of elevational specialization within populations [22], with local adaptation to environmental gradients known from genomic studies [80] and trait data [22,81]. Such specialization within hotspots could potentially promote the retention of taxa within them [82], if species became adapted to narrow environmental conditions or elevational bands, although further studies are needed to explore this idea.

Conifer hotspots do not occur in all mountainous areas, however, suggesting that the raw capacity to harbour species is not the only factor determining hotspot formation. Prior work has suggested that high plant richness results from long-term climatic stability that promotes species differentiation and endemism [10,70,83–85], and our results are consistent with some of these ideas. Conifer hotspots do overlap with areas of climatic stability, centres of plant endemism [10] (electronic supplementary material, figure S5) and all occur in broad regions of low inferred climate change velocity [83], which is thought to both lower extinction rates and promote unique traits [10,86,87]. Hotspot species traits do not consistently suggest that they are adapted to restricted, stable environments or that they are otherwise different from non-hotspot species in their dispersal traits/potential. It therefore appears that climatic stability in hotspot areas results in lower extinction (although exact rates are notoriously difficult to directly calculate [88]), which when coupled with the high number of potential niches in these regions, promotes the accumulation and maintenance of lineages from regional pools regardless of their specific dispersal biology. The global conifer hotspots that we identify then reflect areas that are unique in their geography and history rather than their biology; they are where mountainous topography and climatic stability intersect with a large regional species pools. This may explain why mountains in Europe, for example, do not harbour global conifer richness hotspots, because their regional pool is depauperate compared with East Asia or western North America.

Conifer hotspots could then be thought of like Pleistocene glacial refugia [89–91], although they maintain and accumulate regional diversity over longer time periods. Indeed, the southern China conifer hotspot occurs within a region that has long been considered a global refugium for a variety of plant lineages, including notable ‘living fossil’ conifers like *Pseudolarix* and *Metasequoia* [92,93]. What is new in our results, however, is that almost all conifer hotspots could be thought of as these kinds of refugia, where their specific phylogenetic structure (i.e. whether they contain an abundance of old or young lineages) simply reflects that of their broader geographical region. We do not mean to suggest that all hotspot diversity is only due to aggregation; some grid cells within the Mexican hotspot (figure 3; electronic supplementary material, figure S4), for example, show evidence of recent diversification in a few clades (particularly pines [94]). Nevertheless, most grid cells in most hotspots are not unique relative to their broader species pools, so we envision simple regional sampling and accumulation as the first-order process responsible for building high richness, with more local processes adding additional diversity in some hotspots.

Given the association between topographic heterogeneity and species richness among organisms more generally [23,89–91,95–99], the accumulation process that we propose for the formation of conifer hotspots is likely to be widespread among organisms. The degree to which *in situ* diversification or preservation processes influence biodiversity should depend on the rate at which groups adapt [18]; for example, conifers and other long-lived woody plant groups have low rates of evolution but can show large range shifts over time, and thus may be more likely to form such accumulation hotspots. We suggest, however, that accumulation may be important in the formation of high richness regions generally and should be considered in addition to other processes such as high origination. The spatial variability in conifer diversity also highlights how areas of high species richness may be as much a function of geographical happenstance as of more intrinsic biological processes like diversification. Comprehensive species-level phylogenies combined with detailed trait and geographical range data will make it possible to better assess the extent and importance of accumulation in the generation of high richness and its role in structuring macroecological and biogeographic patterns.

**Data accessibility.** All data have been made available on data dryad: <http://dx.doi.org/10.5061/dryad.23p15n1>. All additional results provided in the electronic supplementary material.

**Authors' contributions.** M.S., A.B.L., M.J.D., W.J. and S.M. conceived the ideas and methodology; M.S., A.B.L., A.F. and D.F. collected the data; M.S. and A.B.L. statistically analysed the data; M.S. and A.B.L. led the writing of the manuscript.

