



# Environmental and historical controls of floristic composition across the South American Dry Diagonal

Danilo M. Neves<sup>1,2\*</sup>, Kyle G. Dexter<sup>1,3</sup>, R. Toby Pennington<sup>1</sup>, Marcelo L. Bueno<sup>2</sup> and Ary T. Oliveira Filho<sup>2</sup>

<sup>1</sup>Royal Botanic Garden Edinburgh, Edinburgh, Midlothian EH3 5LR, UK, <sup>2</sup>Programa de Pós-Graduação em Biologia Vegetal, Universidade Federal de Minas Gerais – UFMG, Campus Pampulha, Belo Horizonte, MG 31270–090, Brazil, <sup>3</sup>School of GeoSciences, University of Edinburgh, King's Buildings, Edinburgh, Midlothian EH9 3JN, UK

## ABSTRACT

**Aim** The aim of this study was to test the role of environmental factors and spatially autocorrelated processes, such as historical fragmentation and dispersal limitation, in driving floristic variation across seasonally dry tropical forests (SDTFs) in eastern South America.

**Location** SDTFs extending from the Caatinga phytogeographical domain of north-eastern Brazil to the Chaco phytogeographical domain of northern Argentina, an area referred to as the Dry Diagonal.

**Methods** We compiled a database of 282 inventories of woody vegetation in SDTFs from across the Dry Diagonal and combined this with data for 14 environmental variables. We assessed the relative contribution of spatially autocorrelated processes and environmental factors to the floristic turnover among SDTFs across the Dry Diagonal using variation partitioning methods. In addition, we used multivariate analyses to determine which environmental factors were most important in explaining the turnover.

**Results** We found that the environmental factors measured (temperature, precipitation and edaphic conditions) explained 21.3% of the variation in species composition, with 14.1% of this occurring independently of spatial autocorrelation. A spatially structured fraction of 4.2% could not be accounted for by the environmental factors measured. The main axis of compositional variation was significantly correlated with a north–south gradient in temperature regime. At the extreme south of the Dry Diagonal, a cold temperature regime, in which frost occurred, appeared to underlie floristic similarities between chaco woodlands and southern SDTFs.

**Main conclusions** Environmental variables, particularly those related to temperature regime, seem to be the most significant factors affecting variation in species composition of SDTFs. Thus historical fragmentation and isolation alone cannot explain the turnover in species composition within SDTFs, as is often assumed. Compositional and environmental heterogeneity needs to be taken into account both to understand the past distribution of SDTFs and to manage and conserve this key tropical biome.

## Keywords

caatinga woodlands, chaco woodlands, environmental niche, Pleistocene Arc Hypothesis, seasonally dry tropical forests, species turnover, variation partitioning.

\*Correspondence: Danilo M. Neves, Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh, Midlothian EH3 5LR, UK. E-mail: danilormn@gmail.com

## INTRODUCTION

Seasonally dry tropical forests (SDTFs) occur as fragments of varying size throughout Latin America, from northern Argentina and north-east Brazil to north-west Mexico (Pennington *et al.*, 2006). These SDTFs comprise a recognized biome (Pennington *et al.*, 2000, 2009) that is characterized by  $< 1600$  mm precipitation year<sup>-1</sup>, a marked dry season of 5–6 months and high rates of deciduousness during the dry season (Bullock *et al.*, 1995; Gentry, 1995; Oliveira-Filho *et al.*, 2006). SDTFs have received increased attention in recent years because of both their highly threatened status (e.g. Mooney *et al.*, 1995; Pennington *et al.*, 2006, 2009; Dirzo *et al.*, 2011) and the influential Pleistocene Arc Hypothesis (PAH; Prado & Gibbs, 1993; Pennington *et al.*, 2000). The PAH postulates that SDTFs had a more widespread and contiguous distribution in South America during Pleistocene glacial phases, including areas that are currently covered by rain forest or savanna. The PAH is based upon the widespread distributions of multiple, unrelated tree species found in disjunct SDTF areas from north-east Brazil through central and southern Brazil to Paraguay and Argentina and, in some cases, as far as the dry inter-Andean valleys of western South America. It is suggested that the putative Pleistocene Arc, a more continuous area of SDTF, connected these areas during the cooler and drier climates of Pleistocene glacial phases (Prado & Gibbs, 1993; Pennington *et al.*, 2000).

Subsequent research, however, has shown that these widespread species represent a small minority of SDTF tree species and that floristic variation among SDTF sites across the Neotropics is high (e.g. Gillespie *et al.*, 2000; Trejo & Dirzo, 2002; Linares-Palomino, 2006; Lott & Atkinson, 2006; Pennington *et al.*, 2006; Queiroz, 2006; Castillo-Campos *et al.*, 2008; Linares-Palomino *et al.*, 2011). For instance, Linares-Palomino *et al.* (2011) found that less than 2% of 3839 SDTF tree species are found in 10 or more of 21 disjunct SDTF nuclei across the Neotropics. On the other hand, the same authors (Linares-Palomino *et al.*, 2011) found higher floristic similarity (Sørensen similarity index values  $> 0.25$ ) among SDTF nuclei in Brazil and southern South America and suggested that this offers some support for the idea that SDTF was more widespread and continuous within this region during Pleistocene glacial phases.

