

ORIGINAL
ARTICLE

A novel phylogenetic regionalization of phytogeographical zones of southern Africa reveals their hidden evolutionary affinities

Barnabas H. Daru^{1,2*}, Michelle van der Bank¹, Olivier Maurin¹, Kowiyou Yessoufou^{1,3}, Hanno Schaefer⁴, Jasper A. Slingsby^{5,6} and T. Jonathan Davies^{1,7}

¹African Centre for DNA Barcoding, University of Johannesburg, APK Campus, PO Box 524, Auckland Park, 2006 Johannesburg, South Africa, ²Department of Plant Science, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa, ³Department of Environmental Sciences, University of South Africa, Florida Campus, Florida 1710, South Africa, ⁴Technische Universität München, Plant Biodiversity Research, Emil-Ramann Strasse 2, 85354 Freising, Germany, ⁵Fynbos Node, South African Environmental Observation Network, Private Bag X7, 7735 Rhodes Drive, Newlands, South Africa, ⁶Statistics in Ecology, Environment and Conservation, Department of Biological Sciences, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa, ⁷Department of Biology, McGill University, Montreal, QC H3A 0G4, Canada

ABSTRACT

Aim Although existing bioregional classification schemes often consider the compositional affinities within regional biotas, they do not typically incorporate phylogenetic information explicitly. Because phylogeny captures information on the evolutionary history of taxa, it provides a powerful tool for delineating biogeographical boundaries and for establishing relationships among them. Here, we present the first vegetation delineation of the woody flora of southern Africa based upon evolutionary relationships.

Location Southern Africa.

Methods We used a published time-calibrated phylogenetic tree for 1400 woody plant species along with their geographical distributions and a metric of phylogenetic beta diversity to generate a phylogenetic delineation of the woody vegetation of southern Africa. We then explored environmental correlates of phylogenetic turnover between them, and the evolutionary distinctiveness of the taxa within them.

Results We identified 15 phylogenetically distinct biogeographical units, here referred to as phyloregions. The largest phyloregion broadly overlaps with Savanna vegetation, while the phyloregion overlapping with the south-western portion of the Fynbos biome is the most evolutionarily distinct. Potential evapotranspiration and mean annual temperature differ significantly among phyloregions and correlate with patterns of phylogenetic beta diversity between them. Our phylogeny-based delimitation of southern Africa's woody vegetation broadly matches currently recognized phytogeographical classifications, but also highlights parts of the Namib Karoo and Greater Limpopo Transfrontier Park as distinct, but previously under-recognized biogeographical units.

Main conclusions Our analysis provides new insights into the structure and phylogenetic relationships among the woody flora of southern Africa. We show that evolutionary affinities differentiate phyloregions closely resembling existing vegetation classifications, yet also identify 'cryptic' phyloregions that are as evolutionarily distinct as some of the recognized African vegetation types.

Keywords

biogeographical regionalization, biomes, cluster analysis, evolutionary distinctiveness, phylogenetic beta diversity, phyloregions, southern Africa, vegetation types, woody flora

*Correspondence: Barnabas H. Daru, Department of Plant Science, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa.
E-mail: barnabas.daru@up.ac.za

INTRODUCTION

Unevenness in the distribution of species diversity across the globe has long been recognized. De Candolle (1820)

proposed one of the first biogeographical classification schemes for plants based on both ecology and history. In Africa, a diverse range of biogeographical schemes have been proposed using various criteria (see Table S1 in Appendix S1

of the Supporting Information). White's (1983) delineation of phytochoria based on species distribution data remains the most widely accepted. More recent studies, focusing on both biogeographical regionalization and vegetation classification, and often centred on southern Africa, have considered criteria including dominant plant growth forms (e.g. trees, shrubs, herbs and grasses) and similarity in ecological conditions (Low & Rebelo, 1996; Rutherford, 1997; Mucina & Rutherford, 2006), fire characteristics (Archibald *et al.*, 2013), and plant species distribution (Linder *et al.*, 2005, 2012). Other efforts have examined particular species groups, such as non-natives (Hugo *et al.*, 2012), bryophytes (Van Rooy & Van Wyk, 2010), flagship and hotspot species (Steenkamp *et al.*, 2005) or focused on particular areas (Irish, 1994). Here, we revisit the definition and delineation of phytogeographical regions (areas where groups of plants occur together in particular climates and form distinct vegetation types; Lomolino *et al.*, 2010) within southern Africa using a phylogenetic approach.

Southern Africa is indeed one of the most floristically diverse regions in the world, with over 22,000 recorded vascular plant species (Klopper *et al.*, 2006). Woody plants (especially trees) represent one of the most distinguishing features of the vegetation in the region (Coates Palgrave, 2002; Germishuizen & Meyer, 2013). They are the largest and longest lived of plant species (O'Brien *et al.*, 2000) and

have been used for modelling vegetation structure and physiognomy in the region and elsewhere (O'Brien, 1993; Kubota *et al.*, 2014). Within southern Africa, woody plant richness is unevenly distributed, increasing from west to east, likely reflecting the shifting water–energy dynamics (O'Brien, 1993; O'Brien *et al.*, 1998, 2000). More locally, woody plant richness is influenced by additional environmental factors, such as physiography and topographical relief (O'Brien *et al.*, 2000).

The spatial heterogeneity in climate is reflected in the plant distribution and diversity, and a number of vegetation types are currently recognized within the region, including Albany thicket, Indian Ocean Coastal Belt, Desert, Inland forest, Fynbos, Grassland, Miombo woodland (a type of woody savanna), Nama Karoo, Savanna (our usage here excludes Miombo) and Succulent Karoo (White, 1983 with modifications and additions from Low & Rebelo, 1996; Olson *et al.*, 2001; Burgess *et al.*, 2004; and Mucina & Rutherford, 2006; Fig. 1). The hyper-diverse Fynbos is renowned for its high richness of local endemics (Goldblatt, 1978; Linder, 2003), whereas the Savanna is perhaps the most widespread vegetation type. Most of these vegetation types are dominated or characterized by woody plants, and consequently phytochoria (geographical areas with relatively uniform composition of plant species) defined using woody plants would likely appear similar in distribution to biomes

