

Unravelling the evolutionary origins of biogeographic assemblages

Barnabas H. Daru^{1,2,*}, Michelle van der Bank³, and T. Jonathan Davies^{3,4}

¹Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

²Department of Plant Sciences, University of Pretoria, Private Bag X20, Hatfield 0028, Pretoria, South Africa

³African Centre for DNA Barcoding, University of Johannesburg, APK Campus, PO Box 524, Auckland Park 2006, Johannesburg, South Africa

⁴Department of Biology, McGill University, 1205, Avenue Docteur Penfield, Montreal, Quebec, Canada

*Correspondence: Barnabas H. Daru, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

Email: barnabas_daru@fas.harvard.edu

RUNNING TITLE: Origin of modern biomes

ARTICLE TYPE: Biodiversity Research

ABSTRACT

Aim Floristic and faunal diversity fall within species assemblages that can be grouped into distinct biomes or ecoregions. Understanding the origins of such biogeographic assemblages helps illuminate the processes shaping present-day diversity patterns, and identifies regions with unique or distinct histories. While the fossil record is often sparse, dated phylogenies can provide a window into the evolutionary past of these regions. Here, we present a novel phylogenetic approach to investigate the evolutionary origins of present-day biogeographic assemblages, and highlight their conservation value.

Location Southern Africa.

Methods We evaluate the evolutionary turnover separating species clusters in space at different time slices to determine the phylogenetic depth at which the signal for their present-day structure emerges. We suggest present-day assemblages with distinct evolutionary histories might represent important units for conservation. We apply our method to the vegetation of southern Africa using a dated phylogeny of the woody flora of the region, and explore how the evolutionary history of vegetation types compares to common conservation currencies, including species richness, endemism and threat.

Results We show the differentiation of most present-day vegetation types can be traced back to evolutionary splits in the Miocene. The woody flora of the Fynbos is the most evolutionarily distinct, and thus has deeper evolutionary roots, whereas the Savanna and Miombo Woodland show close phylogenetic affinities and likely represent a more recent separation. However, evolutionarily distinct phyloregions do not necessarily capture the most unique phylogenetic diversity, nor are they the most species-rich or threatened.

Main conclusions Our approach complements analyses of the fossil record, and serves as a link to the history of diversification, migration and extinction of lineages within biogeographic assemblages that is separate from patterns of species richness and endemism. Our analysis reveals how phyloregions capture conservation value not represented by traditional biodiversity metrics.

Keywords: Miocene, phylogenetic beta diversity, phyloregions, phytogeographic regions, southern Africa

INTRODUCTION

Biodiversity is under increasing anthropogenic pressure, and there is an urgent need to prioritize the protection of areas with high conservation value. Systematic conservation planning aims to optimize the use of limited resources for the conservation of biodiversity (Margules & Pressey, 2000; James *et al.*, 2001), with a common approach being the prioritization of biodiversity hotspots (*i.e.*, geographic clusters of high conservation value; Myers, 1988; Myers *et al.*, 2000). Hotspots may represent locations of important ecological and evolutionary processes, and they have proven useful for guiding effective allocation of limited conservation funds (Mittermeier *et al.*, 2000; Myers *et al.*, 2000). However, conservation efforts based on hotspots alone do not necessarily guarantee optimal conservation solutions. Hotspots are typically based upon only one or a few axes of diversity, most usually species richness, endemism and threat, and might not, therefore, capture other dimensions of biodiversity, such as phylogenetic diversity or latent risk (see Forest *et al.*, 2007; Davies & Cadotte, 2011; Daru *et al.*, 2015). In addition, conservation planners might wish to strive for representation of habitats, vegetation types or ecological communities as well as species within a protected area network (Margules & Pressey, 2000). Phylogenetic regionalization offers a relatively new approach for delineating landscapes or seascapes into geographic units that captures information on the evolutionary structure of the species assemblages within them (Holt *et al.*, 2013; Daru *et al.*, 2016; Daru *et al.*, 2017a), and provides additional information for conservation planners beyond traditional hotspot approaches.

The association of species into distinct phylogenetically delimited biogeographic units (referred to as phyloregions by Daru *et al.*, 2016) captures historical processes such as diversification, niche conservatism, dispersal and extinction, that may have operated over several

millions of years (Davies & Buckley, 2012; Guo *et al.*, 2012; Wu *et al.*, 2016; Daru *et al.*, 2017b). For instance, dispersal limitation will tend to increase phylogenetic turnover among regions and, as a consequence, the evolutionary separation of phyloregions; whereas the phylogenetic composition of areas encompassing species with high dispersal abilities might be more homogeneous spatially. Dispersal limitation is one of several mechanisms that may be linked to the more general processes of niche conservatism – the tendency of species to retain their ancestral niche. The more phylogenetically conserved the niche, the greater the predicted phylogenetic structure in species assemblages. Speciation and extinction can also leave an imprint on the structure of phyloregions. Endemic radiations can increase the evolutionary separation of phyloregions by increasing clustering of closely related species, while vicariance speciation due to geographic isolation, could reduce the evolutionary distinctiveness of phyloregions because sibling species may occur in neighbouring phyloregions. Extinction can structure phyloregions through its effect on both taxon richness and composition. For example, local extinction of old lineages (paleoendemics) might decrease the phylogenetic distance separating phyloregions. However, the loss of evolutionarily derived lineages (neoendemics) will have less impact on the phylogenetic structure of a region, at least initially (Daru *et al.*, 2017b).

The species composition of present-day assemblages can thus provide insights into the mechanistic processes that have shaped them (Pennington *et al.*, 2004), and more evolutionarily distinct phyloregions might deserve greater conservation priority because they may encompass co-occurring lineages with longer ecological histories. However, to fully understand their present day structure we must look backwards into their evolutionary past. Davies & Buckley (2011) examined the diversity of lineages present at different time slices through the phylogenetic tree

to gain insight into the evolutionary past of present-day diversity of mammals. This framework opened a window into the historical processes by which present-day mammalian diversity arose. We present a related framework to explore the evolutionary origin of present-day biogeographic assemblages by successively removing phylogenetic structure in the tree linking regional species assemblages to identify the evolutionary depth at which they collapse into each other (Fig. 1). We suggest that the phylogenetic depth at which biogeographic regions become indistinguishable may help inform estimates of the time at which they differentiated historically.

We illustrate our framework using a dated molecular phylogeny of woody species to investigate the evolutionary history of present-day vegetation types in southern Africa, an area covering Botswana, Lesotho, Mozambique south of the Zambezi river, Namibia, South Africa, Swaziland, and Zimbabwe. On the African continent, southern Africa includes some of the most floristically diverse regions, and is characterized by varied vegetation types including Savanna, Miombo Woodland (a type of woody savanna), Grassland, Forest, Desert, and Fynbos (Olson *et al.*, 2001; Mucina & Rutherford, 2006), which differ greatly in floristic composition. We show that some phyloregions have long evolutionary histories, whereas others represent relatively recent separations. We suggest that these patterns reflect the historical assembly of southern African vegetation into present-day biomes. We then contrast the evolutionary distinctiveness of phyloregions to alternative conservation indices and show how different indices capture distinct dimensions of diversity.