When explaining the disjunct distributions of SDTF species and floristic variation of SDTFs in general, recent studies have focused primarily on the role of historical biogeographical processes, namely habitat fragmentation and dispersal limitation (e.g. Pennington *et al.*, 2006, 2009). However, environmental conditions are also expected to affect species distribution. This study is the first to examine the role of environmental, primarily climatic, determinants in shaping the distribution of tree species across a large expanse of Neotropical SDTF sites.

We aimed to test the role of historical fragmentation and dispersal limitation versus environmental factors in driving floristic differentiation among SDTFs across eastern South

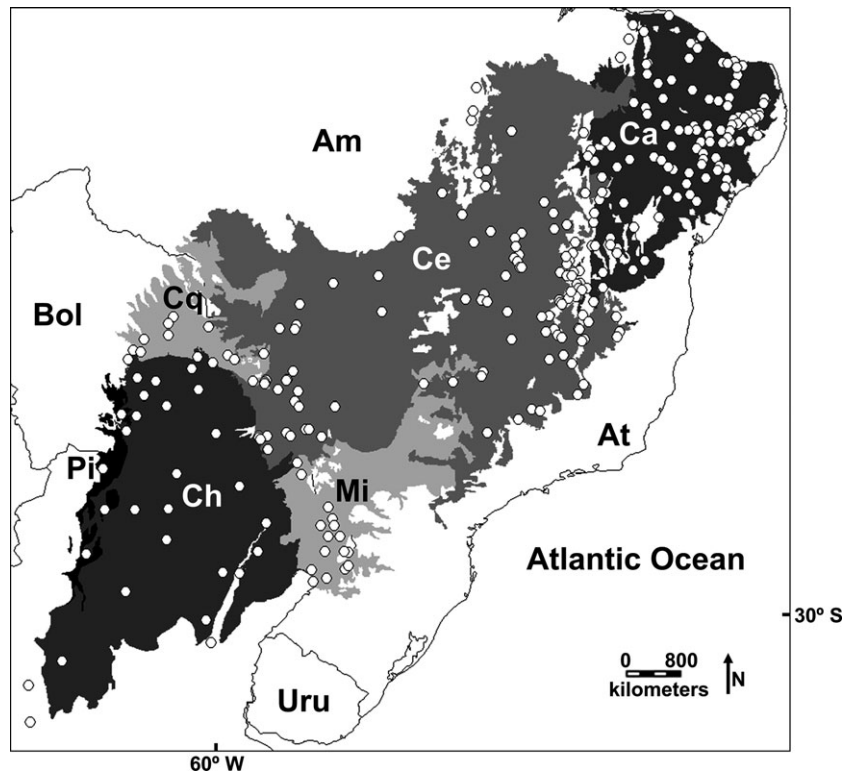
America. Understanding the roles of these two processes is an important and unsolved issue in many systems (e.g. in the Amazon: Haffer, 1969; ter Steege *et al.*, 2006; in the cerrado vegetation: Werneck *et al.*, 2012a; in South American subtropical forests: Oliveira-Filho *et al.*, 2013). Variation partitioning methods provide a means of assessing the relative contribution of measured environmental factors and spatially autocorrelated processes (potentially including unmeasured environmental factors) in driving species turnover (Legendre *et al.*, 2012). This approach has been used mostly at small spatial scales (e.g. within plots; Legendre *et al.*, 2009) to address the influence of steady-state, small-scale dispersal limitation. However, it can also be used at much broader spatial scales (Vetaas & Ferrer-Castán, 2008) to address habitat fragmentation and isolation at historical time-scales, and may be particularly useful in ecosystems with limited palaeo-ecological data, such as SDTFs.

We compiled a database of SDTF floristic inventories from sites extending from the caatinga woodlands of north-eastern Brazil to the chaco woodlands of northern Argentina, an area referred to as the Dry Diagonal (following Prado & Gibbs, 1993; also see Vanzolini, 1963; Werneck *et al.*, 2012b). This region was particularly appropriate because it was the main area studied by Prado & Gibbs (1993) in the development of the PAH and because it has been the target of previous studies on SDTF biogeography (de Andrade-Lima, 1954, 1982; Ratter *et al.*, 1988; Caetano & Naciri, 2011; Werneck *et al.*, 2011; Collevatti *et al.*, 2012, 2013a,b). If the floristic variation of SDTFs is solely the result of fragmentation and subsequent dispersal limitation between SDTF patches, climatic and edaphic variables should not be important factors in explaining the variation in community composition. Alternatively, if previously unrecognized environmental conditions within SDTFs are the principal factors driving floristic turnover, then climatic and edaphic factors should explain a large proportion of the variation in community composition.