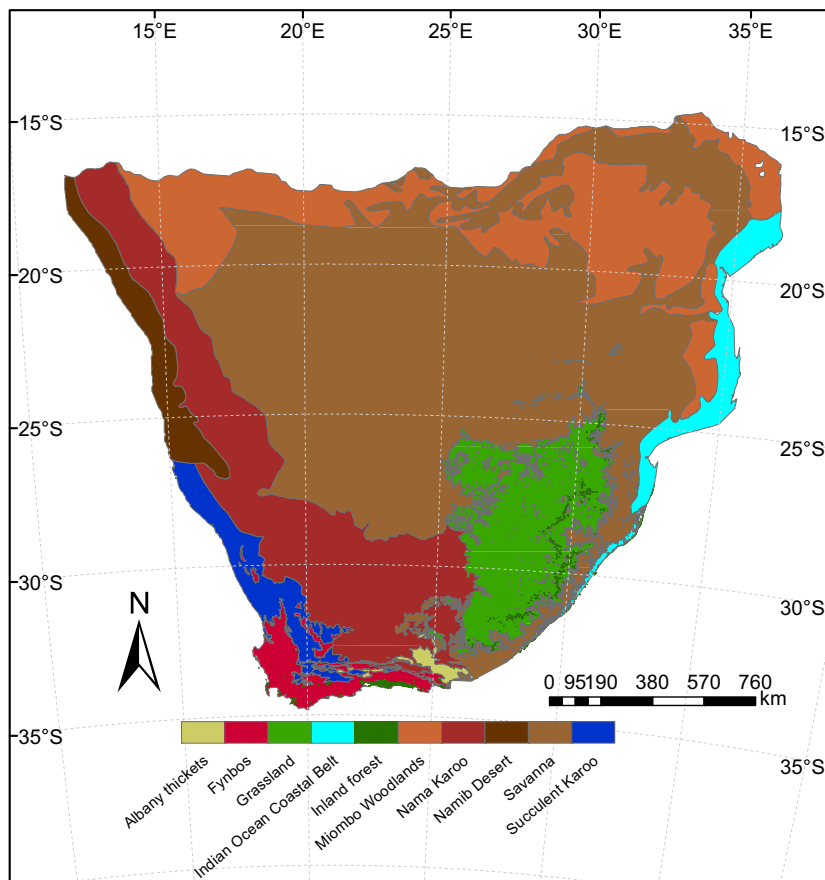


Figure 1 Map of the currently recognized broad (biome-level) vegetation types (phytochoria) of southern Africa by White (1983) with modifications and additions from Low & Rebelo (1996), Olson *et al.* (2001), Burgess *et al.* (2004) and Mucina & Rutherford (2006).

or finer vegetation units. Most current approaches for the delineation of vegetation types did not take into account possible differences in the phylogenetic relationships among species. However, considering only the number of shared species among regions might capture only poorly their evolutionary affinities, and underrate the importance of historical factors. Phylogeny provides an additional tool for differentiating biogeographical regions by allowing quantification of the phylogenetic turnover (phylogenetic beta diversity) among them (Graham & Fine, 2008; Holt *et al.*, 2013). Early phylogenetic studies showed southern Africa to be a hotspot of diversification (e.g. Linder, 2001; Davies *et al.*, 2005) with high local phylogenetic diversity (e.g. Procheş *et al.*, 2006), and thus suggest that there is likely to be strong spatial structure in the phylogenetic structure of the region's flora.

Here, we use a dated molecular phylogeny for the native woody flora of southern Africa to present a novel classification of phylogenetically defined vegetation types or 'phyloregions'. We identify phyloregions with particularly evolutionarily distinct floras, and we test the contribution of ecological factors in shaping phylogenetic membership within them using a suite of environmental variables.

MATERIALS AND METHODS

Description of data

The analysis in this study is based on the largest dated phylogenetic tree yet available for the woody plants of southern Africa (Africa south of the Zambezi river, including the countries of Botswana, southern Mozambique, Zimbabwe, Namibia, South Africa, Swaziland and Lesotho). For details on the tree and phylogenetic methods, see Maurin *et al.* (2014). We defined woody plants here as phanerophytes with stems or pseudostems > 0.5 m in height, following previous work (O'Brien, 1993; O'Brien *et al.*, 1998; Kubota *et al.*, 2014). Non-native species were excluded from our analyses. Subspecies and varieties were combined under the parent species binomials (see Table S2 in Appendix S1). Our data set includes monocots, gymnosperms and angiosperms, and is comprised of 115 families, 541 genera and 1400 species. While monocots do not produce vascular cambium, secondary growth and thickening through special kinds of cambium do occur in certain species e.g. palms and some arborescent *Aloe* species (Scarpella & Meijer, 2004), and were thus included in this study. The woody plant data set was compiled using a checklist of the woody species from the literature (Coates Palgrave, 2002; Schmidt *et al.*, 2007; Boon, 2010; Van Wyk *et al.*, 2011; Germishuizen & Meyer, 2013), and crosschecked with the African Plants database (www.ville-ge.ch/cjb/) and The Plant List (www.plantlist.org) for synonyms. The family classification followed the APG III (Angiosperm Phylogeny Group, 2009). Note that our data set undersamples the woody flora of the Fynbos region because of limitations on the availability of DNA sequence data.

Species distribution data

Range maps representing the geographical distributions for all 1400 woody plant species were sourced from Coates Palgrave (2002) and Van Wyk *et al.* (2011). These maps represent the extent-of-occurrence of the native distributions of species based on herbarium specimens, observations in the field, published monographs and field guides validated by botanical experts. All maps were scanned at 300 dpi, processed in ARCMAP 10.0 and overlaid onto a 50 km × 50 km grid (Behrmann projection) to generate a species richness map (Fig. 2) and a matrix of species presence/absences for the 1563 grid cells. Coastal grid cells with < 50% land area were excluded to minimize the influence of unequal sampling area. Although the use of range maps will tend to overestimate species area of occupancy (Quinn *et al.*, 1996; Gaston, 2003) owing to false presences, the maps used here represent the best current and predicted extent of the species distribution. Because the clustering of vegetation types is likely more sensitive to false absences (Kreft & Jetz, 2010; Holt *et al.*, 2013), we believe our regionalization analysis, based on the precise group of organisms defining vegetation units, to be robust.

Computing phylogenetic beta diversity

We evaluated taxonomic turnover between grid cells using the β_{sim} metric of beta diversity expressed as:

$$\beta_{\text{sim}} = 1 - \frac{a}{\min(b, c) + a}$$

where a is the number of shared species between two grid cells, and b and c are the numbers of species unique to each grid cell. β_{sim} values range from 0 (when species composition is identical between grid cells), to a maximum value of 1 (when there is no shared taxa). Importantly, β_{sim} is independent of differences in species richness among grid cells (Koleff *et al.*, 2003; Kreft & Jetz, 2010) and therefore provides a robust estimate of community similarity. Nonetheless, to evaluate the sensitivity of our results we additionally compared β_{sim} values (from Simpson's index) with those obtained from the Jaccard Index (β_{jaccard}).

Phylogenetic turnover was quantified using phylogenetic beta diversity ($p\beta_{\text{sim}}$), adapted from β_{sim} where the proportion of shared phylogenetic branch lengths of the dated phylogenetic tree between cells are substituted for species (Graham & Fine, 2008). Indices of $p\beta_{\text{sim}}$ range from 0 (species are identical and share the same branch lengths) to 1 (species share no phylogenetic branches). We then generated pairwise distance matrices of phylogenetic beta diversity ($p\beta_{\text{sim}}$) and beta diversity (β_{sim}) between all grid cells at the 50 × 50 km scale. We also tested the sensitivity of the phylogenetic results by comparing $p\beta_{\text{sim}}$ values (from Simpson's index) with Jaccard Index ($p\beta_{\text{jaccard}}$). All analyses were conducted in R (R Core Team, 2013), using the R-libraries *ape* (Paradis *et al.*, 2004) and *picante* (Kembel *et al.*, 2010) to integrate the phylogeny with the community matrix of