METHODS

Taxon sampling and phylogenetic reconstruction

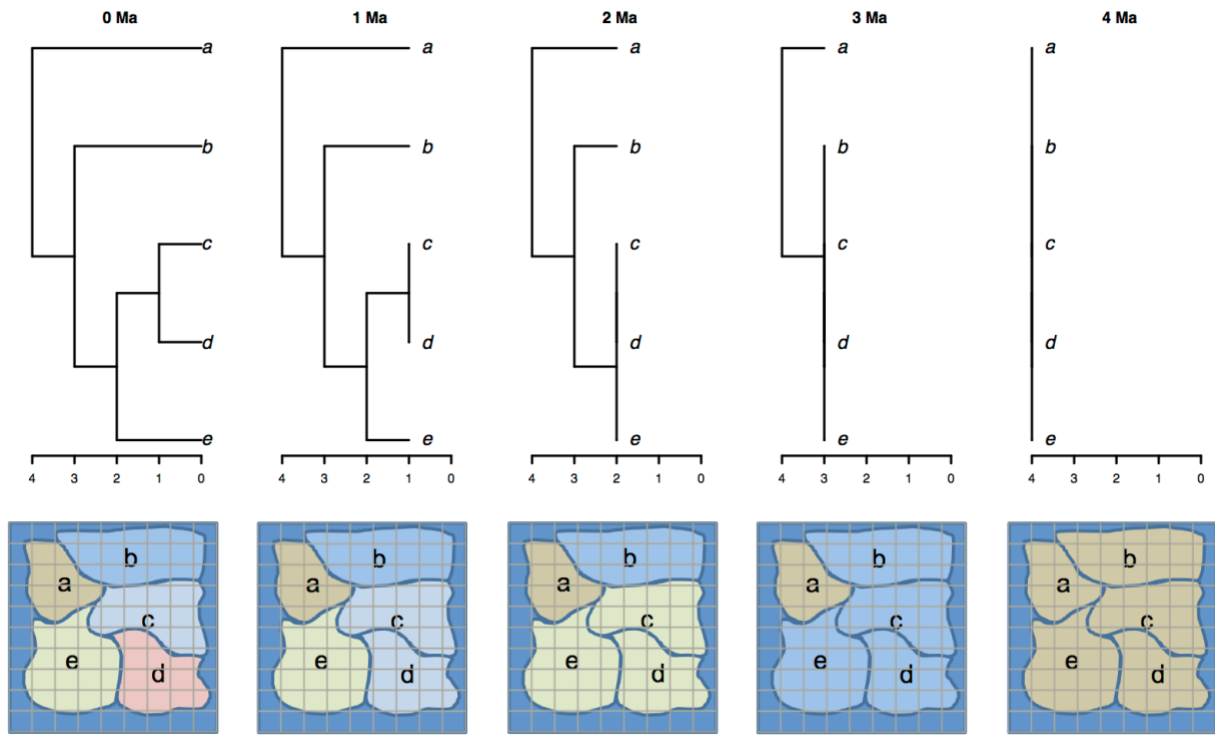


Fig. 1. Cartoon demonstrating the successive removal of phylogenetic structure in a tree for five species *a-e* in five separate phylogeographic regions, assuming each phylogeographic region is represented by one tip on the phylogeny. As the tree is deconstructed at 1-Myr intervals from 0 to 4 Ma, information on more recent branching patterns is sequentially lost, and only deeper splits in the tree are informative with respect to biogeographic groupings. For empirical data, phylogeographic regions can be defined by multiple tips in the phylogeny.

Following Daru *et al.* (2016), we used a dated molecular phylogenetic tree for the woody flora of southern Africa from Maurin *et al.* (2014) to generate a classification of phyloregions representing the major vegetation types in southern Africa. Our definition of woody species follows previous studies (*e.g.*, O'Brien, 1993; FitzJohn *et al.*, 2014), as plant species with above-ground stems >0.5 m, including species that produce secondary xylem tissues such as some tree *Aloe* species, palms, and bamboo. The total woody flora of southern Africa is estimated at ~2200 species within 117 families representing 543 genera of angiosperms and gymnosperms (Palgrave, 2002; Schmidt *et al.*, 2007; Boon, 2010; Van Wyk *et al.*, 2011), from which our analysis samples 1400 species within 115 families representing 541 genera; sampling is thus almost complete for higher taxonomic levels (see Table S1 in Supporting Information). We used extent-of-occurrence maps for all 1400 species from the literature (Palgrave, 2002; van Wyk *et al.*, 2011), and extracted species presences/absences within a matrix of 50 km × 50 km grid cells (Behrmann equal-area cylindrical coordinate system) as described in Daru *et al.* (2015).

Full details of the phylogeny reconstruction are provided in Maurin *et al.* (2014) and Daru *et al.* (2016). The phylogenetic tree used here was estimated using Bayesian analysis of 1400 species and 1633 bp of chloroplast DNA sequences derived from a combination of *matK* and *rbcLa*, assuming an uncorrelated relaxed molecular clock model, using the program BEAST v.1.7.5 (Drummond & Rambaut, 2007). Branch lengths were calibrated in millions of years using a Bayesian MCMC approach by enforcing topological constraints assuming APG III backbone from Phylomatic v.3 (Webb & Donoghue, 2005) and 18 fossil calibration points from Bell *et al.* (2010). A table with GenBank accession numbers is provided in the supporting information to Daru *et al.* (2016).

Construction of Phyloregions

Phyloregions were generated by combining the presence or absence of each species distributions within grid cells with branch length information from a dated phylogenetic tree to generate a pairwise distance matrix of phylogenetic beta diversity ($p\beta_{sim}$; Holt *et al.*, 2013) between all pairs of grid cells using R (R Core Team, 2015). We tested the performance of eight clustering algorithms for the regionalization purpose including single linkage, complete linkage, unweighted pair-group method using arithmetic averages (UPGMA), unweighted pair-group method using centroids, weighted pair-group method using arithmetic averages, weighted pair-group method using centroids, Ward's minimum variance and DIANA's divisive hierarchical method. The best performing algorithm was identified using cophenetic correlation coefficient (Sokal & Rohlf, 1962) and Gower's distance (Gower, 1983).

In a second step, we used the 'elbow' method of Salvador & Chan (2004) to 'cut' the dendrogram and determined the optimal number of meaningful clusters (phyloregions) using the "elbow" function in the R package GMD (Zhao *et al.*, 2011). Last, we described the relationships among the resulting phylogenetically-delimited clusters ("phyloregions"; Daru *et al.*, 2016) using hierarchical dendrogram of dissimilarity and non-metric multidimensional scaling (NMDS) ordination, which represents the original distance matrix in a few (usually 2 or 3) dimensions with a minimum loss of information. A full description of how phyloregions were delimited is included in Daru *et al.* (2016). For convenience, when present day phyloregions correspond spatially to a recognized vegetation type, we refer to the phyloregion by this domain name (Daru *et al.*, 2016); however, phyloregions do not represent a one-to-one match to the named vegetation type.

Reconstructing the evolutionary history of phyloregions

First, we explored the phylogenetic signal in the distribution of species within phyloregions using a binary presence-absence matrix for each phyloregion. Phylogenetic signal was evaluated using the D-statistic from Fritz & Purvis (2010) implemented in the R package Caper (Orme *et al.*, 2012). The D statistic estimates phylogenetic conservatism by scaling the sum of sister-clade differences with those expected under a random model versus a Brownian motion model, and significance is assessed by shuffling the trait values (1000 times) at the tips of the phylogeny. A D-value of 1 indicates a random distribution of traits at the tip of the phylogeny, whereas a D-value of 0 indicates a Brownian motion model (Fritz & Purvis, 2010).

The deeper branches of a phylogenetic tree (those towards the root of the tree) may capture the signature of past environments that might be different to their present-day descendants (Rosauer *et al.* 2014). To explore the phylogenetic depth at which the spatial signature of present-day biogeographic assemblages emerges, we successively removed phylogenetic structure in the tree from the tips to the root (*i.e.*, collapse younger phylogenetic branches into older ones) and iterated the clustering algorithm at each time slice (see Fig. 1). R-code for manipulating the phylogenetic tree and clustering using phylogenetic beta diversity are provided as supplemental information Data S1 and S2.