**Competing interests.** We declare we have no competing interests.

**Funding.** We received no funding for this study.

**Acknowledgements.** We would like to thank two anonymous reviewers for their helpful comments.

## References

1. Pianka E. 1966 Latitudinal gradients in species diversity: a review of concepts. *Am. Nat.* **100**, 33–46. (doi:10.1086/282398)
2. Rosenzweig M. 1995 *Species diversity in space and time*. Cambridge, UK: Cambridge University Press.
3. Chown SL, Gaston KJ. 2000 Area, cradles and museums: the latitudinal gradient in species richness. *Trends Ecol. Evol.*



- 15, 311–315. (doi:10.1016/S0169-5347(00)01910-8)
4. Jablonski D, Kaustuv R, Valentine JW. 2006 Out of the tropics: evolutionary diversity gradient. *Science* **314**, 102–106. (doi:10.1126/science.1130880)
5. Orme CDL *et al.* 2005 Global hotspots of species richness are not congruent with endemism or threat. *Nature* **436**, 1016–1019. (doi:10.1038/nature03850)
6. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000 Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858. (doi:10.1038/35002501)
7. Mittermeier RA, Myers N, Thomsen JB, Fonseca GA, Olivieri S. 1998 Hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conserv. Biol.* **12**, 516–520. (doi:10.1046/j.1523-1739.1998.012003516.x)
8. Jetz W, Rahbek C. 2002 Geographic range size and determinants of avian species richness. *Science* **297**, 1548–1551. (doi:10.1126/science.1072779)
9. Wiens JJ, Donoghue MJ. 2004 Historical biogeography, ecology and species richness. *Trends Ecol. Evol.* **19**, 639–644. (doi:10.1016/j.tree.2004.09.011)
10. Harrison S, Noss R. 2017 Endemism hotspots are linked to stable climatic refugia. *Ann. Bot.* **119**, 207–214. (doi:10.1093/aob/mcw248)
11. McKenna D, Farrell B. 2006 Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. *Proc. Natl Acad. Sci. USA* **103**, 10 947–10 951. (doi:10.1073/pnas.0602712103)
12. Rolland J, Condamine FL, Jiguet F, Morlon H. 2014 Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biol.* **12**, e1001775. (doi:10.1371/journal.pbio.1001775)
13. Linder HP. 2008 Plant species radiations: where, when, why? *Phil. Trans. R. Soc. B* **363**, 3097–3105. (doi:10.1098/rstb.2008.0075)
14. Stebbins G. 1974 *Flowering plants: evolution above the species level*. London, UK: Arnold.
15. Médail F, Quézel P. 1999 Biodiversity hotspots in the Mediterranean basin: setting global conservation priorities. *Conserv. Biol.* **13**, 1510–1513. (doi:10.1046/j.1523-1739.1999.98467.x)
16. Raven PH, Axelrod DI. 1978 *Origin and relationships of the California flora*. Berkeley, CA: University of California Publications in Botany.
17. Stebbins GL, Major J. 1965 Endemism and speciation in the California flora. *Ecol. Monogr.* **35**, 1–35. (doi:10.2307/1942216)
18. Rangel TF, Edwards NR, Holden PB, Diniz-Filho JAF, Gosling WD, Coelho MTP, Cassemiro FAS, Rahbek C, Colwell RK. 2018 Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums, and graves. *Science* **361**, eaar5452. (doi:10.1126/science.aar5452)