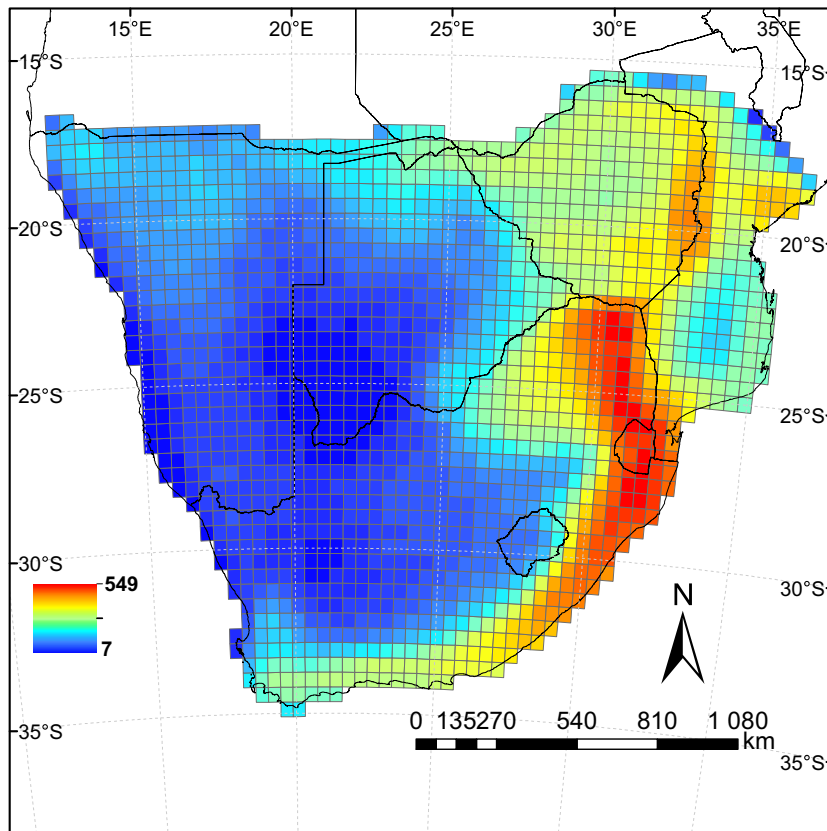


Figure 2 Woody plant species richness. Native distributions of woody flora in southern Africa across 50 × 50 km grid cells.

species presence/absences, and R-libraries cluster (Maechler *et al.*, 2013), and doSNOW (Revolution Analytics & Weston, 2013) to cluster cells, as described below.

Cluster algorithm selection and validation

To identify spatial clusters, we contrasted eight hierarchical clustering algorithms on the $p\beta_{sim}$ and β_{sim} matrices: single linkage, complete linkage, unweighted pair-group method using arithmetic averages (UPGMA), unweighted pair-group method using centroids (UPGMC), weighted pair-group method using arithmetic averages (WPGMA), weighted pair-group method using centroids (WPGMC), Ward's minimum variance and DIANA's divisive hierarchical method. These models were assessed for their degree of data distortion using Sokal & Rohlf's (1962) cophenetic correlation coefficient and Gower's distance (Gower, 1983). The cophenetic correlation defines the relationship between the terminals of a dendrogram with the original distance matrix for all cluster methods and has a value between 0 (poor correlation) and 1. Gower's distance measures the sum of squared difference between the original and cophenetic distances (Gower, 1983). A cluster method that returns the smallest Gower distance corresponds to a good clustering model for the distance matrix (Borcard *et al.*, 2011). UPGMA was identified as the best performing clustering algorithm and was thus used to define vegetation types.

Determining the appropriate stopping criteria to identify the optimal number of clusters is challenging (Milligan &

Cooper, 1985). Here, we used the 'L method' described by Salvador & Chan (2004), which corresponds to the 'elbow' in an evaluation graph, i.e. the point of optimal curvature, using the function 'elbow' within the R package GMD (Zhao *et al.*, 2011). The 'elbow' method identifies the optimal number of clusters based on the range of explained variances using squared Euclidean distances (Zhao *et al.*, 2011). We additionally evaluated the sensitivity of our analysis by setting the number of defined clusters to match the number of currently recognized vegetation types (White, 1983 with modifications and additions from Low & Rebelo, 1996; Olson *et al.*, 2001; Burgess *et al.*, 2004; and Mucina & Rutherford, 2006).

We delimited clusters (hereafter referred to as 'phyloregions') using a non-metric multidimensional scaling (NMDS) ordination and hierarchical dendrogram of dissimilarity on the $p\beta_{sim}$ matrix. The units for axes of the NMDS ordination plot are arbitrary; where similar regions occur closer together in two-dimensional space. The degree of fit of the NMDS ordination is characterized by the stress value and should be minimized. The NMDS ordination plots are shown in hue, chroma and lightness (HCL) colour space to depict the variation among vegetation types such that similar colours indicate cells that are similar in their shared species or branch lengths. The hierarchical dendrogram quantifies the topology and depth of relationships among clusters. Cluster results are presented as maps, NMDS plots and dendrograms using matching colour combinations in HCL colour space.

We estimated the evolutionary distinctiveness of each phyloregion as the mean $p\beta_{sim}$ value between the focal phyloregion and all other phyloregions following Holt *et al.* (2013). This approach estimates evolutionary distinctiveness as the average sum of branch length shared with other regions thereby identifying regions that enclose endemic radiations (unshared branches).

Last, we assessed the relative congruence between currently recognized vegetation types (White, 1983 with modifications and additions from Low & Rebelo, 1996; Olson *et al.*, 2001; Burgess *et al.*, 2004; and Mucina & Rutherford, 2006) and the regional clusters defined by species composition versus phylogeny by comparing the $p\beta_{sim}$ and β_{sim} distance matrices against a third matrix representing the geographical clustering of recognized vegetation types. The clustering of recognized vegetation types was generated as for β_{sim} (using beta diversity) except that columns of the presence/absence matrix represent vegetation types and rows correspond to grid cells. We then compared the distance matrices for $p\beta_{sim}$ and β_{sim} against the equivalent distance matrix for vegetation types using the cophenetic correlation coefficient.

Environmental comparisons between phyloregions

We evaluated differences in various environmental variables (mean annual temperature, mean annual precipitation, net primary productivity, potential evapotranspiration, elevation and annual fire return frequency) among phyloregions using analysis of similarity (ANOSIM) with a Monte Carlo randomization test of significance (Clarke, 1993). Precipitation, temperature and elevation were obtained from the WorldClim database (Hijmans *et al.*, 2005); net primary productivity (Kucharik *et al.*, 2000), fire return frequency from Archibald *et al.* (2010) and potential evapotranspiration (PET) was obtained from Willmott & Matsuura (2001) and follows Thornthwaite's equation (Thornthwaite, 1948). Potential evapotranspiration is a measure of the amount of incident solar energy and potential loss of water into the atmosphere by evaporation and plant transpiration and is used to describe energy regime. Net primary productivity was derived from remotely sense data using radiometer sensors (Moderate Resolution Imaging Spectroradiometer), and measured as the difference between the amount of CO_2 absorbed by plants and CO_2 released during respiration (Kucharik *et al.*, 2000). As fire regimes vary across different biomes (Archibald *et al.*, 2013), we obtain data on fire regime (annual fire return frequency) to test its potential in delineating vegetation types. Environmental variables were extracted as the mean within 50×50 km grid cells.

We generated a matrix of Euclidean distances from the geographical coordinates of each cell (based on a Behrmann equal area cylindrical projection) and converted each environmental variable into a matrix of pairwise distances. We ran ANOSIM (1000 replicates) on each of the six environmental matrices using phyloregions as the grouping factor. We then used a partial Mantel test (Mantel, 1967) to test for

correlation between regional $p\beta_{sim}$ values and each environmental variable in turn, while correcting for spatial autocorrelation among grid cells using the Euclidean geographical distance matrix of the grid cells.

Statistical analyses were conducted using the following R packages: APE (Paradis *et al.*, 2004), CLVALID (Brock *et al.*, 2008), CLUSTER (Maechler *et al.*, 2013), GMD (Zhao *et al.*, 2011), MCLUST (Fraley & Raftery, 2012), RASTER (Hijmans, 2015), PICANTE (Kembel *et al.*, 2010) and VEGAN (Oksanen *et al.*, 2015). Geographical data were processed in ARCMAP 10.0 (ESRI (Environmental Systems Resource Institute), 2010).