At each phylogenetic depth, we compare our newly generated phyloregions to the original clusters estimated using the fully resolved phylogeny, and to the geographical distribution of currently recognized phytogeographic regions using a map of the vegetation zones of southern Africa (White, 1983 with modifications and additions from Low & Rebelo, 1996; Olson *et al.*, 2001; Burgess *et al.*, 2004; Mucina & Rutherford, 2006). We examined the

phylogenetic depth of phyloregions at two temporal scales. First, we considered relatively coarse evolutionary intervals spanning from 260 Ma to present-day (0 Ma) in 20-million year (myr) time slices, and sequentially collapsed nodes and truncated branch lengths, such that the final phylogenetic topology had a total depth of zero. Second, we focus at a finer phylogenetic depth between 20 and 0 Ma (present-day), slicing the tree in 4-myr intervals. The latter time period corresponds to the Miocene-Pliocene (~20 Ma to 3 Ma), which is hypothesized to represent a period of major biome reorganization in Earth's history (Osborne & Beerling, 2006; Kürschner *et al.*, 2008; Edwards *et al.*, 2010).

Floristic relationships among phyloregions are represented using non-metric multidimensional scaling (NMDS) and a hierarchical dendrogram of dissimilarity.

Dimensions of biodiversity

To evaluate how phyloregions captured different dimensions of diversity, we mapped five commonly used indices of biodiversity onto each phyloregion: species richness (SR), species endemism richness (SE), phylogenetic diversity endemism (PD-endemism), evolutionary distinctiveness (ED), and ‘evolutionary distinctiveness and global endangerment’ (EDGE). SR is the total count of species in each phyloregion, SE was calculated as the richness of endemic species per phyloregion, PD-endemism measures the degree to which phylogenetic diversity is restricted to a phyloregion (Faith *et al.*, 2004); and ED is the amount of unique evolutionary history represented by a species with few or no extant relatives and calculated based on the fair proportion metric (Isaac *et al.*, 2007), and can be multiplied by each species global endangerment (GE) from IUCN conservation categories to form EDGE scores (species that are

phylogenetically isolated and threatened with extinction). ED and EDGE were represented as the mean of the species values within each phyloregion. We tested the performance of phyloregions as a conservation decision making tool by overlaying the top richest grid cells (*i.e.*, hotspot cells) of five types of diversity metrics (SE, EDGE, ED, PD-endemism and SR) on the map of phyloregions to identify phyloregions that do not overlap any hotspots.

Phyloregions at risk

Last, we examined the potential evolutionary future of present-day phyloregions by removing all currently threatened species (IUCN conservation categories VU, EN and CR), recalculating phylogenetic beta diversity metrics for the remaining extant species, and then reconstructing a new set of phyloregions for this reduced taxon set following the methods described above.

RESULTS

Membership of a particular phyloregion appears to be a relatively labile trait, and most phyloregions sample from clades across the phylogenetic tree (Fig. 2). Although lineages show significant phylogenetic conservatism in phytogeographic affinity, phylogenetic signal is generally weak. However, there is some evidence for stronger phylogenetic signal in the Fynbos phyloregion and weaker signal in the Savanna and Grassland phyloregions ($D_{\text{Fynbos}} = 0.4$, $D_{\text{Savanna}} = 0.7$, $D_{\text{Grassland}} = 0.8$; all $P < 0.001$).

Phyloregions corresponding closely to present-day vegetation start to emerge with the inclusion of more recent evolutionary splits within the last ~60 myr. Between 40-20 Ma the grassland phyloregion, encompassing present-day Grassland, Nama Karoo, Indian Ocean Coastal

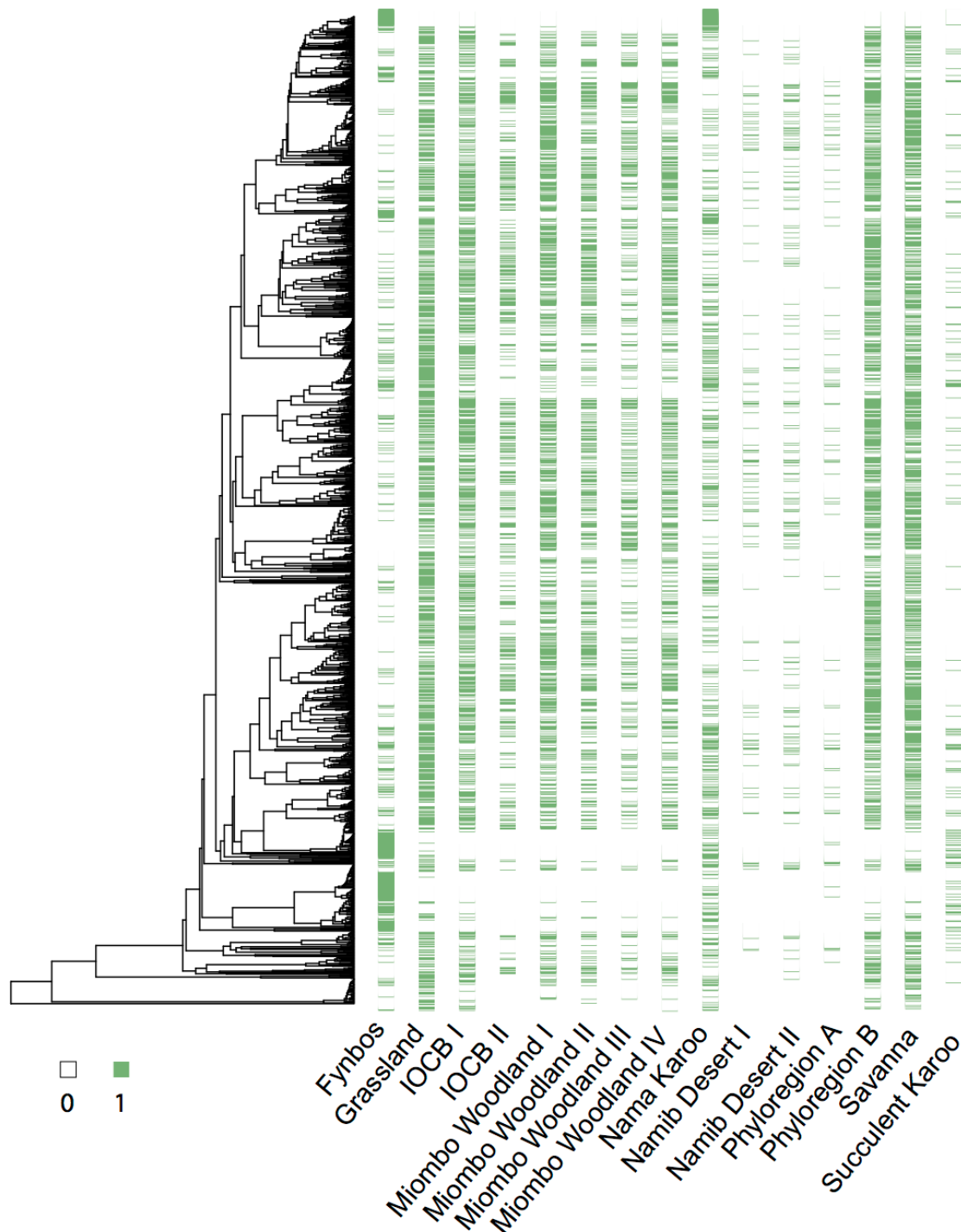


Fig. 2. Distribution of lineages among present-day phyloregions of southern Africa. The presence or absence of a species within a phyloregion is indicated by the color bars. The tree is a maximum clade credibility tree obtained from Bayesian analysis of combined of *matk* and *rbcLa* DNA regions of woody species of southern Africa, and dated using 18 secondary calibration points (see Maurin *et al.*, 2014).