19. Fine PVA. 2015 Ecological and evolutionary drivers of geographic variation in species diversity. *Annu. Rev. Ecol. Evol. Syst.* **46**, 369–392. (doi:10.1146/annurev-ecolsys-112414-054102)
20. Hopper SD. 2009 OCBIL theory: towards an integrated understanding of the evolution, ecology and conservation of biodiversity on old, climatically buffered, infertile landscapes. *Plant Soil* **322**, 49–86. (doi:10.1007/s11104-009-0068-0)
21. Hopper SD, Silveira FAO, Fiedler PL. 2016 Biodiversity hotspots and Ocbil theory. *Plant Soil* **403**, 167–216. (doi:10.1007/s11104-015-2764-2)
22. Ohsawa T, Ide Y. 2008 Global patterns of genetic variation in plant species along vertical and horizontal gradients on mountains. *Glob. Ecol. Biogeogr.* **17**, 152–163. (doi:10.1111/j.1466-8238.2007.00357.x)
23. Jetz W, Rahbek C, Colwell RK. 2004 The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecol. Lett.* **7**, 1180–1191. (doi:10.1111/j.1461-0248.2004.00678.x)
24. Farjon A. 2010 *A handbook of the world's conifers*. Leiden, The Netherlands: Brill.
25. Farjon A, Filer D. 2013 *An atlas of the world's conifers: an analysis of their distribution, biogeography, diversity and conservation status*. Leiden, The Netherlands: Brill.
26. Rai HS, Reeves PA, Peakall R, Olmstead RG, Graham SW. 2008 Inference of higher-order conifer relationships from a multi-locus plastid data set. *Botany* **86**, 658–669. (doi:10.1139/B08-062)
27. Leslie AB, Beaulieu JM, Rai HS, Crane PR, Donoghue MJ, Mathews S. 2012 Hemisphere-scale differences in conifer evolutionary dynamics. *Proc. Natl Acad. Sci. USA* **109**, 16 217–16 221. (doi:10.1073/pnas.1213621109)
28. Givnish TJ. 1980 Ecological constraints on the evolution of breeding systems in seed plants: dioecy and dispersal in gymnosperms. *Evolution* **34**, 959–972. (doi:10.2307/2408001)
29. Leslie AB, Beaulieu JM, Crane PR, Donoghue MJ. 2013 Explaining the distribution of breeding and dispersal syndromes in conifers explaining the distribution of breeding and dispersal syndromes in conifers. *Proc. R. Soc. B* **280**, 20131812. (doi:10.1098/rspb.2013.1812)
30. Pateiro-López B, Rodríguez-Casal A. 2010 Generalizing the convex hull of a sample: the R package alphahull. *J. Stat. Softw.* **34**, 1–28. (doi:10.18637/jss.v034.i05)
31. Gaston KJ, Fuller RA. 2009 The sizes of species' geographic ranges. *J. Appl. Ecol.* **46**, 1–9. (doi:10.1111/j.1365-2664.2008.01596.x)
32. IUCN Standards and Petitions Subcommittee. 2017 Guidelines for using the IUCN Red List categories and criteria. Version 13. See <http://www.iucnredlist.org/documents/RedListGuidelines.pdf>.
33. Getis A, Ord JK. 1992 The analysis of spatial association by use of distance statistics. *Geogr. Anal.* **24**, 189–206. (doi:10.1111/j.1538-4632.1992.tb00261.x)
34. Ord JK, Getis A. 1995 Local spatial autocorrelation statistics: distributional issues and an application. *Geogr. Anal.* **27**, 286–306. (doi:10.1111/j.1538-4632.1995.tb00912.x)