RESULTS

The geographical distribution of woody plant species richness among our sampled lineages is concentrated in the east from the Mtata district in the Eastern Cape to KwaZulu-Natal, Swaziland, Mpumalanga, Limpopo, southern Zimbabwe and parts of Mozambique (Fig. 2). Of the alternative clustering algorithms, UPGMA (unweighted pair-group method using arithmetic averages) provided the best fit between the dendrogram and the original distance matrix for both beta (β_{sim}) and phylogenetic beta ($p\beta_{sim}$) diversity (cophenetic $r = 0.726$ and 0.641 , and Gower distance = 35610.5 and 15357.2 for β_{sim} and $p\beta_{sim}$ respectively; Table S3 in Appendix S1). Single linkage performed worst for both indices (cophenetic $r = 0.292$ and 0.197 , and Gower distance = 434241.1 and 103237.9 for β_{sim} and $p\beta_{sim}$ respectively). The Jaccard Index ($p\beta_{jaccard}$) showed a strong correlation with $p\beta_{sim}$ (Mantel test; $r = 0.94$, $P < 0.001$, 999 permutations); for brevity we focus here on results for Simpson's index ($p\beta_{sim}$ and β_{sim}).

The elbow criterion of Salvador & Chan (2004) identified 15 optimal clusters for both $p\beta_{sim}$ and β_{sim} (see Fig. S1 in Appendix S1), which we used to delineate vegetation types (Fig. 3; see Fig. S2 in Appendix S1 for the comparable distribution of clusters defined by β_{sim}). We label each $p\beta_{sim}$ cluster (phyloregion) by the vegetation type with which it most closely coincides, although we note that they do not necessarily represent equivalent phytogeographical regions, and as our analysis included woody species only, we would not expect an exact match between vegetation types and our phyloregions in any case. The largest phyloregion (402 grid cells; $1,005,000 \text{ km}^2$) corresponds closely to the distribution of Savanna (excluding Miombo) – hence the Savanna phyloregion – and covers the interior part of southern Africa, followed by Miombo Woodland I with 271 cells ($677,500 \text{ km}^2$). The smallest cluster is the Miombo Woodland II, with 18 cells ($45,000 \text{ km}^2$; Table 1). Although the $p\beta_{sim}$ and the β_{sim} values are reasonably strongly correlated (Mantel test: $r = 0.84$, $P = 0.001$), perhaps surprisingly, the distance matrix for $p\beta_{sim}$ is a closer fit to the geographical clustering of currently recognized vegetation types than the distance matrix for β_{sim} (cophenetic $r = 0.331$ and 0.311 for the comparison with $p\beta_{sim}$ and β_{sim} respectively).

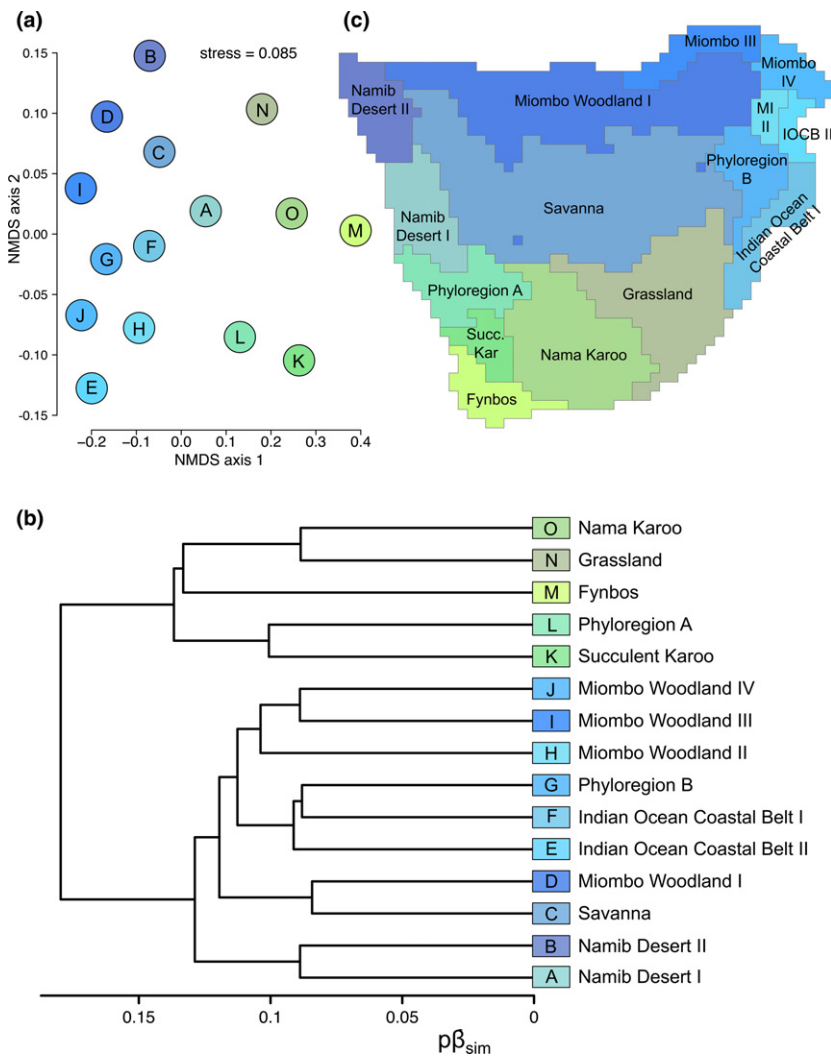


Figure 3 Relationship among phyloreions for the woody flora of southern Africa presented as an NMDS ordination plot (a), a dendrogram of dissimilarity (b) and in geographical space (c), as delimited by UPGMA hierarchical clustering of phylogenetic beta diversity ($p\beta_{sim}$) in 50×50 km grid cells. Grid cells cluster into 15 phyloreions based upon the 'L method' described by Salvador & Chan (2004). Colours differentiating between phyloreions in the NMDS plot, dendrogram and map are identical. IOCB, Indian Ocean Coastal Belt; Miombo = Miombo Woodland; Succ. Kar., Succulent Karoo.

Significantly, we also differentiate two phyloreions that are not widely recognized by current species-level vegetation classification schemes, labelled Phylloreions A and B (Fig. 3). Phylloreion A is located over portions of the Succulent Karoo (containing 146 woody species) and the Namib Desert (containing 275 woody species), with closest phylogenetic affinity to the former (Fig. 3b,c). Phylloreion B, is found to the east between Savanna/Miombo Woodland and the Indian Ocean Coastal Belt (IOCB), with phylogenetic affinities to the IOCB, and perhaps represents a transitional vegetation between Miombo Woodland and IOCB. Both Phylloreion A and Phylloreion B are retained even when the grain size of grid cells is reduced.

When we reduced the number of clusters to match the number of currently recognized vegetation types ($n = 10$), phyloreions characterized by distinct vegetation types, such as Fynbos, Namib Desert and Succulent Karoo remained, whereas phyloreions with closer phylogenetic affinities collapsed into each other, for example, Miombo Woodland merged with Savanna, Grassland merged into the Nama Karoo and the Indian Ocean Coastal Belt merged into Phylloreion B (see Fig. S3 in Appendix S1). Notably,

Phylloreion A was still differentiated, suggesting its phylogenetic signature is at least as unique as other more well-recognized vegetation types.