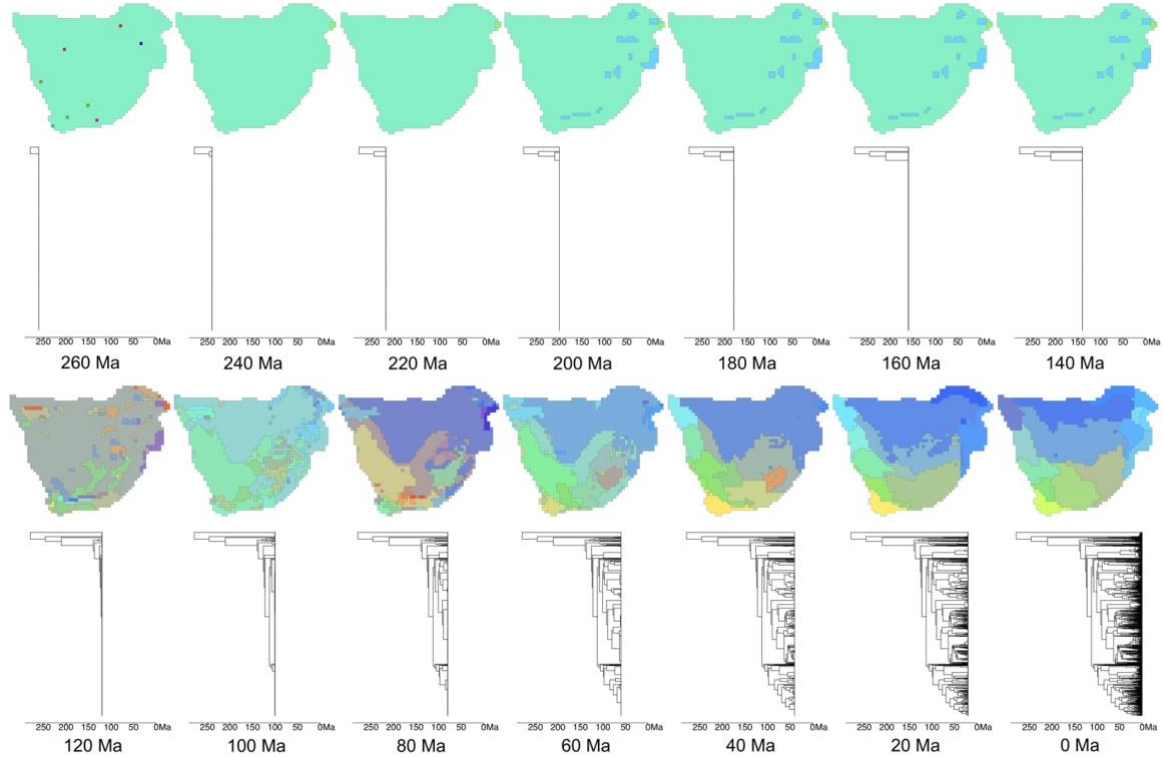


Fig. 3. Temporal window into the evolutionary past of the present-day phyloregions of southern Africa at successive phylogenetic depths over the last 260 myr. The phylogenetic tree is cut at 20-myrr intervals (below), and the corresponding geographic clustering is illustrated above. The map is plotted in Berhmann (equal-area) projection using 50×50 km grid cells. The time slice at which phyloregions become indistinguishable may be indicative of the time at which they differentiated historically. The spatial representation of aggregated phyloregions as we step back in time should not be interpreted as representative of their true historical distribution as past climates likely differed from the present day, but is simply provided to illustrate spatially the phylogenetic linkages between present-day phyloregions. Colors are in NMDS space, and indicate levels of differentiation of the flora in different phyloregions. Phyloregions with similar colors have similar clades and those with different colors differ in the plant clades they enclose.

Belt and Albany thickets, becomes distinct (Fig. 3), but it is only when we include evolutionary splits within the last 20 myr that we are able to fully differentiate among modern vegetation types.

To better characterize the emergence of vegetation types within the last 20 Ma, we explored this time window using finer phylogenetic slices at 4-myr intervals. We found that the Miombo Woodland phyloregion can be differentiated from the Savanna phyloregion at a phylogenetic depth of ~12 Ma, indicating that present-day Savanna may have arisen during this period. The Grassland and Nama Karoo phyloregions, which encompass parts of the Indian Ocean Coastal Belt and Albany thickets differentiate ~8 Ma. It was also during this time period (~8 Ma) that we observe the emergence of a formerly unclassified vegetation type, which we previously referred to as “Gariiep Karoo” (see Daru *et al.*, 2016), located between the Succulent Karoo and the Namib Desert. Last, between ~4 Ma to the present, a second novel phyloregion, which we refer to as the “Zambebian transition zone” (Daru *et al.*, 2016), emerges at the border between South Africa, Zimbabwe and Mozambique (Fig. 4). This is the youngest phyloregion (*i.e.*, is only differentiated by recent phylogenetic splits), and shows close phylogenetic affinity to the flora of the Indian Ocean Coastal Belt (Fig. 4).

The order of differentiation of phyloregions is supported by the NMDS ordination and hierarchical dendrogram of dissimilarity (Fig. S1 in Appendix S1). As we step back through time, adjacent phyloregions collapse into each other sequentially. Notably, around 20 Ma the NMDS ordination plots support a major split separating a northern group encompassing Savanna, Miombo Woodland, Indian Ocean Coastal Belt, and Desert, and a southern group encompassing Fynbos, Nama Karoo, Grassland, and Albany thickets (Fig. S1 in Appendix S1).

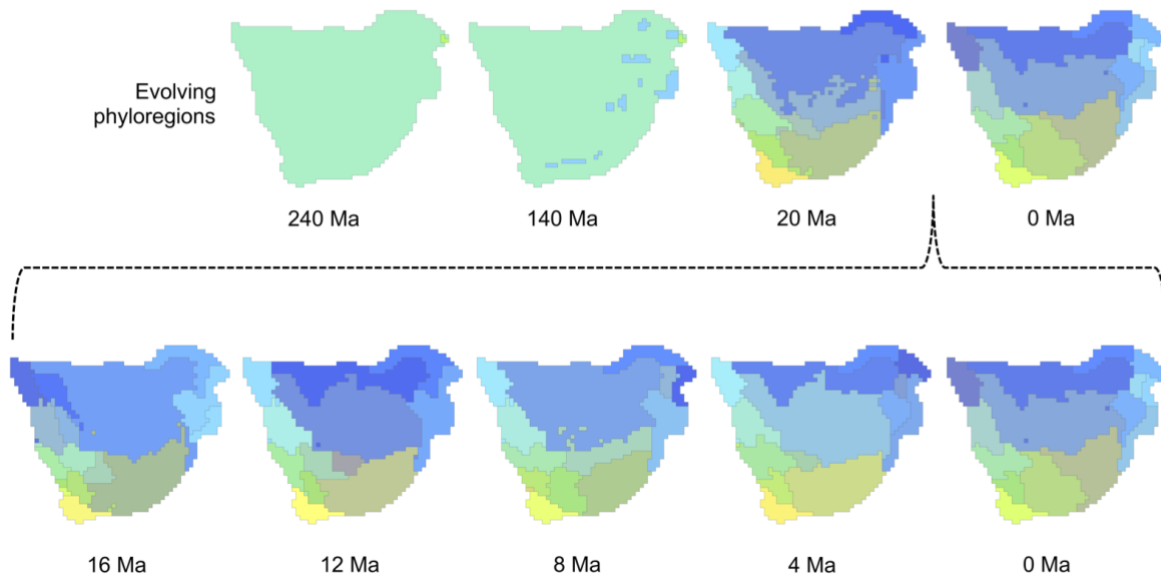


Fig. 4. Equivalent representation of the evolution of modern phyloregions to that shown in Fig. 3, but focusing on 20 Ma to present. Five maps resulting from the phylogenetic beta diversity ($p\beta_{sim}$) values for species assemblages in 50×50 km grid cells, cut at phylogenetic depths of 16, 12, 8, 4 and 0 Ma are illustrated. Colors are in NMDS space, and indicate levels of differentiation of the flora in different phyloregions. Phyloregions with similar colors have similar clades and those with different colors differ in the plant clades they enclose. The color gradient in the shading of the maps depicts northern regions in darker colors and southern regions in lighter colors. The map is plotted using Berhmann (equal-area) projection.