35. Caldas de Castro M, Singer BH. 2006 Controlling the false discovery rate: a new application to account for multiple and dependent test in local statistics of spatial association. *Geogr. Anal.* **38**, 180–208. (doi:10.1111/j.0016-7363.2006.00682.x)
36. Baselga A, Orme CDL. 2012 Betapart: an R package for the study of beta diversity. *Methods Ecol. Evol.* **3**, 808–812. (doi:10.1111/j.2041-210X.2012.00224.x)
37. Baselga A. 2012 The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Glob. Ecol. Biogeogr.* **21**, 1223–1232. (doi:10.1111/j.1466-8238.2011.00756.x)
38. Leprieux F, Albouy C, de Bortoli J, Cowman PF, Bellwood DR, Mouillot D. 2012 Quantifying phylogenetic beta diversity: distinguishing between 'true' turnover of lineages and phylogenetic diversity gradients. *PLoS One* **7**, e42760. (doi:10.1371/journal.pone.0042760)
39. Kreft H, Jetz W. 2010 A framework for delineating biogeographical regions based on species distributions. *J. Biogeogr.* **37**, 2029–2053. (doi:10.1111/j.1365-2699.2010.02375.x)
40. Borcard D, Gillet F, Legendre P. 2018 *Numerical ecology with R*. New York, NY: Springer.
41. Fick SE, Hijmans RJ. 2017 Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315. (doi:10.1002/joc.5086)
42. Hartmann J, Moosdorf N. 2012 The new global lithological map database GLiM: a representation of rock properties at the Earth surface. *Geochem. Geophys. Geosyst.* **13**, 1–37. (doi:10.1029/2012GC004370)
43. Post WM, Zobler L. 2000 Data from: Global soil types, 0.5-degree grid (Modified Zobler). Oak Ridge, TN: ORNL DAAC. See <https://doi.org/10.3334/ORNLDAAC/540>.
44. Aerts R. 1995 The advantages of being evergreen. *Trends Ecol. Evol.* **10**, 402–407. (doi:10.2307/2934676)
45. Cressie NAC. 1981 *Statistics for spatial data*. New York, NY: Wiley.
46. Bivand R *et al.* 2006 Spatial dependence: weighing schemes, statistics and models. R package 'spdep' version 0.5-92.
47. Cliff AD, Ord JK. 1981 *Spatial processes*. London, UK: Pion.
48. Fortin MJ, Dale MRT. 2005 *Spatial analysis: a guide for ecologists*. New York, NY: Cambridge University Press.
49. Schmitt S, Pouteau R, Josteau D, de Boissieu F, Birnbaum P. 2017 ssdm: an R package to predict distribution of species richness and composition based on stacked species distribution models. *Methods Ecol. Evol.* **8**, 1795–1803. (doi:10.1111/2041-210X.12841)
50. Leishman MR, Wright IJ, Moles AT, Westoby M. 2000 The evolutionary ecology of seed size. In *Seeds: the ecology of regeneration in plant communities* (ed. M Fenner), pp. 31–57. Wallingford, UK: CAB International.
51. Moles AT, Westoby M. 2004 Seedling survival and seed size: a synthesis of the literature. *J. Ecol.* **92**, 372–383. (doi:10.1111/j.0022-0477.2004.00884.x)
52. Westoby M, Leishman M, Lord J. 1996 Comparative ecology of seed size and dispersal. *Phil. Trans. R. Soc. B* **351**, 1309–1318. (doi:10.1098/rstb.1996.0114)
53. Westoby M, Falster DS, Moles AT, Vesk PA, Wright IJ. 2002 Plant ecological strategies: some leading