We further explored phylogenetic relationships among phyloreions using NMDS ordination (Fig. 3a) and a hierarchical dendrogram (Fig. 3b). The NMDS ordination indicates a separation between Fynbos-Nama Karoo-Succulent Karoo-Grassland-Phylloreion A and the other phyloreions, which is supported in the dendrogram with a major split at a $p\beta_{sim}$ of about 0.2, suggesting important phytogeographical divisions along the north-south axis of southern Africa (Fig. 3c). We also investigated the evolutionary distinctiveness (ED) of phyloreions, and found that Fynbos, Succulent Karoo and Nama Karoo have the highest mean $p\beta_{sim}$ (mean $p\beta_{sim}$ between Fynbos, Succulent Karoo and Nama Karoo and all other phyloreions = 0.447, 0.376 and 0.347 respectively; Fig. 4), emphasizing the high evolutionary uniqueness of these regions. In contrast, the ED for Savanna and Miombo Woodland I is relatively low (mean $p\beta_{sim}$ = 0.271 and 0.3125 respectively), indicating that these regions are less phylogenetically distinct (Fig. 3B). Interestingly, we note that the Miombo Woodland phyloreions do not all cluster

together on the NMDS ordination and hierarchical dendrogram, but form part of a larger cluster with Savanna and the IOCB.

Table 1 Summary of vegetation types based on UPGMA clustering of phylogenetic beta diversity ($p\beta_{sim}$) values for woody species assemblages within 50×50 km grid cells in southern Africa.

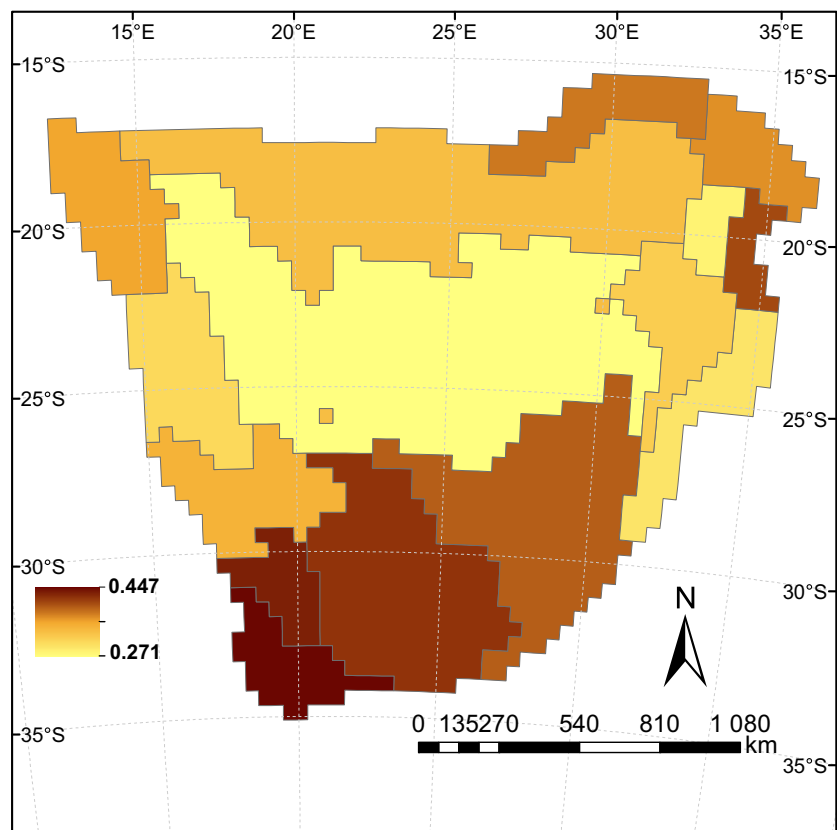
Phyloreregion (Common name)	Area in km ² (number of grid cells)	Total PD (SR)
Savanna	1,005,000 (402)	16067.65 (824)
Miombo Woodland I	677,500 (271)	14747.47 (705)
Grassland	455,000 (182)	16292.95 (885)
Nama Karoo	430,000 (172)	11221.52 (461)
Namib Desert I	190,000 (76)	3934.11 (106)
Phyloreregion A (New; between Succulent Karoo and Namib Desert)	185,000 (74)	3341.37 (82)
Namib Desert II	167,500 (67)	5564.97 (169)
Phyloreregion B (New; between Savanna, Miombo woodland and Indian Ocean Coastal Belt)	145,000 (58)	15655.22 (803)
Indian Ocean Coastal Belt I	132,500 (53)	15151.39 (732)
Miombo Woodland III	122,500 (49)	9757.67 (378)
Fynbos	120,000 (48)	8364.91 (358)
Miombo Woodland IV	95,000 (38)	13712.76 (567)
Succulent Karoo	87,500 (35)	4819.87 (146)
Indian Ocean Coastal Belt II	50,000 (20)	10424.53 (437)
Miombo Woodland II	45,000 (18)	13090.09 (554)

PD, phylogenetic diversity; SR, species richness.

Monocots and gymnosperms could potentially have a strong impact on the delineation of phyloregions because they represent evolutionarily distinct clades. Therefore, to examine sensitivity of our results, we re-ran our analysis excluding these two groups ($n = 36$ and 43 taxa for monocots and gymnosperms respectively). Using Mantel tests, we found that the $p\beta_{sim}$ values for the entire species assemblages and when excluding these two groups are strongly correlated ($r = 0.98$, $P < 0.001$; 999 permutations). While the optimal number of clusters (phyloregions) decreased to 13, we still resolve Phyloregions A and B as distinct, although Phylore-region A is now recognized as two separate regions.

To evaluate potential explanations for the separation of phyloregions, we explored differences in environmental factors among them. We found that phyloregions differ significantly in several key environmental variables (see Fig. S4 in Appendix S1). Spatial turnover in potential evapotranspiration (PET) and mean annual temperature (MAT), are the most important environmental variables distinguishing between phyloregions (ANOSIM, $r = 0.49$ and 0.48 , respectively, both $P = 0.001$ from 999 permutations). Mean annual precipitation (MAP), elevation and net primary productivity (NPP), are also significant (ANOSIM, $r = 0.44$, 0.39 and 0.37 respectively, all $P = 0.001$). However, fire return frequency, thought to be important in the genesis of Savanna (Beerling & Osborne, 2006), is less effective in differentiating among phyloregions (ANOSIM, $r = 0.04$, $P = 0.012$), although it is retained as a significant variable in fire-dominated vegetation types, such as Savanna and Miombo

Figure 4 Evolutionary distinctiveness of woody plant species within the 15 phyloregions, quantified as the mean of pairwise $p\beta_{sim}$ values between each phyloreregion and all other phyloregions. Darker regions indicate regions of higher evolutionary distinctiveness.



Woodland (see Fig. S4 in Appendix S1). We further tested the influence of these environmental variables on phylogenetic turnover by examining partial r values of the correlation between the $p\beta_{\text{sim}}$ matrix and environmental factors, accounting for geographical non-independence among grid cells by including the Euclidean distance matrix. Again, we found strong correlation of $p\beta_{\text{sim}}$ values with matrices of MAT, PET and elevation (partial Mantel test, $r = 0.46, 0.35$ and 0.14 , all $P = 0.001$, 999 permutations), but not for fire return frequency (partial Mantel test, $r = 0.015$, $P = 0.20$, 999 permutations), MAP (partial Mantel test, $r = -0.10$, $P = 1$, 999 permutations) or NPP (partial Mantel test, $r = -0.187$, $P = 1$, 999 permutations).