## MATERIALS AND METHODS

### Study area

The Dry Diagonal stretches from north-eastern Brazil to Argentina, Paraguay and Bolivia, and occurs primarily in three phytogeographical domains (IBGE, 1993) that are named after the main vegetation types that occur within them: Caatinga (tropical semi-arid thorn woodlands, called caatinga woodlands in Brazil), Cerrado (seasonal woody savannas, called cerrado woodlands in Brazil) and Chaco (subtropical/tropical semi-arid thorn woodlands, called chaco woodlands in Brazil and elsewhere) (Fig. 1). Prado & Gibbs (1993) defined three major nuclei of SDTFs in the Dry Diagonal: (1) the Caatinga nucleus, which occurs in north-eastern Brazil and is the largest nucleus (essentially the same area as the Caatinga phytogeographical domain); (2) the Misiones nucleus, which extends from the lower course of the Paraguay and Paraná rivers to the upper Uruguay River in Misiones,



**Figure 1** Map showing the location of tree species surveys and major vegetation types used for the study of environmental and historical controls of floristic composition across the South American Dry Diagonal. Circles denote the location of tree species surveys used in this study ( $n = 282$ ). The phytogeographical domains that form the South American Dry Diagonal are shaded (Ca, Caatinga; Ce, Cerrado; Ch, Chaco; Cq, Chiquitania; Pi, Piedmont; Mi, Misiones) while moist forest phytogeographical domains are white (Am, Amazon rain forest; At, Atlantic rain forest). The phytogeographical domains are named after the major vegetation type that occurs within them (e.g. chaco woodlands, caatinga woodlands and woody savannas; the latter is called cerrado woodlands in Brazil). Note that each phytogeographical domain may contain several additional vegetation types (e.g. the Cerrado phytogeographical domain contains SDTF and gallery forest in addition to the main vegetation type of savanna).

Argentina; and (3) the Piedmont nucleus, which occurs along the eastern base of the Andes in north-eastern Argentina. SDTFs also occur in the Chiquitania region of Bolivia (Linares-Palomino *et al.*, 2011) and in patches scattered throughout the Cerrado phytogeographical domain (Ratter *et al.*, 1988; Pennington *et al.*, 2000) in areas of fertile soil.

One environmental element in the definition of SDTFs is that they are a frost-free vegetation (Prado, 1993a,b; Pennington *et al.*, 2000), which would suggest that the chaco woodlands, found in northern Argentina, western Paraguay, south-eastern Bolivia and the extreme western edge of Mato Grosso do Sul state in Brazil, are not SDTFs because they suffer regular winter frost. However, large areas delimited as SDTF in the schematic maps of previous studies (e.g. Piedmont and Misiones; Prado & Gibbs, 1993; Oliveira-Filho *et al.*, 2006; Pennington *et al.*, 2009) also receive significant frost (see Results). Thus we chose to include the chaco woodlands in our analyses because they meet other criteria of SDTF vegetation as defined by the same authors (Prado, 1993a,b; Pennington *et al.*, 2000), such as seasonality of rainfall and deciduousness.

### The SDTF database

We extracted floristic inventory data from the NeoTropTree database (available at <http://prof.icb.ufmg.br/treetatlan>). This database consists of > 2000 georeferenced localities for which lists of tree species (trees defined here as woody plants > 3 m in height) were compiled from an extensive survey of published and unpublished literature. A NeoTropTree site is defined by a single vegetation type contained within a circular area with a 5-km radius. Therefore, where two or more vegetation types co-occur in one 10-km area, there may be two geographically overlapping sites in the NeoTropTree database, each for a distinct vegetation type. The NeoTropTree data were derived primarily from inventories of woody vegetation (i.e. plots, transects and vegetation surveys). Surveys of specimens at major herbaria were then used to augment the lists for each NeoTropTree site with any tree species that were collected within the 5-km radius of the original NeoTropTree site and within the same vegetation type. Moreover, all NeoTropTree species were checked regarding their taxonomy and synonymies by using the Flora do Brasil

database for plants that occur in Brazil, and the Tropicos database for plants that do not occur in Brazil (available at <http://floradobrasil.jbrj.gov.br/> and <http://tropicos.org/>, respectively). NeoTropTree does not include occurrence records without an indication or evidence of vegetation type, nor sites with very incomplete species lists. The latter is an important filter because different sampling efforts across sites may bias their descriptive power.