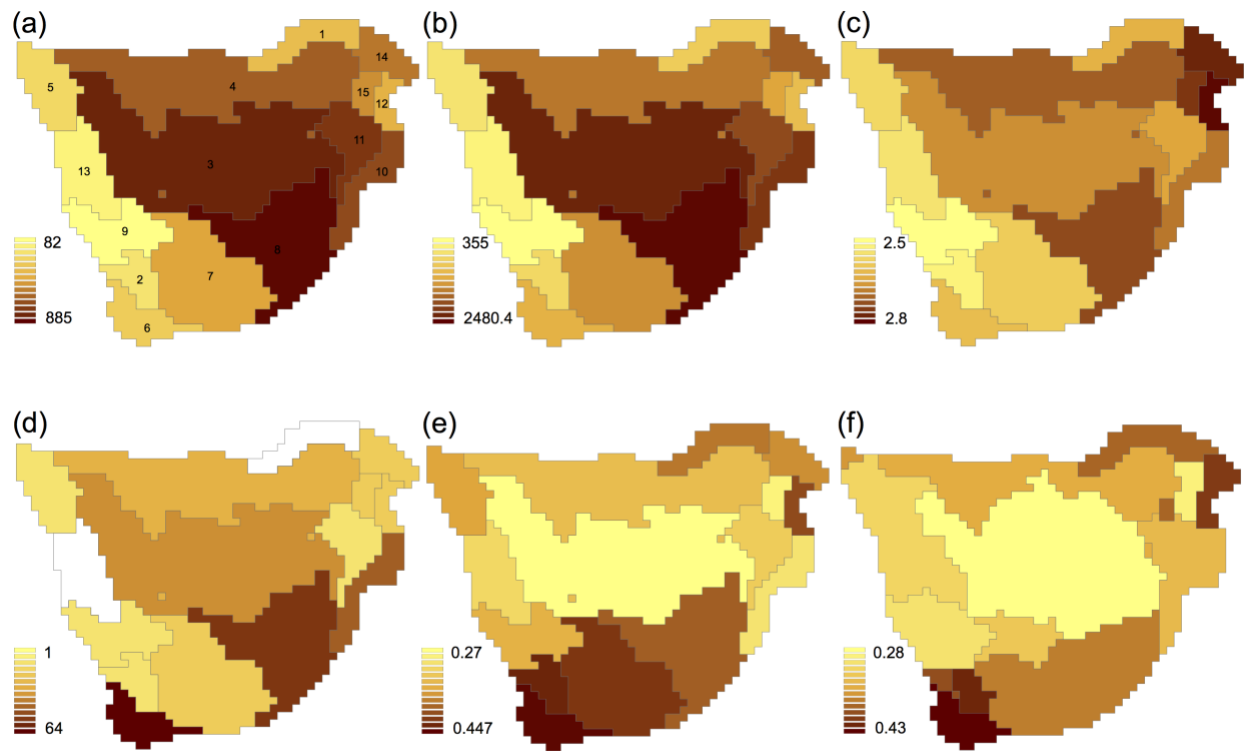


Fig. 5. Diversity metrics for the woody plants of southern Africa within phyloregions. (a) species richness, (b) PD-endemism, (c) average EDGE, (d) species endemism (e) evolutionary distinctiveness, and (f) evolutionary distinctiveness of a subset of non-threatened taxa, assuming currently threatened species go extinct. 1 = Miombo Woodland III, 2 = Succulent Karoo, 3 = Savanna, 4 = MiomboWoodland I, 5 = Namib Desert II, 6 = Fynbos, 7 = Nama Karoo, 8 = Grassland, 9 = Gariiep Karoo, 10 = Indian Ocean Coastal Belt I, 11 = Zambezian transition zone, 12 = Indian Ocean Coastal Belt II, 13 = Namib Desert I, 14 = MiomboWoodland IV, 15 = Miombo Woodland II (see Daru *et al.*, 2016).

This split is also reflected in the hierarchical dendrogram of dissimilarity, and indicates that the differentiation of the present-day vegetation types may have initially progressed along a north-south axis.

Present-day phyloregions differ in relative rank depending on the biodiversity index evaluated. For example, Fynbos emerged as the most evolutionarily distinct phyloregion (mean $p\beta_{sim} = 0.447$) (Fig. 5), whereas the Grassland phyloregion that overlaps parts of the Eastern Cape to Drakensberg in KwaZulu-Natal and Gauteng captured the most unshared phylogenetic diversity (PD-endemism = 2480.36). In contrast, the Indian Ocean Coastal Belt and Miombo Woodland phyloregion in Mozambique had the highest proportion of phylogenetic diversity under threat (mean EDGE; Fig. 5c). We also found that hotspots based on grid cells capture only 13 out of 15 phyloregions, and any one single conservation metric prioritises a smaller set of phyloregions (see Table S2).

With the removal of currently threatened species, present-day phyloregions became less distinct, and with this increasing homogeneity we are no longer able to distinguish some phyloregions (Fig. 5f). For instance, we lose the newly defined ‘Zambezi transition zone’ (Daru *et al.*, 2016) that overlaps south-western Mozambique at the border between Zimbabwe and South Africa (Fig. 5f), and we are also unable to differentiate the Namib Desert I and Nama Karoo phyloregions, which now collapse into the Grassland and Namib Desert respectively (Fig. 5f).

DISCUSSION

Here we have presented a novel approach for exploring the evolutionary history and geographic organization of biotic diversity using information from a dated phylogenetic tree of extant taxa. By exploring phylogenetic affinities at successive phylogenetic depths we are able to make inferences on the historical sequence of the formation of biogeographical assemblages and gain insights into the history of dispersal, extinction, and speciation of lineages between and within them. In the flora of southern Africa, we show that the emergence of modern vegetation types was recent, and probably within the last 40-20 myr, suggesting that the present day species distributions may retain some signature of the earlier migration and habitat filtering of lineages across Africa. By looking backwards in time, we identify phyloregions with long evolutionary histories and those that represent relatively recent separations. We suggest not only that phyloregions represent important conservation units, but just as species with long evolutionary histories might deserve more conservation attention (Vane-Wright *et al.*, 1991), so too might equivalent phyloregions. Such phyloregions encompass lineages that have co-evolved over many millions of years and capture the biogeographical processes that have shaped modern diversity patterns (Holt *et al.*, 2013).

Undoubtedly, there have been major floral and faunal shifts over the past several million years, and the geographic distribution of extant species we observe today obviously does not necessarily reflect the geography of the past. Our depiction of the spatial aggregation of vegetation types as we step back in time is only provided to illustrate spatially the phylogenetic linkages among them, and should not be interpreted as representative of historical distributions. Nonetheless, present-day vegetation types tend to be characterized by a phylogenetically distinct set of lineages, and using a dated phylogenetic tree of the woody flora, we show how phylogeny

provides a window into their evolutionary history that complements approaches using fossil data and climate reconstructions.

Phyloregions and the origins of the vegetation types of southern Africa

Our results indicating a relatively recent origin of present-day vegetation types in southern Africa match to previous studies focused on particular taxa (*e.g.*, Richardson *et al.*, 2001; Linder & Hardy, 2004) and changes in atmospheric CO₂ (Kürschner *et al.*, 2008). The evolution of the Fynbos of the Cape floristic region (CFR) has attracted perhaps the most attention, and has been linked to the upwelling of cold waters – the Benguela upwelling system (BUS) – along the Atlantic seaboard of southern Africa during late Miocene about 10-8 Ma (Siesser, 1980). Varying dates for the development of the Fynbos biome have been proposed, depending on the lineages examined. Some studies have inferred younger dates, <5 Ma (Cowling & Pressey, 2001), whereas others suggest much older dates within the Oligocene or early Miocene (Goldblatt *et al.*, 2002, Linder & Hardy, 2004). Our study supports a Miocene origin; however, the Fynbos vegetation type might only have become more widespread later, explaining the relative sparsity of Fynbos vegetation in the fossil record. The strong phylogenetic signal in membership of the Fynbos and Karoo vegetation types likely reflects recent *in situ* radiation of particular clades, such as *Protea* and *Leucadendron*, in these phyloregions (Linder, 2003), further emphasizing their evolutionary distinctiveness.