- dimensions of variation between species. *Annu. Rev. Ecol. Syst.* **33**, 125–159. (doi:10.1146/annurev.ecolsys.33.010802.150452)
54. Reeves G, Francis D, Davies MS, Rodgers H, Hodkinson T. 1998 Genome size is negatively correlated with altitude in natural populations of *Dactylis glomerata*. *Ann. Bot.* **82**, 99–105. (doi:10.1006/anbo.1998.0751)
55. Vinogradov AE. 2003 Selfish DNA is maladaptive: evidence from the plant Red List. *Trends Genet.* **19**, 609–614. (doi:10.1016/j.tig.2003.09.010)
56. Knight CA, Molinari NA, Petrov DA. 2005 The large genome constraint hypothesis: evolution, ecology and phenotype. *Ann. Bot.* **95**, 177–190. (doi:10.1093/aob/mci011)
57. Procheş Ş, Wilson JRU, Richardson DM, Rejmánek M. 2012 Native and naturalized range size in *Pinus*: relative importance of biogeography, introduction effort and species traits. *Glob. Ecol. Biogeogr.* **21**, 513–523. (doi:10.1111/j.1466-8238.2011.00703.x)
58. Bainard JD, Bainard LD, Henry TA, Fazekas AJ, Newmaster SG. 2012 A multivariate analysis of variation in genome size and endoreduplication in angiosperms reveals strong phylogenetic signal and association with phenotypic traits. *New Phytol.* **196**, 1240–1250. (doi:10.1111/j.1469-8137.2012.04370.x)
59. Murray B, Leitch I, Bennett M. 2012 Data from: gymnosperm DNA C-value database. Kew Royal Botanic Gardens. See <http://data.kew.org/cvalues/CvalServlet?querytype=4>.
60. Ho LST, Ané C. 2014 A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* **63**, 397–408. (doi:10.1093/sysbio/syu005)
61. Ives AR, Garland T. 2010 Phylogenetic logistic regression for binary dependent variables. *Syst. Biol.* **59**, 9–26. (doi:10.1093/sysbio/syp074)
62. Revell LJ. 2012 phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223. (doi:10.1111/j.2041-210X.2011.00169.x)
63. Tucker CM *et al.* 2017 A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biol. Rev.* **92**, 698–715. (doi:10.1111/brv.12252)
64. Isaac NJB, Turvey ST, Collen B, Waterman C, Baillie JEM. 2007 Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS ONE* **2**, e296. (doi:10.1371/journal.pone.0000296)
65. Redding DW, Hartmann K, Mimoto A, Bokal D, DeVos M, Mooers A. 2008 Evolutionarily distinctive species often capture more phylogenetic diversity than expected. *J. Theor. Biol.* **251**, 606–615. (doi:10.1016/j.jtbi.2007.12.006)
66. Jetz W, Thomas GH, Joy JB, Hartmann K, Redding D, Mooers AO. 2014 Distribution and conservation of global evolutionary distinctness in birds. *Curr. Biol.* **24**, 1–12. (doi:10.1016/j.cub.2014.03.011)
67. Leslie AB, Beaulieu JM, Holman G, Campbell CS, Mei W, Raubeson LR, Mathews S. 2018 An overview of extant conifer evolution from the perspective of the fossil record. *Am. J. Bot.* **105**, 1531–1544. (doi:10.1002/ajb2.1143)
68. Laffan SW, Lubarsky E, Rosauer DF. 2010 Biodiverse, a tool for the spatial analysis of biological and related diversity. *Ecography* **33**, 643–647. (doi:10.1111/j.1600-0587.2010.06237.x)
69. Jetz W, Fine PVA. 2012 Global gradients in vertebrate diversity predicted by historical area-productivity dynamics and contemporary environment. *PLoS Biol.* **10**, e1001292. (doi:10.1371/journal.pbio.1001292)
70. Jansson R. 2003 Global patterns in endemism explained by past climatic change. *Proc. R. Soc. B* **270**, 583–590. (doi:10.1098/rspb.2002.2283)
71. Ricklefs RE. 2008 Disintegration of the ecological community. *Am. Nat.* **172**, 741–750. (doi:10.1086/593002)
72. Ricklefs RE. 1987 Community diversity: relative roles of local and regional processes. *Science* **235**, 167–171. (doi:10.1126/science.235.4785.167)
73. Zhao M, Heinsch FA, Nemani RR, Running SW. 2005 Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sens. Environ.* **95**, 164–176. (doi:10.1016/j.rse.2004.12.011)
74. Zhao M, Running SW. 2010 Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* **329**, 940–943. (doi:10.1126/science.1192666)
75. Ehlers J, Ehlers J, Gibbard PL, Hughes PD. 2011 *Quaternary glaciations: extent and chronology: a closer look*. Amsterdam, The Netherlands: Elsevier.