The 282 NeoTropTree inventories used in this study comprised 134 SDTF sites from the Caatinga phytogeographical domain, 14 from the Misiones nucleus, five from the Piedmont nucleus, 91 from SDTF enclaves scattered throughout the Cerrado phytogeographical domain, six from the Bolivian Chiquitania and 32 chaco woodland sites from the Chaco phytogeographical domain. This does not represent an even geographical coverage of inventories, but we believe our results are not excessively biased because nuclei covered by just a few sites often had many species, whereas nuclei with many sites could only have a few species. For example, the Misiones nucleus was represented by 398 species from just 14 lists, whereas 891 species were recorded from 134 lists from the Caatinga phytogeographical domain. The final species matrix contained presence–absence data for 1765 tree species across the 282 SDTF sites, with a total of 25,650 presences (Fig. 1).

The NeoTropTree database also included environmental data for all its sites, derived from multiple sources. Eight variables were related to edaphic substrate: mean, minimum and maximum monthly soil moisture (%), obtained from the International Soil Moisture Network (<http://www.ipf.tuwien.ac.at/insitu/>); prevailing slope and aspect, obtained from CGIAR-CSI (2006); and soil fertility (% base saturation), soil coarseness (% sand) and substrate rockiness (% surface of rock outcrops), extracted from a detailed map of soil types produced by EMBRAPA & IBGE (2003) using the RadamBrasil Soil Survey of the 1970s and 1980s. Climatic variables included the mean duration (days) and severity (mm) of water-deficit periods, both extracted from Walter diagrams (Walter, 1985); the 19 bioclimatic variables produced by WorldClim 1.4, a high-resolution (1-km) database of global climate layers created by Hijmans *et al.* (2005); mean frequency of frosts (days), cloud cover (%) and cloud interception (mm), obtained from gridded datasets produced by Jones & Harris (2008); and three additional variables, potential and actual evapotranspiration (mm) and an aridity index, derived from WorldClim by Zomer *et al.* (2007, 2008). A full description with details of the protocols for NeoTropTree can be found at <http://prof.ic-b.ufmg.br/treetlan>.

### Analyses of species composition

We first explored the patterns of similarity in community composition by conducting a hierarchical clustering analysis of all 282 sites. We used Jaccard distance as the dissimilarity metric and unweighted paired groups as the linkage method

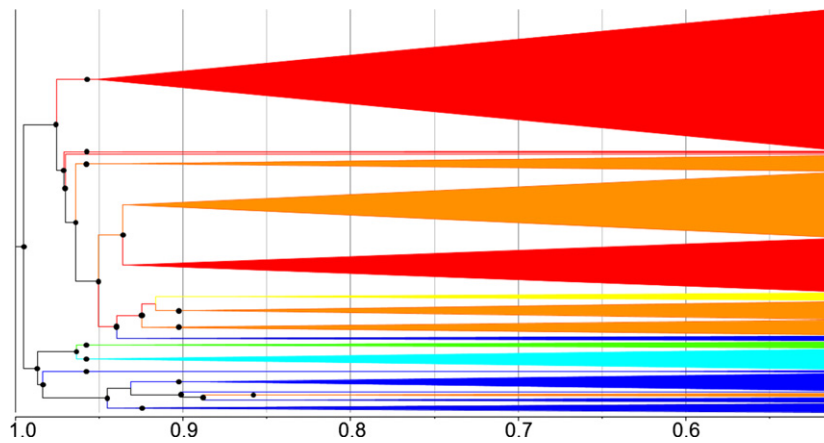
(McCune & Grace, 2002). We assessed the confidence for each node of the hierarchical cluster using multi-scale bootstrapping (Shimodaira, 2004) in the *PVCLUST* package (Suzuki & Shimodaira, 2011) in the R statistical environment (R Core Team, 2015). Because singletons (species found at a single site) commonly increase the intrinsic noise in most analyses without contributing information (Lepš & Šmilauer, 2003), we excluded the 318 singleton species from this and all subsequent analyses.

We obtained the relative contribution of measured environmental factors and unmeasured spatially autocorrelated factors in explaining variation in community composition by following methods similar to those proposed by Dray *et al.* (2012) and Legendre *et al.* (2012). This routine comprises (1) the compilation of significant spatial and environmental variables through a forward selection method for canonical correspondence analysis (CCA), with a permutation-based test for each variable added (Borcard *et al.*, 2011); (2) an additional and progressive elimination of collinear variables based on their variance inflation factor (VIF) until maintaining only those with VIF < 10 (Quinn & Keough, 2002); and (3) variation partitioning of the community composition matrix with respect to the significant spatial and environmental variables. As spatial variables, we used Moran's eigenvector maps (MEMs), which represent the diagonalization of a centred spatial weighting matrix (Dray *et al.*, 2012). We tested the overall significance of the spatial fraction (controlled for measured environmental variation) and the environmental fraction (controlled for spatial autocorrelation) by applying a permutation test (999 permutations) for CCA. All variation partitioning analyses were conducted using the *SPACEMAKER* (Dray, 2010), *SPDEP* (Bivand, 2012), *TRIPACK* (Gebhardt, 2009) and *VEGAN* (Oksanen *et al.*, 2012) packages in the R statistical environment (R Core Team, 2015).