Our analysis also sheds light on the differentiation of modern-day Savanna and Miombo Woodland around the mid-Miocene ~12-8 Ma. The Miocene was a period characterized by extreme climatic fluctuations and followed a major cooling event at the Oligocene-Miocene

boundary that has been suggested to coincide with the origin of many modern terrestrial biomes worldwide (Kürschner *et al.*, 2008). The spread of Savanna in Africa during the Miocene coincides with the rise to dominance of flammable C₄ grasses and shifts in atmospheric CO₂ concentrations (Osborne & Beerling, 2006; Kürschner *et al.*, 2008). The Miombo Woodland, which covers much of Zimbabwe, Mozambique and stretches across the continent to Angola (Campbell *et al.*, 1996) has been variously categorized as Savanna (Huntley, 1982), Woodland (White, 1983) and Forest (Freson *et al.*, 1974; Malaisse, 1978), although it differs from these formations by the dominance of three Fabaceae genera: *Brachystegia*, *Julbernardia* and *Isoberlinia* (Campbell *et al.*, 1996). Geological records indicate that Miombo Woodland also arose during the Miocene ~25-7 Ma, corresponding to the formation of the Central African plateau (Lister, 1987). Our results highlight the close phylogenetic affinities between Savanna and Miombo Woodland, which are also reflected in the weaker phylogenetic conservatism observed in the Savanna vegetation, and favors the younger date for the differentiation between the two vegetation types.

Another feature of interest that emerges from our results is the differentiation of the Grassland vegetation. The present-day Grassland in southern Africa is dominated by C₄ grasses of the Poaceae family and occurs inland in the summer rainfall areas of South Africa covering the highveld areas, eastern seaboard, and the high mountains along the east coast (Mucina & Rutherford, 2006). However, plants other than grasses such as herbaceous elements (forbs) and woody species constitute an important part of the native grassy biomes. These woody lineages are adapted to the unique ecological and abiotic features of the grassland biome, which include particular rainfall regimes, nutrient availability, and generally low competition for light (Parr *et*

al., 2014), and they thus allow us to define the biome even though they do not represent the dominant lifeforms within it. Paleocological evidence indicates that grasses evolved ~70-55 Ma (Kellogg, 2001) and gradually extended their distribution into tropical woodland (Bredenkamp *et al.*, 2002), but that the rapid spread of grasses came only later ~45-30 Ma triggered by increases in global aridity (Zachos *et al.*, 2001) and mediated by frequent disturbances such as fire and animal grazing (Bond *et al.*, 2003; Woodward *et al.*, 2004). Flammable C₄ grasses appeared ~30-25 Ma in the tropical and arid regions of Africa when atmospheric CO₂ was low, and climatic conditions were hot and dry (Kellogg, 2001), and are then thought to have spread rapidly into other plant biomes ~8-4 Ma (Woodward *et al.*, 2004). Once again, the concordance with our results is striking. We find phyloregions corresponding to present-day Grassland and encompassing parts of the Nama Karoo, Indian Ocean Coastal Belt and Albany thickets appear between 40 Ma and 20 Ma, whereas the differentiation of the Grassland from the Nama Karoo occurs ~8 Ma.

The timescales of the past and future changes are very different (current species extinctions and climate change are occurring on scales of decades, but historical events rolled over 10's or 100's of thousands of years). Secondly, we suggest that, while past climate change likely influenced species distributions, it is possible that co-adapted species shifted in synchrony. For example, the Fynbos phyloregion may have shifted and expanded southwards as the climate became drier, but the composition of lineages within it may have remained relatively stable. Thus, we can still draw inference on the timing of the evolutionary origins of the Fynbos and Savanna biomes, and we show that our age estimates fit well with current theory. Indeed, we

show that lineages dated to 20 Ma form distinct phyloregions that resemble the signature of present-day phyloregions.

Phyloregions as units for conservation

The phyloregions presented here represent evolutionary coherent units, a product of species assembly processes and co-evolutionary dynamics that have taken place over many millions of years. While we do not mean to suggest that phyloregions are better conservation units than more standard measures of species richness or endemism, phyloregions can help identify floras overlooked by traditional floristic studies, such as the newly defined Gariep Karoo and Zambezian transition zone in southern Africa (Daru *et al.*, 2016).

We identified phyloregions that are the most evolutionarily distinct; here, the Fynbos stands out, enclosing lineages that tend to be conserved in their ancestral areas of origin (Pennington *et al.*, 2006). Phyloregions with high evolutionary distinctness might represent geographic regions that harbor species with rare features not found elsewhere (*i.e.*, locally restricted radiations), and thus rank highly for conservation purposes (Holt *et al.*, 2013; Daru *et al.* 2016; Daru *et al.*, 2017a). However, evolutionarily distinct phyloregions do not necessarily capture the greatest taxonomic richness or unique phylogenetic diversity. Nonetheless, it is possible to integrate more traditional conservation metrics, such as richness, and threat within a phyloregions framework.

We indexed phyloregions using common conservation currencies, and show that while the Fynbos is the most evolutionarily distinct, phyloregions of the Eastern Cape, Drakensberg and Gauteng, capture greater species and unique phylogenetic diversity (*i.e.*, PD-endemism).

While high evolutionary distinctness of the Fynbos might reflect its history as a centre of speciation (*e.g.*, Forest *et al.*, 2007), the Grassland phyloregion, with high PD-endemism, is a likely centre of neo- and paleo-endemism (Mishler *et al.*, 2014). The Grassland phyloregion thus ranks highly with respect to irreplaceability. However, we are losing some parts of the phylogenetic tree more rapidly than other parts (Vane-Wright *et al.*, 1991; Davies, 2015; Veron *et al.*, 2017). If we wish to minimize the loss of evolutionary history, we may therefore want to focus conservation efforts on those areas where evolutionary history is currently most at risk, for this our priority shifts to Mozambique where the phyloregions of the Indian Ocean Coastal Belt and Miombo Woodland enclose a high proportion of species which are both threatened and evolutionarily distinct.

Missing taxa and phylogenetic error

Although our study focused on woody plants, we predict that analyses of non-woody taxa such as grasses, herbs, forbs might show similar patterns to those revealed here, but currently the data is not available to explore this further. We recognize that incomplete sampling (*i.e.*, missing taxa) or misplaced taxa on the phylogeny could potentially influence reconstructed phyloregions, especially when such biases are geographically non-random. We did not explicitly evaluate the impact of phylogenetic error in our analyses; however, we have reason to believe that any effect might be relatively small, and certainly less biased than when estimating diversification rates from phylogenies, where exact branch lengths are critical. First, our sampling includes all the major woody lineages, so missing taxa will contribute relatively little additional unique phylogenetic diversity. Second, phyloregions are generated by analyzing multiple pairwise

distances between many taxa. One or a few misplaced taxa would thus have only small leverage. Third, phyloregions represent information on both species composition and phylogenetic distance, and similar clusters can be generated considering only species composition (although all data on evolutionary distinctiveness, age of assemblages and their phylogenetic relationships is obviously lost). Last, phylogenetic errors will tend to homogenize phyloregions, reducing the apparent separation between them, rather than generate spurious structure. At the extreme, if species were distributed at random across space, we would only return a single homogeneous phyloregion. Nonetheless, it is still possible that omitting an endemic radiation within a particular phyloregion could add error to our estimates. However, our phyloregions demonstrate a remarkable congruence with currently recognized vegetation types both in distribution and age, giving us confidence in our findings.

The evolutionary future

If currently at-risk species become extinct, the vegetation of the African flora might be very different in the future, and the evolutionary coherence of phyloregions may be lost. For instance, we would no longer be able to distinguish the newly defined 'Zambezi transition zone' (Daru *et al.*, 2016) overlapping south-western Mozambique, Zimbabwe and South Africa. This region has been characterized as a 'mixing zone' for long-separated African flora migrating from the Cape or the Drakensberg in the south to South Central Africa in the north (Weimarck, 1941; Van Wyk & Smith, 2001; Galley *et al.*, 2007). However, part of this phyloregion in Mozambique is already experiencing high degree of threat from timber logging (Fath, 2001), and we risk losing many of the evolutionarily distinct and globally endangered species that are found within it

(Daru *et al.*, 2015). The loss of distinct phyloregions might suggest greater phylogenetic homogenization, potentially destabilizing centres of diversification and endemism.