DISCUSSION

We used data on the phylogenetic relatedness and spatial distribution of woody plant species to delimit phylogenetically distinct vegetation types for southern Africa, an area well known for its hyper-diverse flora (Goldblatt, 1978). We identified 15 distinct regions, which we term phyloregions, defined by the phylogenetic affinities of the woody taxa within them. Our phylogenetic regionalization of the woody flora broadly matches with previous delineations of the major vegetation types of southern Africa (White, 1983; Olson *et al.*, 2001; Mucina & Rutherford, 2006), including the Fynbos, Savanna and Succulent Karoo (Low & Rebelo, 1996; Mucina & Rutherford, 2006). However, we also identify two distinct phyloregions that have been largely overlooked in past vegetation classification schemes.

We show that, on average, species in the Fynbos are the most evolutionarily distinct, perhaps indicating some evidence for phylogenetic niche conservatism (Wiens, 2004), and a tendency for lineages to maintain their ancestral area of origin (Pennington *et al.*, 2006) within this vegetation type. The Fynbos has often been the focus of increased attention as a theatre of spectacular evolutionary events (Linder, 2003; Verboom *et al.*, 2003), and many of the species within this vegetation type represent endemic radiations. One potential bias in our study was the under-sampling of woody taxa within the Fynbos vegetation.

While Savanna was the largest phyloregion in southern Africa, we found that the Grassland phyloregion hosted the highest woody plant species richness (while acknowledging the sampling limits within Fynbos) and phylogenetic diversity. The Grassland phyloregion occurs predominantly in South Africa, extending from the central plains of Eastern Cape to mountainous areas of Drakensberg in KwaZulu-Natal, and the Highveld in Gauteng (Mucina & Rutherford, 2006). This vegetation type has been described as a transitional vegetation zone that arose during the late Miocene (perhaps due to the mixing of long separated biota), coinciding with the development of the Nama Karoo and the Fynbos vegetation types (Scott *et al.*, 1997; Linder, 2003), and has been considered as a stepping-stone habitat for several Cape taxa migrating into tropical Africa (Galley *et al.*, 2007).

Our study shows that it also harbours a distinct woody flora, and a number of locally endemic woody species such as *Cussonia paniculata* (Araliaceae), *Greyia sutherlandii* (Greyiaceae) and the bamboo *Thamnocalamus tessellatus* (Poaceae) that contribute to the rich evolutionary history of the area. However, our finding that the Grassland phyloregion is the richest in woody plants may in part also reflect the inclusion of taxa from the Indian Ocean Coastal Belt with which the Grassland phyloregion overlaps, and the patchy distribution of Afromontane, mistbelt and subtropical forests, and thicket vegetation types that are nested within it. In our analyses, therefore, the Grassland phyloregion represents a species rich amalgamation of several less distinct vegetation types.

The Savanna and Miombo Woodland phyloregions demonstrate strong geographical and phylogenetic affinities. Savanna, one of the defining features of the African landscape, is the dominant vegetation in southern Africa, covering about 40% of the interior part of the region. Miombo woodland dominates much of Zimbabwe and some parts of Mozambique. The close association between these two vegetation types has been noted previously, based on their similar growth forms (grasses mixed with trees or shrubs) and climate (long periods of drought, and high frequency of fire), and they have sometimes been considered a single vegetation unit (Low & Rebelo, 1996; Mucina & Rutherford, 2006; Archibald *et al.*, 2010). Our analysis supports this grouping, while also pointing out a certain degree of distinctness. African savannas are believed to have emerged from a forest vegetation that spread during the Miocene, coincident with the rise to dominance of flammable C_4 grasses and expansion of the Savanna biome world-wide (Osborne & Beerling, 2006). Thus, the separation of Miombo Woodland and Savanna was likely relatively recent.

We characterize two vegetation types (Phyloregion A and Phyloregion B) that are not widely recognized in many current classification schemes (e.g. White, 1983), but bear similarity to habitat types defined by some diversity analyses of plant species distribution data (Van Wyk & Smith, 2001). We show that these under-recognized phyloregions are at least as phylogenetically distinct as some of the more familiar vegetation types in the region, and they remain distinguishable in each of our sensitivity analyses.

We refer to Phyloregion A as 'Gariep Karoo' based on its close phylogenetic and geographical similarity to the Succulent Karoo on the west coast (Fig. 3c). This region is predominantly desert, characterized by sand dunes and rocky outcrops in the Fish River Canyon–Great Karas region of Namibia. It has 82 species of woody plants with a high representation of narrow-ranged and evolutionarily distinct species – those with few or no close-living relatives – including *Aloe dichotoma* (Asphodelaceae), *Ceraria fruticulosa* (Portulacaceae), *Pachypodium namaquanum* (Apocynaceae) and *Stoeberia utilis* (Mesembryanthemaceae). The south-western portion of this phyloregion corresponds to the Gariep centre of endemism recognized by Nordenstam (1974). While Linder *et al.* (2005) lumped the Gariep centre of endemism

within the Namib-Karoo phytogeographical region based upon species distribution data, our results indicate that the two are phylogenetically distinct, although we only consider woody species, a relatively small fraction of the flora in these regions. We suggest that Gariep Karoo deserves to be included in systematic conservation actions targeted at representing the diversity as well as the richness of the southern African flora (DWAF, 2005).

We liken Phyloregion B to the 'Zambeian transition zone' that has affinities with Indian Ocean Coastal Belt, occurring in south-western Mozambique at the border between Zimbabwe and South Africa. It is characterized by 803 species of woody plants, including many narrow-ranged endemics such as *Combretum padoides* (Combretaceae), *Dovyalis hispidula* (Salicaceae), *Toddaliopsis bremekampii* (Rutaceae) and *Zanthoxylum humile* (Rutaceae) that contribute to the region's high species and phylogenetic diversity. The phyloregion also coincides with the boundaries of the Greater Limpopo Transfrontier Park (GLTP; Spenceley, 2006), a region that straddles several parks and game reserves (e.g. Kruger National Park in South Africa, Gonarezhou National Park in Zimbabwe, Zinave and Banhine National Parks and Coutada 16 Wildlife Utilisation Area in Mozambique; Wolmer, 2003). Although the delineation of the GLTP was based predominantly upon political and economic rather than ecological criteria (Wolmer, 2003), our study suggests that it also represents a distinct vegetation type.

Ecological processes acting as environmental filters (Keddy, 1992) have been key to many previous biogeographical classification schemes (e.g. Olson *et al.*, 2001), and we also found that phyloregions differed significantly in potential evapotranspiration, mean annual temperature, net primary productivity, mean annual precipitation and elevation, even when our analysis showed them to be phylogenetically close. For example, while Savanna and Miombo Woodland are phylogenetically similar, they differ notably in temperature, energy regime (potential evapotranspiration), precipitation and elevation. Our study highlights the important role of PET and MAT in structuring the distribution of phylogenetic beta diversity of woody plants, and complements previous work that has emphasized the importance of water-energy dynamics in determining patterns of species richness (O'Brien, 1993; O'Brien *et al.*, 1998). However, while our findings indicate that phyloregions might occupy different climate envelopes, climate alone cannot capture their complex evolutionary histories, including patterns of dispersal, speciation and extinction. In addition, past climate was very different from that observed today, and historical events such as the origination of the Benguela Upwelling System in the Miocene, following the separation of Antarctica from South America (Siesser, 1980; Dupont *et al.*, 2011), likely had a major impact on vegetation across the region.