We also performed a detrended correspondence analysis (DCA) to derive orthogonal eigenvectors representing gradients in floristic variation and used post hoc correlations to assess their relationship with environmental factors. We tested the significance of the post hoc correlation coefficients by applying a permutation test (999 permutations). There was a negligible increase in explained variation with the addition of > 2 DCA axes, and we therefore focused on the first two axes for comparisons with environmental variables. We also explored the results visually by plotting the site scores of the DCA axes and the values of the environmental variables in geographical space. These analyses were performed in the statistical package *VEGAN* (Oksanen *et al.*, 2012) and the maps were designed using the package *MAPTOOLS* (Lewin-Koh & Bivand, 2012), both in the R statistical environment.

### RESULTS

The hierarchical clustering results were largely congruent with the previously defined SDTF nuclei. The classification



**Figure 2** Hierarchical clustering of seasonally dry tropical forests (SDTFs) and chaco woodlands across the South American Dry Diagonal based on their tree species composition. The dissimilarity measure and linkage methods used were Jaccard and unweighted group average, respectively. Black bullets at the base of clusters indicate approximately unbiased bootstrap values  $\geq 0.90$  (Suzuki & Shimodaira, 2011). The size of the triangles is proportional to the number of sites, and the floristic nuclei are discriminated by different colours: red, 134 caatinga woodlands; orange, 91 Cerrado SDTFs; yellow, six SDTFs from Chiquitania; green, five SDTFs from Piedmont; light blue, 14 SDTFs from Misiones; dark blue, 32 chaco woodlands.

segregated the 282 sites into two principal groups, hereafter termed the northern and southern groups. The northern group consisted of 232 sites from the Caatinga nucleus, Cerrado phytogeographical domain SDTFs and Bolivian Chiquitania, whereas the southern group consisted of 50 sites from the Misiones and Piedmont nuclei and the chaco woodlands from the Chaco phytogeographical domain (Fig. 2).

The forward selection procedure retained 14 out of 36 environmental variables for modelling variation in species composition (Table 1). After partitioning the variation explained by the selected spatial and environmental predictors, we found that the measured environmental factors explained 21.3% of the variation in species composition, with 14.1% of this occurring independently of spatial autocorrelation. A spatially structured fraction of 4.2% could not be accounted for by the measured environmental factors; 74.5% of the variation remained unexplained (Fig. 3).

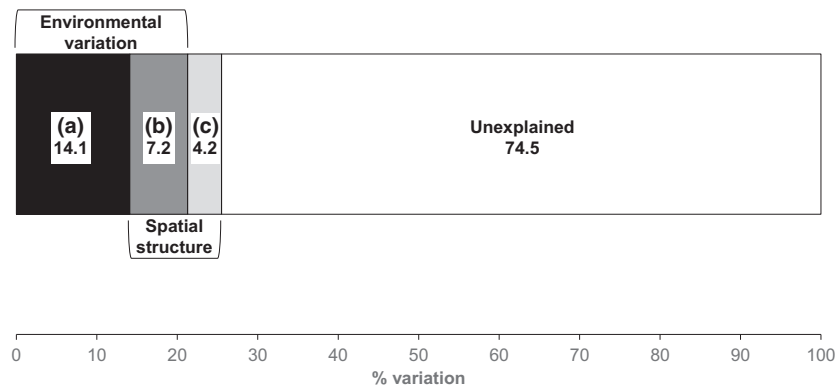
When investigating which environmental variables explained species turnover across the Dry Diagonal, we found that the variation in species composition summarized by the first DCA axis revealed a large-scale gradient (Fig. 4a) that was correlated with a north–south gradient in the mean minimum temperature of the coldest month (Fig. 4c, Table 1). The first DCA axis was also congruent with a north–south segregation in days of frost (dashed black line in Fig. 4a). The mean minimum of the coldest month in southern SDTFs ranged from 0.1 °C to 15.4 °C, while some southern SDTFs, such as those from the Piedmont nucleus, could experience up to 18 days of frost. In the northern group, four sites experienced up to 3 days of frost, while the other 228 sites experienced none. The first DCA axis was also correlated with a north–south gradient in isothermality (Fig. 4e, Table 1), suggesting that the high seasonality, including both hot summers and winter frosts in the south-

**Table 1** Environmental variables selected for the study of environmental controls of floristic composition across the South American Dry Diagonal. The variables shown are ordered by the post hoc correlation coefficient with the first detrended correspondence analysis (DCA) axis. The post hoc correlation coefficient of each environmental variable with the second DCA axis is also shown. VIF, variance inflation factor, obtained using the  $r$ -squared value of the regression of one variable against all other explanatory variables; n.s., not significant;  $P < 0.05$ , except for n.s. correlation coefficients.