76. Moreau CS, Bell CD. 2013 Testing the museum versus cradle tropical biological diversity hypothesis: phylogeny, diversification, and ancestral biogeographic range evolution of the ants. *Evolution* **67**, 2240–2257. (doi:10.1111/evo.12105)
77. Davis M, Shaw R. 2001 Range shifts and adaptive responses to quaternary climate change. *Science* **292**, 673–679. (doi:10.1126/science.292.5517.673)
78. Williams JW, Shuman BN, Webb T, Bartlein PJ, Leduc PL. 2004 Late-quaternary vegetation dynamics in North America: scaling from taxa to biomes. *Ecol. Monogr.* **74**, 309–334. (doi:10.1890/02-4045)
79. Svenning JC, Skov F. 2007 Ice age legacies in the geographical distribution of tree species richness in Europe. *Glob. Ecol. Biogeogr.* **16**, 234–245. (doi:10.1111/j.1466-822x.2006.00280.x)
80. Yeaman S *et al.* 2016 Convergent local adaptation to climate in distantly related conifers. *Science* **353**, 23–26. (doi:10.1126/science.aaf7812)
81. Eiga S, Sakai A. 1984 Altitudinal variation in freezing resistance of Saghalien fir (*Abies sachalinensis*). *Can. J. Bot.* **62**, 156–160. (doi:10.1139/b84-025)
82. Janzen DH. 1967 Why mountain passes are higher in the tropics. *Am. Nat.* **101**, 233–249. (doi:10.1086/282487)
83. Sandel B, Arge L, Dalsgaard B, Davies RG, Gaston KJ, Sutherland WJ, Svenning J-C. 2011 The influence of late quaternary climate change velocity on species endemism. *Science* **334**, 660–664. (doi:10.1126/science.1210173)
84. Ma Z, Sandel B, Svenning JC. 2016 Phylogenetic assemblage structure of North American trees is more strongly shaped by glacial-interglacial climate variability in gymnosperms than in angiosperms. *Ecol. Evol.* **6**, 3092–3106. (doi:10.1002/ece3.2100)
85. Sandel B, Monnet AC, Govaerts R, Vorontsova M. 2016 Late quaternary climate stability and the origins and future of global grass endemism. *Ann. Bot.* **119**, 279–288. (doi:10.1093/aob/mcw178)
86. Dynesius M, Jansson R. 2000 Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc. Natl Acad. Sci. USA* **97**, 9115–9120. (doi:10.1073/pnas.97.16.9115)
87. Jansson R, Dynesius M. 2002 The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. *Annu. Rev. Ecol. Syst.* **33**, 741–777. (doi:10.1146/annurev.ecolsys.33.010802.150520)
88. Schluter D, Pennell MW. 2017 Speciation gradients and the distribution of biodiversity. *Nature* **546**, 48–55. (doi:10.1038/nature22897)
89. Braun EL. 1955 The phytogeography of unglaciated eastern United States and its interpretation. *Bot. Rev.* **21**, 297–375. (doi:10.1007/BF02872433)
90. Whittaker R. 1961 Vegetation history of the Pacific coast states and the 'central' significance of the Klamath region. *Madroño* **16**, 5–23.
91. Gavin D, Heath KD, Mcglone M, Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD. 2014 Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytol.* **204**, 37–54. (doi:10.1111/nph.12929)
92. López-Pujol J, Zhang FM, Sun HQ, Ying TS, Ge S. 2011 Centres of plant endemism in China: places for survival or for speciation? *J. Biogeogr.* **38**, 1267–1280. (doi:10.1111/j.1365-2699.2011.02504.x)
93. Tang CQ. 2015 *The subtropical vegetation of southwestern China: plant distribution, diversity and ecology*, pp. 185–274. Heidelberg, Germany: Springer.
94. Farjon A. 1996 Biodiversity of *Pinus* (Pinaceae) in Mexico: speciation and palaeo-endemism. *Bot. J. Linn. Soc.* **121**, 365–384. (doi:10.1111/j.1095-8339.1996.tb00762.x)
95. Pryke JS, Samways MJ. 2008 Conservation of invertebrate biodiversity on a mountain in a global biodiversity hotspot, Cape Floral Region. *Biodivers. Conserv.* **17**, 3027–3043. (doi:10.1007/s10531-008-9414-4)
96. Fjeldsa J, Lovett JC. 1997 Geographical patterns of old and young species in African forest. *Biodivers. Conserv.* **6**, 325–346. (doi:10.1023/A:1018356506390)
97. Körner C. 2004 Mountain biodiversity, its causes and function. *Ambio* **13**, 11–17. (doi:10.2307/25094582)
98. Araújo MB. 2002 Biodiversity hotspots and zones of ecological transition. *Conserv. Biol.* **16**, 1662–1663. (doi:10.1046/j.1523-1739.2002.02068.x)
99. Steinbauer MJ, Irl SDH, Beierkuhnlein C. 2013 Elevation-driven ecological isolation promotes diversification on Mediterranean islands. *Acta Oecol.* **47**, 52–56. (doi:10.1016/j.actao.2012.11.004)