It is possible that climate change could shift the spatial distribution of species, and thus phyloregions, perhaps over shorter periods of time. However, phyloregion responses may tend to lag behind individual species responses (*i.e.*, phyloregions will tend to be more stable over time than individual species distributions).

ACKNOWLEDGMENTS

We thank the Government of Canada through Genome Canada and Ontario Genomics Institute (2008-OGI-ICI-03), International Development Research Centre (IDRC), and University of Johannesburg for financial support; and the Research Computing Clusters of Harvard University (<https://rc.fas.harvard.edu/>) for statistics support. We thank three anonymous reviewers for improving earlier versions of the manuscript.

References

- Bell, C.D., Soltis, D.E. & Soltis, P.S. (2010) The age and diversification of the angiosperms revisited. *American Journal of Botany*, **97**, 1296–1303.
- Bond, W.J., Midgley, G.F. & Woodward F.I. (2003) The importance of low atmospheric CO₂ and fire in promoting the spread of grasslands and savannas. *Global Change Biology*, **9**, 973–982.
- Boon, D. (2010) *Pooley's Trees of Eastern South Africa – A complete guide*. Flora and Fauna Publication Trust, Durban, South Africa.
- Bredenkamp, G.J., Spada, F. & Kazmierczak E. (2002) On the origin of northern and southern hemisphere grasslands. *Plant Ecology*, **163**, 209–229.
- Burgess, N., Hales, J.D., Underwood, E., Dinerstein, E., Olson, D., Itousa, I., Schipper, J., Ricketts, T. & Newman, K. (2004) *The terrestrial ecoregions of Africa and Madagascar*. Island Press, Washington DC, USA.
- Campbell, B.M., Frost, P. & Byron, N. (1996) Miombo woodlands and their use: overview and key issues. *The Miombo in Transition: Woodlands and Welfare in Africa* (ed. by B. Campbell), pp. 1-10. Center for International Forestry Research, Indonesia.
- Cowling, R.M. & Pressey R.L. (2001) Rapid plant diversification: Planning for an evolutionary future. *Proceedings of the National Academy of Sciences USA*, **98**, 5452–5457.
- Daru, B.H., Holt, B.G., Lessard, J.P., Yessoufou, K. & Davies, T.J. (2017a) Phylogenetic regionalization of marine plants reveals close evolutionary affinities among disjunct temperate assemblages. *Biological Conservation*, **213**, 351–356.

- Daru, B.H., Elliott, T.L., Park, D.S. & Davies, T.J. (2017b) Understanding the processes underpinning patterns of phylogenetic regionalization. *Trends in Ecology and Evolution*, doi: 10.1016/j.tree.2017.08.013
- Daru, B.H., Van der Bank, M. & Davies, T.J. (2015) Spatial incongruence among hotspots and complementary areas of tree diversity in southern Africa. *Diversity and Distributions*, **21**, 769–780.
- Daru, B.H., Van der Bank, M., Maurin, O., Yessoufou, K., Schaefer, H., Slingsby, J.A. & Davies, T.J. (2016) A novel phylogenetic regionalization of the phytogeographic zones of southern Africa reveals their hidden evolutionary affinities. *Journal of Biogeography*, **43**, 155–166.
- Davies, T. J. & Cadotte M.W. (2011) Quantifying biodiversity –does it matter what we measure? *Biodiversity Hotspots* (ed. by F.E. Zachos & J.C. Habel), pp. 43-60. Springer Berlin Heidelberg.
- Davies, T.J. (2015) Losing history: how extinctions prune features from the tree-of-life. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**, 20140006.
- Davies, T.J. & Buckley, L.B. (2011) Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2414–2425.
- Davies, T.J. & Buckley, L.B. (2012) Exploring the phylogenetic history of mammal species richness. *Global Ecology and Biogeography*, **21**, 1096–1105.
- Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

- Edwards, E.J., Osborne, C.P., Strömberg, C.A.E., Smith, S.A. & C₄ Grasses Consortium. (2010) The origins of C₄ Grasslands: integrating evolutionary and ecosystem science. *Science*, **328**, 587–591.
- Fath, H. (2001) *Commercial timber harvesting in the natural forests of Mozambique*. Food and Agriculture organization of the United Nations, Rome.
- Faith, D.P., Reid, C.A.M. & Hunter, J. (2004) Integrating phylogenetic diversity, complementarity, and endemism for conservation assessment. *Conservation Biology*, **18**, 255–261.
- FitzJohn, R.G., Pennell, M.W., Zanne, A.E., Stevens, P.F., Tank, D.C. & Cornwell W.K. (2014) How much of the world is woody?. *Journal of Ecology*, **102**, 1266–1272.
- Forest, F., Grenyer, R., Rouget, M., Davies, T.J., Cowling, R.M., Faith, D.P., Balmford, A., Manning, J.C., Proches, S., van der Bank, M., Reeves, G., Hedderson, T.A. & Savolainen, V. (2007) Preserving the evolutionary potential of floras in biodiversity hotspots. *Nature*, **445**, 757–760.
- Freson, R., Goffinet, G. & Malaisse, F. (1974) Ecological effects of the regressive succession in muhulu-miombo-savanna in Upper Shaba, Zaire. *Proceedings of the first international congress of ecology*, pp. 365-371. Structure, functioning and management of ecosystems, PUDOC, Wageningen.
- Fritz, S.A. & Purvis, A. (2010) Selectivity in mammalian extinction risk and threat types: A new measure of phylogenetic signal strength in binary traits. *Conservation Biology*, **24**, 1042–1051.

- Galley, C., Bytebier, B., Bellstedt, D.U. & Linder, H.P. (2007) The Cape element in the Afrotropical flora: from Cape to Cairo? *Proceedings of the Royal Society B: Biological Sciences*, **274**, 535–543.
- Goldblatt, P., Savolainen, V., Porteous, O., Sostaric, I., Powell, M., Reeves, G., Manning, J.C., Barraclough, T.G. & Chase, M.W. (2002) Radiation in the Cape flora and the phylogeny of peacock irises *Moraea* (Iridaceae) based on four plastid DNA regions. *Molecular Phylogenetics and Evolution*, **25**, 341–360.
- Gower, J.C. (1983) Comparing classifications. *Numerical taxonomy* (ed. by J. Felsenstein), pp. 137–155. Springer, Berlin, Germany.
- Guo, Y.Y., Luo, Y.B., Liu, Z.J. & Wang, X.Q. (2012) Evolution and biogeography of the Slipper Orchids: Eocene vicariance of the *Conducuplicate* genera in the Old and New World tropics. *PLoS ONE*, **7**, e38788.
- Holt, B.G., Lessard, J.P., Borregaard, M.K., Fritz, S.A., Araujo, M.B., Dimitrov, D., Fabre, P.H., Graham, C.H., Graves, G.R., Jønsson, K.A., Nogues-Bravo, D., Wang, Z., Whittaker, R.J., Fjeldsa, J. & Rahbek, C. (2013) An update of Wallace’s zoogeographic regions of the world. *Science*, **339**, 74–78.
- Huntley, B.J. (1982) Southern African savannas. *Ecology of tropical savannas* (eds. B.J. Huntley & B.H. Walker), pp. 101-109. Springer-Verlag, Berlin, Germany.
- Isaac, N.J., Turvey, S.T., Collen, B., Waterman, C. & Baillie, J.E. (2007) Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS ONE*, **2**, e296.
- James, A., Gaston, K.J. & Balmford, A. (2001) Can we afford to conserve biodiversity? *BioScience*, **51**, 43–52.