Variables	DCA 1	DCA 2	VIF
Isothermality	0.62	0.06	7.64
Minimum temperature of the coldest month	0.52	n.s.	8.43
Severity of water deficit	0.48	0.02	8.87
Precipitation of the driest month	0.34	0.03	7.57
Cloud cover	0.32	0.12	4.08
Sediment coarseness	0.31	n.s.	4.85
Days of frost	0.21	n.s.	3.08
Soil fertility	0.13	0.02	5.43
Soil water storage capacity	0.09	0.44	9.29
Precipitation of the coldest quarter	0.03	0.04	3.74
Precipitation of the wettest month	0.02	0.43	7.47
Elevation	0.01	0.04	3.69
Surface rockiness	0.01	n.s.	2.02
Cloud interception	0.01	0.02	1.45

ern group, was a relevant factor driving variation in community composition.

The community composition variation summarized by the second DCA axis revealed a rainfall gradient that was most strongly correlated with the precipitation of the wettest month (Fig. 4d, Table 1). In fact, the species composition variation summarized by the second DCA axis was congruent



**Figure 3** Variation partitioning by canonical correspondence analysis to determine how much of the spatial variation in floristic composition across the South American Dry Diagonal was accounted for by the environmental variables measured. Fraction (a) represents non-spatial environmental variation, fraction (b) spatially structured environmental variation, i.e. the overlap between environmental and spatial components, and fraction (c) the spatial structure not explained by the measured environmental variables. All values are expressed as percentages.

with most of the precipitation-related variables (Table 1), segregating a dry group (Caatinga nucleus, chaco woodlands, Piedmont nucleus and Bolivian Chiquitania) from a wet group (Cerrado phytogeographical domain SDTFs and Misiones nucleus). Furthermore, the second DCA axis correlated with maximum soil water storage capacity (Fig. 4f), suggesting an interaction between rainfall and soil characteristics in driving the turnover in species composition summarized by this axis.

## DISCUSSION

Previous large-scale floristic analyses of South American SDTFs have focused on historical fragmentation and dispersal limitation when explaining patterns of species turnover amongst communities (e.g. Pennington *et al.*, 2006, 2009; Linares-Palomino *et al.*, 2011). Our analyses indicate that climatic and edaphic conditions also merit attention. Because of the strong environmental correlations we found (Table 1), it is inappropriate to designate historical fragmentation and dispersal limitation as the sole factors shaping species composition in SDTFs. Most of the variation in species composition that we were able to explain was assigned to environmental factors that were independent of spatial autocorrelation (14.1%). On the other hand, another 7.2% of the variation was accounted for by spatially structured environmental variation, signifying that the predictions made by spatial and niche-based models converged. However, considering that correlations between community composition and environmental variables are known to be important to plants (discussed below), we believe that it is more parsimonious to attribute most of the spatially structured environmental variation to niche-based controls (cf. Legendre *et al.*, 2009).

At a broad scale within the South American tropics, precipitation regime is likely to be a primary determinant of the presence versus absence of an SDTF; at a much smaller spatial scale (within north-eastern and central-western Brazil),