We did not find a strong correlation with fire, despite suggestions that fire is a major force structuring the savanna-forest mosaic (Beerling & Osborne, 2006). However, historical fire occurrences in southern Africa are recently highly controlled, often for land management purposes (Roy *et al.*,

2005). Current fire regime might therefore be only poorly correlated with fire regimes across evolutionary time-scales, and perhaps our results showing little relationship between fire return frequency and the differentiation of phyloregions might not be unexpected.

Our taxon sampling represents the largest estimate of phylogeny for woody flora in the region (1400 species, across all vascular plant lineages). Woody plants (or their absence) are perhaps the most significant distinguishing feature of vegetation types in southern Africa. They are well represented in most plant families, and occur across a diverse range of habitat types. It would be interesting to explore whether herbs, grasses and other plant groups follow similar evolutionary patterns. However, while a more complete representation of the flora might allow a finer partitioning of vegetation types, we do not expect dramatic changes to the major phyloregions identified here (see e.g. Linder *et al.*, 2012).

We do not propose that our phylogenetic regionalization should replace existing classification schemes. In fact, these earlier schemes provide the foundations upon which our study rests, and the match between our phyloregions and currently recognized vegetation types is reasonably strong. Rather, our study highlights that woody plants in southern Africa can be grouped within distinct 'phyloregions' based upon their evolutionary affinities, and allows for the development of hypotheses about the origins and evolutionary histories of the woody flora within each region (e.g. the separation of Savanna and Miombo Woodland, and the evolutionary dynamics that define the Fynbos). In addition, using phylogeny we were able to distinguish two vegetation types with distinct evolutionary histories that are not well recognized by previous vegetation classification schemes. Our study thus reveals how comprehensive phylogenetic data can provide new insights into the spatial structure of biodiversity, and reveal hidden evolutionary affinities between vegetation types.

ACKNOWLEDGEMENTS

We acknowledge with gratitude the following institutions for financial and technical support: Government of Canada through Genome Canada and the Ontario Genomics Institute (2008-OGI-ICI-03), the International Development Research Centre (IDRC) Canada, the University of Johannesburg. BHD acknowledges financial support for this study from the South African National Research Foundation (NRF). Many individuals have provided very valuable and insightful comments on earlier drafts of the manuscript. Particular thanks go to Pieter Bester, Bezeng S. Bezeng, Ben G. Holt, Jean-Philippe Lessard and two anonymous referees.

REFERENCES

Angiosperm Phylogeny Group (2009) An update of the Angiosperm Phylogeny Group classification for the orders

- and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, **161**, 105–121.
- Archibald, S., Scholes, R.J., Roy, D.P., Roberts, G. & Boschetti, L. (2010) Southern African fire regimes as revealed by remote sensing. *International Journal of Wildland Fire*, **19**, 861–878.
- Archibald, S., Lehmann, C.E.R., Gómez-Dans, J.L. & Bradstock, R.A. (2013) Defining pyromes and global syndromes of fire regimes. *Proceedings of the National Academy of Sciences USA*, **110**, 6442–6447.
- Beerling, D.J. & Osborne, C.P. (2006) The origin of the savanna biome. *Global Change Biology*, **12**, 2023–2031.
- Boon, D. (2010) *Pooley's trees of eastern South Africa – a complete guide*. Flora and Fauna Publication Trust, Durban, South Africa.
- Borcard, D., Gillet, F. & Legendre, P. (2011) *Numerical ecology with R*. Springer, New York, USA.
- Brock, G., Pihur, V., Datta, S. & Datta, S. (2008) *clVALID*, an R package for cluster validation. *Journal of Statistical Software*, **25**, 1–22.
- 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.
- Clarke, K.R. (1993) Nonparametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, **18**, 117–143.
- Coates Palgrave, M. (2002) *Keith Coates Palgrave trees of southern Africa*, 3rd edn. Struik, Cape Town, South Africa.
- Davies, T.J., Savolainen, V., Chase, M.W., Goldblatt, P. & Barraclough, T.G. (2005) Environment, area and diversification in the species-rich flowering plant family Iridaceae. *The American Naturalist*, **166**, 418–425.
- De Candolle, A.P. (1820) *Géographie botanique. Dictionnaire des Sciences Naturelles* (ed. by F.G. Levrault), pp. 359–422. Strasbourg, France.
- Dupont, L.M., Linder, H.P., Rommerskirchen, F. & Schefuß, E. (2011) Climate-driven rampant speciation of the Cape flora. *Journal of Biogeography*, **38**, 1059–1068.
- DWAF (2005) *Systematic conservation planning for the forest biome of South Africa. Approach, methods and results of the selection of priority forests for conservation action*. Department of Water Affairs and Forestry, Pretoria, South Africa.
- ESRI (Environmental Systems Resource Institute) (2010) *ArcMap 10.0, ArcGIS Desktop*. ESRI Redlands, California, USA.
- Fraley, C. & Raftery, A.E. (2012) *Package 'mclust'*. Model-based clustering/normal mixture modeling. <http://cran.rproject.org/web/packages/mclust/mclust.pdf>
- 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.
- Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford, UK.
- Germishuizen, G. & Meyer, N.L. (2013) *Plants of southern Africa: an annotated checklist*. Strelitzia 14, National Botanical Institute, Pretoria, South Africa.
- Goldblatt, P. (1978) An analysis of the flora of southern Africa: its characteristics, relationships, and origins. *Annals of the Missouri Botanical Garden*, **65**, 369–436.
- Gower, J.C. (1983) Comparing classifications. *Numerical taxonomy* (ed. by J. Felsenstein), pp. 137–155. Springer, Berlin, Germany.
- Graham, C.H. & Fine, P.V.A. (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, **11**, 1265–1277.
- Hijmans, R.J. (2015) *raster: Geographic data analysis and modeling*. R package version 2.4-15. <http://CRAN.R-project.org/package=raster>
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Holt, B.G., Lessard, J.P., Borregaard, M.K., Fritz, S.A., Araújo, M.B., Dimitrov, D., Fabre, P.H., Graham, C.H., Graves, G.R., Jönsson, K.A., Nogués-Bravo, D., Wang, Z., Whittaker, R.J., Fjeldsø, J. & Rahbek, C. (2013) An update of Wallace's zoogeographic regions of the world. *Science*, **339**, 74–78.
- Hugo, S., Van Rensburg, B.J., Van Wyk, A.E. & Steenkamp, Y. (2012) Alien phytogeographic regions of southern Africa: numerical classification, possible drivers, and regional threats. *PLoS ONE*, **7**, e36269.
- Irish, J. (1994) The biomes of Namibia, as determined by objective categorisation. *Navorsinge van die Nasionale Museum Bloemfontein*, **10**, 549–592.
- Keddy, P.A. (1992) Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science*, **3**, 157–164.
- Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463–1464.
- Klopper, R.R., Chatelain, C., Bänninger, V., Habashi, C., Steyn, H.M., De Wet, B.C., Arnold, T.H., Gautier, L., Smith, G.F. & Spichiger, R. (2006). Checklist of the flowering plants of Sub-Saharan Africa. An index of accepted names and synonyms. South African Botanical Diversity Network Report No. 42. SABONET, Pretoria, South Africa.
- Koleff, P., Gaston, K.J. & Lennon, J.J. (2003) Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, **72**, 367–382.
- Kreft, H. & Jetz, W. (2010) A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, **37**, 2029–2053.
- Kubota, Y., Hirao, T., Fujii, S.-J., Shiono, T. & Kusumoto, B. (2014) Beta diversity of woody plants in the Japanese archipelago: the roles of geohistorical and ecological processes. *Journal of Biogeography*, **41**, 1267–1276.