- Kellogg, E.A. (2001) Evolutionary history of the grasses. *Plant Physiology*, **125**, 1198–1205.
- Kürschner, W.M., Kvacek, Z. & Dilcher, D.L. (2008) The impact of Miocene atmospheric carbon dioxide fluctuations on climate and the evolution of terrestrial ecosystems. *Proceedings of the National Academy of Sciences USA*, **105**, 449–453.
- Linder, H.P. (2003) The radiation of the Cape flora, southern Africa. *Biological Reviews*, **78**, 597–638.
- Linder, H.P. & Hardy, C.R. (2004) Evolution of the species rich Cape flora. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1623–1632.
- Lister, L.A. (1987) *The erosion surfaces of Zimbabwe*. Zimbabwe Geological Survey Bulletin No. 90. Geological Survey Department, Harare.
- Low, A.B. & Rebelo, A.G. (1996) *Vegetation of South Africa, Lesotho and Swaziland*. Department of Environmental Affairs and Tourism, Pretoria.
- Malaisse, F.P. (1978) *The miombo ecosystem*. Natural Resources Research 14. UNESCO, Paris.
- Margules, C.R. & Pressey R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.
- Maurin, O., Davies, T.J., Burrows, J.E., Daru, B.H., Yessoufou, K., Muasya, A.M., van der Bank, M. & Bond, W.J. (2014) Savanna fire and the origins of the “underground forests” of Africa. *New Phytologist*, **204**, 201–214.
- Mishler, B.D., Knerr, N., González-Orozco, C.E., Thornhill, A.H., Laffan, S.W. & Miller, J.T. (2014) Phylogenetic measures of biodiversity and neo- and paleo-endemism in Australian *Acacia*. *Nature Communications*, **5**, 4473.
- Mittermeier, R.A., Myers, N., Gill, P.C. & Mittermeier, C.G. (2000) *Hotspots: Earth’s Richest and Most Endangered Terrestrial Ecoregions*. CEMEX, Mexico City.

- Mucina, L. & Rutherford, M.C. (2006) *The vegetation of South Africa, Lesotho and Swaziland. Strelitzia 19*. South African National Biodiversity Institute, Pretoria.
- Myers, N. (1988) Threatened biotas: 'hot-spots' in tropical forests. *The Environmentalist*, **8**, 187–208.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.
- O'Brien, E.M. (1993) Climatic gradients in woody plant species richness: towards an explanation based on an analysis of southern Africa's woody flora. *Journal of Biogeography*, **20**, 181–198.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C. Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. (2001) Terrestrial ecoregions of the worlds: A new map of life on Earth. *BioScience*, **51**, 933–938.
- Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, I. & Pearse, W. (2012) *caper: Comparative Analyses of Phylogenetics and Evolution in R*. R package version 0.5. URL <http://CRAN.R-project.org/package=caper>
- Osborne, C.P. & Beerling, D.J. (2006) Nature's green revolution: the remarkable evolutionary rise of C₄ plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **361**, 73–194.
- Palgrave, C.M. (2002) *Keith Coates Palgrave trees of southern Africa*, 3rd edn. Struik, Cape Town.

- Parr, C.L., Lehmann, C.E.R., Bond, W.J., Hoffmann, W.A. & Andersen, A.N. (2014) Tropical grassy biomes: misunderstood, neglected, and under threat. *Trends in Ecology & Evolution*, **29**, 205–213.
- Pennington, R.T., Cronk, Q.C.B. & Richardson, J.A. (2004) Introduction and synthesis: plant phylogeny and the origin of major biomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1455–1464.
- Pennington, R.T., Richardson, J.E. & Lavin, M. (2006) Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytologist*, **172**, 605–616.
- Richardson, J.E., Weitz, F.M., Fay, M.F., Cronk, Q.C.B., Linder, H.P., Reeves, G. & Chase, M.W. (2001) Rapid and recent origin of species richness in the Cape flora of South Africa. *Nature*, **412**, 181–183.
- R Core Team (2015) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rosauer, D.F., Ferrier, S., Williams, K.J., Manion, G., Keogh, J.S. & Laffan, S.W. (2014) Phylogenetic generalised dissimilarity modelling: a new approach to analysing and predicting spatial turnover in the phylogenetic composition of communities. *Ecography*, **37**, 21–32.
- Salvador, S. & Chan, P. (2004) Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. *Proceedings of the Sixteenth IEEE International Conference on tools with artificial intelligence*, pp. 576–584. Institute of Electrical and Electronics Engineers, New Jersey.

- Schmidt, E., Lotter, M. & McClelland, W. (2007) *Trees and shrubs of Mpumalanga and Kruger National Park*. Jacana Education, Johannesburg.
- Siesser, W.G. (1980) Late Miocene origin of the Benguela upwelling system off northern Namibia. *Science*, **208**, 283–285.
- Sokal, R.R. & Rohlf, F.J. (1962) The comparison of dendrograms by objective methods. *Taxon*, **11**, 33–40.
- Van Wyk, A.E. & Smith, G.F. (2001) *Regions of floristic endemism in southern Africa: a review with emphasis on succulents*. Umdaus Press, Pretoria, South Africa.
- Van Wyk, B., van der Berg, E., Coates Palgrave, M. & Jordaan, M. (2011) *Dictionary of names for southern African trees*, 1st edn. Briza, Pretoria.
- Vane-Wright, R.I., Humphries, C.J. & Williams, P.H. (1991) What to protect? Systematics and the agony of choice. *Biological Conservation*, **55**, 235–254.
- Veron, S., Davies, T.J., Cadotte, M.W., Clergeau, P. & Pavoine, S. (2017) Predicting loss of evolutionary history: Where are we?. *Biological Reviews*, **92**, 271–291.
- Webb, C.O. & Donoghue, M.J. (2005) Phylomatic: tree assembly for applied phylogenetics. *Molecular Ecology Notes*, **5**, 181–183.
- Weimarck, H. (1941) Phytogeographical groups, centres and intervals within the Cape flora. *Lunds Universitets Arsskrift*, **37**, 1–143.
- White, F. (1983) *The vegetation of Africa*. Natural Resources Research 20. UNESCO, Paris.
- Woodward, F.I., Lomas, M.R. & Kelly, C.K. (2004) Global climate and the distribution of plant biomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1465–1476.

- Wu, K., Chen, C.-N.N. & Soong, K. (2016) Long distance dispersal potential of two seagrasses *Thalassia hemprichii* and *Halophila ovalis*. *PLoS ONE*, **11**, e0156585.
- Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 686–693.
- Zhao, X., Valen, E., Parker, B.J. & Sandelin, A. (2011) Systematic clustering of transcription start site landscapes. *PLoS ONE*, **6**, e23409.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.

Appendix S1 Supporting table (Table S1), figure (Fig. S1) and data (Data S1 and S2).

Table S1 Voucher information and GenBank/EBI accession numbers for trees in southern Africa. Taxa with only genus names are those which are not southern African, but are included to have a representation of the lineage.

Table S2 Presence/absence of five types of diversity metrics within existing phyloregions of southern Africa. Values indicate the number of hotspots cells overlapping phyloregions. CWE = corrected weighted endemism, EDGE = evolutionary distinctiveness and global endangerment, PD = phylogenetic diversity, PE = phylogenetic endemism, SR = species richness.

Figure S1 Evolution of modern phyloregions in southern Africa between 20 Ma to present-day. Relationships among phyloregions at each depth are depicted as hierarchical dendrograms, NMDS ordinations and in geographic space based on UPGMA clustering of phylogenetic beta diversity ($p\beta_{sim}$) within 50×50 km grid cells (see Methods). Colors differentiating between phyloregions depict the amount of phylogenetic turnover among phyloregions.

Data S1 R script for computing phylogenetic beta diversity using Simpson's index ($p\beta_{sim}$). The code uses parallel processing, with a default number of clusters set to 4. The function returns a $p\beta_{sim}$ matrix.

Data S2 Sample R scripts for slicing the phylogenetic tree at various time depths.