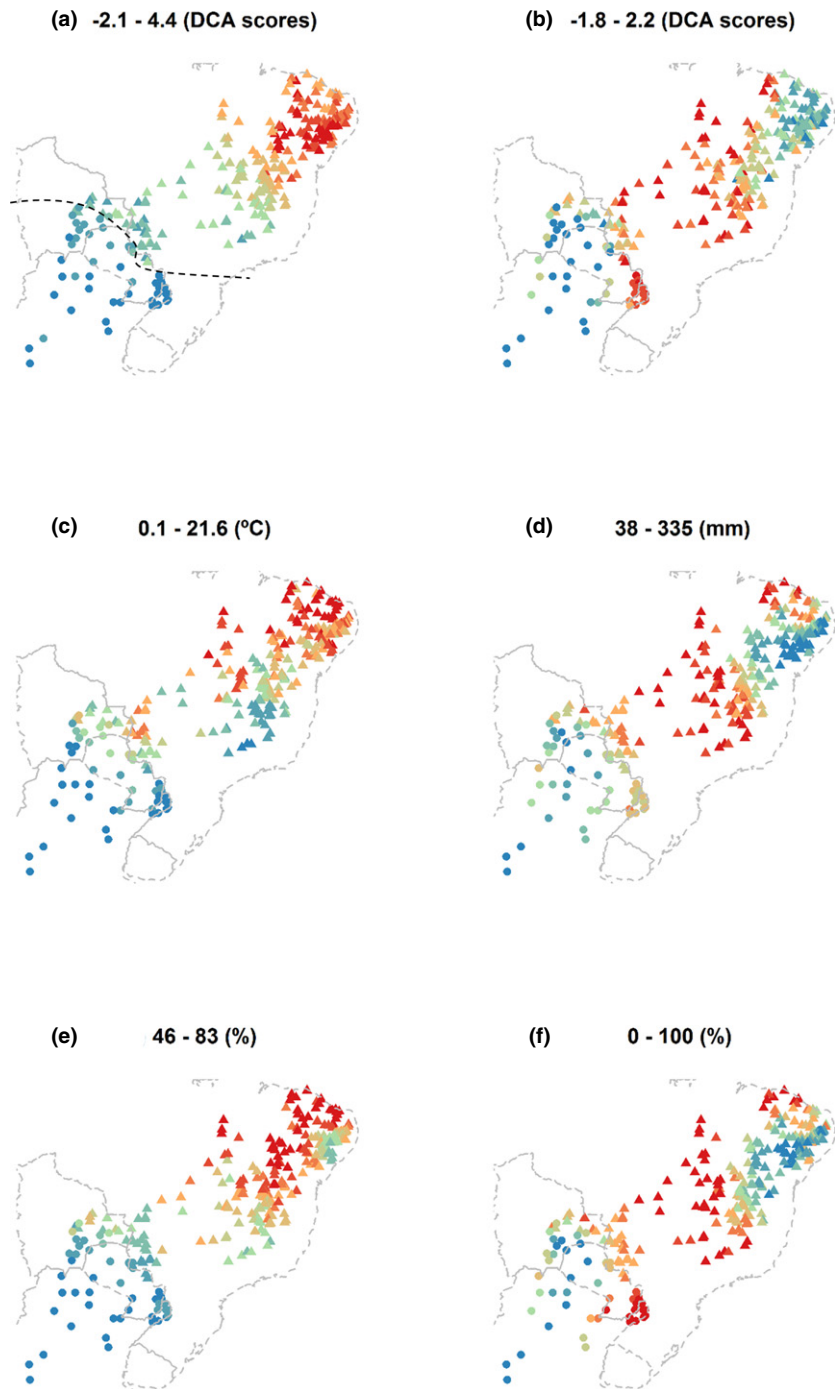
Santos *et al.* (2012) found that precipitation regime was the chief factor determining the species composition of SDTF communities. Nevertheless, in analysing SDTF communities at a larger spatial scale across the entire Dry Diagonal, we found temperature-related variables to be the most significant factors affecting variation in species composition within SDTFs and chaco woodlands. In agreement with this result, the fundamental division in our clustering analysis was between two geographically segregated groups, with the northern group experiencing mild temperatures in the coldest month, no frost (except for three sites with  $\leq 3$  days frost year<sup>-1</sup>) and a low annual temperature range compared with the southern group.

This fundamental latitudinal division reflects marked differences in the frequencies of species from different families. Species in the Leguminosae, Bignoniaceae, Moraceae and Annonaceae are more frequent in the northern group, which indicates that the northern flora primarily comprise species incapable of coping with freezing temperatures. Punyasena *et al.* (2008) classified these families as temperature-sensitive because their abundance and richness decrease in low temperatures. Meanwhile, Lauraceae, Asteraceae and Melastomataceae have higher frequencies in the cooler southern group.

Rainfall regime correlated strongly with the second gradient of species composition, as summarized by the second DCA axis. The precipitation division reflects the fact that the relative frequencies of Leguminosae, Capparaceae and Polygonaceae are higher in drier areas, whereas the relative frequencies of Rubiaceae, Moraceae, Arecaceae and Annonaceae are higher in wetter areas. The lower frequencies of the latter families in the drier areas match the results of Punyasena *et al.* (2008), who showed that their abundance and richness decrease in areas with marked drought.

### The distinctiveness of the chaco woodlands

Previous authors have excluded the chaco woodlands from the SDTF biome, citing floristic differences, with the expla-



**Figure 4** Geographical variation in seasonally dry tropical forest (SDTF) and chaco woodland characteristics across the South American Dry Diagonal. The northern group from the hierarchical clustering analysis is discriminated by triangles and the southern group by bullets (see Figs 1 and 2 and Results). Values are illustrated by the colour of the symbols: warmer colours indicate higher values. Numbers in the header for each panel indicate minimum and maximum values. (a) Scores from the first axis of the detrended correspondence analysis (DCA) analysis of species community composition. The black dashed line delimits frost-free (northern SDTFs) and frost-affected (southern SDTFs and chaco woodlands) areas; (b) scores from the second DCA axis; (c) mean minimum temperature of the coldest month; (d) precipitation of the wettest month; (e) isothermality; and (f) maximum soil water storage capacity.