- Kucharik, C.J., Foley, J.A., Delire, C., Fisher, V.A., Coe, M.T., Lenters, J.D., Young-Molling, C., Ramankutty, N., Norman, J.M. & Gower, S.T. (2000) Testing the performance of a dynamic global ecosystem model: water balance, carbon balance, and vegetation structure. *Global Biogeochemical Cycles*, **14**, 795–825.
- Linder, H.P. (2001) Plant diversity and endemism in sub-Saharan tropical Africa. *Journal of Biogeography*, **28**, 169–182.
- Linder, H.P. (2003) The radiation of the Cape flora, southern Africa. *Biological Reviews*, **78**, 597–638.
- Linder, H.P., Lovett, J.C., Mutke, J., Barthlott, W., Jürgens, N., Rebelo, T. & Küper, W. (2005) A numerical re-evaluation of the sub-Saharan phytochoria of mainland Africa. *Biologiske Skrifter*, **55**, 229–252.
- Linder, H.P., De Klerk, H.M., Born, J., Burgess, N.D., Fjeldsø, J. & Rahbek, C. (2012) The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. *Journal of Biogeography*, **39**, 1189–1205.
- Lomolino, M.V., Riddle, B.R., Whittaker, R.J. & Brown, J.H. (2010) *Biogeography*, 4th edn. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- Low, A.B. & Rebelo, A.G. (1996) *Vegetation of South Africa, Lesotho and Swaziland*. Department of Environmental Affairs and Tourism, Pretoria, South Africa.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2013) *cluster: cluster analysis basics and extensions*. R package version 1.14.4.
- Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209.
- Maurin, O., Davies, T.J., Burrows, J.E., Daru, B.H., Yesoufou, 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.
- Milligan, G.W. & Cooper, M.C. (1985) An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, **50**, 159–179.
- Mucina, L. & Rutherford, M.C. (2006) *The vegetation of South Africa, Lesotho and Swaziland*. Strelitzia 19. South African National Biodiversity Institute, Pretoria, South Africa.
- Nordenstam, B. (1974) The flora of the Brandberg. *Dinteria*, **11**, 3–67.
- 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.
- O’Brien, E.M., Whittaker, R.J. & Field, R. (1998) Climate and woody plant diversity in southern Africa: relationships at species, genus and family levels. *Ecography*, **21**, 495–509.
- O’Brien, E.M., Field, R. & Whittaker, R.J. (2000) Climatic gradients in woody plant (tree and shrub) diversity: water-energy dynamics, residual variation, and topography. *Oikos*, **89**, 588–600.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O’Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2015) *vegan: Community ecology package*. R package version 2.3-0. <http://CRAN.R-project.org/package=vegan>
- 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., Wetzel, W.W., Hedao, P. & Kassem, K.R. (2001) Terrestrial ecoregions of the worlds: a new map of life on Earth. *BioScience*, **51**, 933–938.
- 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**, 173–194.
- Paradis, E., Claude, J. & Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.
- 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.
- Procheş, S., Wilson, J.R.U. & Cowling, R.M. (2006) How much evolutionary history in a 10 × 10 m plot? *Proceedings of the Royal Society B: Biological Sciences*, **273**, 1143–1148.
- Quinn, R.M., Gaston, K.J. & Arnold, H.R. (1996) Relative measures of geographic range size: empirical comparisons. *Oecologia*, **107**, 179–188.
- R Core Team (2013) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Revolution Analytics & Weston, S. (2013) *doSNOW: Foreach parallel adaptor for the snow package*. R package version 1.0.12 <http://CRAN.R-project.org/package=doSNOW>
- Roy, D.P., Frost, P., Justice, C., Landmann, T., Le Roux, J., Gumbo, K., Makungwa, S., Dunham, K., Du Toit, R., Mhwandagara, K., Zacarias, A., Tacheba, B., Dube, O., Pereira, J., Mushove, P., Morissette, J., Santhana Vannan, S. & Davies, D. (2005) The southern Africa fire network (SAFNet) regional burned area product validation protocol. *International Journal of Remote Sensing*, **44**, 4265–4292.
- Rutherford, M.C. (1997) Categorization of biomes. *Vegetation of southern Africa* (ed. by R.M. Cowling, D.M. Richardson and S.M. Pierce), pp. 91–98. Cambridge University Press, Cambridge, UK.
- 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, Piscataway, New Jersey, USA.
- Scarpella, E. & Meijer, A.H. (2004) Pattern formation in the vascular system of monocot and dicot plant species. *New Phytologist*, **164**, 209–242.

- Schmidt, E., Lotter, M. & McClelland, W. (2007) *Trees and shrubs of Mpumalanga and Kruger National Park*. Jacana Education, Johannesburg, South Africa.
- Scott, L., Anderson, H.M. & Anderson, J.M. (1997) Vegetation history. *Vegetation of southern Africa* (ed. by R.M. Cowling, D.M. Richardson and S.M. Pierce), pp. 62–84. Cambridge University Press, Cambridge, UK.
- 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.
- Spenceley, A. (2006) Tourism in the Great Limpopo Transfrontier Park. *Development Southern Africa*, **23**, 649–667.
- Steenkamp, Y., Van Wyk, A.E., Smith, G.F. & Steyn, H. (2005) Floristic endemism in southern Africa: a numerical classification at generic level. *Biologiske Skrifter*, **55**, 253–271.
- Thorntwaite, C.W. (1948) An approach toward a rational classification of climate. *Geographical Review*, **38**, 55–94.
- Van Rooy, J. & Van Wyk, A.E. (2010) The bryofloristic regions of southern Africa. *Journal of Bryology*, **32**, 80–91.
- Van Wyk, A.E. & Smith, G.F. (2001) *Regions of floristic endemism in southern Africa. A review with emphasis on succulents*. Umdaus Press, Cape Town, South Africa.
- Van Wyk, B., dervan Berg, E., Coates Palgrave, M. & Jordaan, M. (2011) *Dictionary of names for southern African trees*, 1st edn. Briza, Pretoria, South Africa.
- Verboom, G.A., Linder, H.P., Stock, W.D. & Baum, D. (2003) Phylogenetics of the grass genus *Ehrharta*: evidence for radiation in the summer-arid zone of the South African cape. *Evolution*, **57**, 1008–1021.
- White, F. (1983) *The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa Natural Resources Research*, 20. UNESCO, Paris, France.
- Wiens, J.J. (2004) Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution*, **58**, 193–197.
- Willmott, C.J. & Matsuura, K. (2001) *Terrestrial water budget data archive: monthly time series (1950–1999)*. Available at: http://climate.geog.udel.edu/~climate/html_pages/README.wb_ts2.html.
- Wolmer, W. (2003) *Transboundary conservation: the politics of ecological integrity in the Great Limpopo Transfrontier Park*. Sustainable livelihoods in Southern Africa Research Paper no. 4, Institute of Development Studies, Brighton.
- 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:

Appendix S1 Summary review (S1), supporting tables (Tables S1–S3) and figures (Figs S1–S4).

BIOSKETCH

Barnabas H. Daru is a Lecturer in the Department of Plant Science at University of Pretoria, South Africa. His research integrates ecology, biogeography and conservation biology, mostly using phylogenetic comparative methods and spatial statistical modelling.

Author contributions: B.H.D., M.V.D.B., and T.J.D. designed the research; B.H.D. performed the analysis; O.M., K.Y.; H.S., and J.S. contributed analytic tools; B.H.D. and T.J.D. wrote the paper with contributions from all co-authors.

Editor: Şerban Procheş