nation that SDTF occurs only in frost-free areas (Prado, 1993a,b; Pennington *et al.*, 2000; Werneck *et al.*, 2011). However, we found that the sites we analysed in the Chaco domain were compositionally similar to southern SDTFs (Fig. 4a; additionally some Chaco domain sites clustered with SDTFs from the Cerrado domain and vice versa in Fig. 2). In fact, the chaco woodlands from our dataset shared 82% of their species with at least one of the SDTF nuclei of the Dry Diagonal. These results, combined with the finding that many sites considered as SDTF experience significant

frost (e.g. in the Piedmont and Misiones nuclei), suggest that the chaco woodlands may not be abruptly distinct from SDTFs and may represent one extreme of a floristic gradient driven primarily by temperature. In support of the idea of chaco woodlands being at one end of an SDTF temperature gradient, a previous study (Sarkinen *et al.*, 2011) found that 8.4% of SDTF specialist species occurred in the Chaco phytogeographical domain. However, further analyses of sites across the Chaco domain and surrounding areas are needed to address its distinctness versus similarity compared with

Neotropical SDTFs, in particular analyses that focus on the potentially distinct edaphic conditions often present in the Chaco phytogeographical domain (i.e. hypersaline soils) (Prado, 1993a).

### Affinities of the arboreal caatinga

Within northern SDTFs, Santos *et al.* (2012) classified the arboreal caatinga (sites that occur along the border with the Cerrado phytogeographical domain) as part of the Caatinga nucleus. However, performing the hierarchical clustering with a larger database we found that the arboreal caatinga (shown as red within the orange Cerrado phytogeographical domain cluster in Fig. 2) is more similar to SDTFs found in the Cerrado phytogeographical domain (support value  $\geq 0.9$  in Fig. 2), which could be related to edaphic similarities between arboreal caatinga and the the Cerrado phytogeographical domain SDTFs (calcareous, rich soils). The soil fertility (measured as a percentage of base saturation), for instance, of the arboreal caatinga and Cerrado phytogeographical domain SDTFs from our database were  $57 \pm 2$  (mean  $\pm$  SD) and  $46 \pm 13$ , respectively, whereas other caatinga woodlands had poorer soils ( $35 \pm 14$ ).

### Unexplained variation

While the environmental correlates of species composition in SDTF were fairly clear and straightforward to interpret, we still failed to explain a large fraction (74.5%) of the variation in species composition. There are many potentially important factors in determining the species composition of assemblages that we have not accounted for adequately, such as: (1) ecological drift (cf. Hubbell, 2001) driving stochastic rearrangements of species distribution ranges through time; (2) biotic processes that were not measured (e.g. competition and natural enemies); (3) false absences in the surveys (e.g. researcher error and misidentification of species); (4) unmeasured descriptors of habitat quality (e.g. succession stage and degree of land degradation); (5) unsaturation, i.e. when species do not occupy all patches of suitable habitat (Titeux *et al.*, 2004); and (6) other spatially and non-spatially structured environmental factors that were not measured. Moreover, a high proportion of unexplained variation, ranging from about 33% to 75% (e.g. Legendre *et al.*, 2009; Oliveira-Filho *et al.*, 2013; reviewed in Soininen, 2014), is a common outcome in studies of floristic composition over similar spatial scales, as statistical noise is very high in analyses with species presence-absence data (ter Braak, 1986; Guisan *et al.*, 1999).

### Concluding remarks

Our demonstration that the floristic composition of SDTFs is highly correlated with temperature regime (variation in

species composition summarized by the first DCA axis in Table 1) suggests that SDTF formations would have been affected by Pleistocene climatic changes (cf. Prado & Gibbs, 1993; Pennington *et al.*, 2000), assuming that species occupied similar climatic niches in the Pleistocene as they do at present. However, these climatic relationships may have had more complex effects than producing a simple Pleistocene arc of SDTF. Palaeoecological studies are beginning to reveal this complexity (Behling & Lichte, 1997; Burbridge *et al.*, 2004; Whitney *et al.*, 2011, 2013) and more such studies are needed. Palaeodistribution modelling also holds promise for understanding the past distribution of SDTFs (cf. Werneck *et al.*, 2011) but, given the high floristic variation of SDTFs documented here, SDTFs clearly cannot be treated as a single unit (as in Werneck *et al.*, 2011). Rather, the distributions of floristically homogeneous subunits of SDTFs and/or individual species should be modelled, both to understand the past distribution of SDTF and its future under global climate change.

Regarding the identity of SDTFs of the Dry Diagonal, our results showed that the chaco woodlands are not as floristically distinct from the other SDTFs of the Dry Diagonal as has been assumed previously. Also, the second DCA axis supports the view that these southern sites have floristic similarities with various sites of the northern group (i.e. the Misiones nucleus grouping with the Cerrado phytogeographical domain SDTFs and chaco woodlands grouping with the Caatinga nucleus; Fig. 4b). Therefore, we suggest that the identity of the chaco woodlands and how they relate to surrounding SDTFs merits further attention.

Finally, we highlight the necessity of revisiting conservation priorities for dry forests in Brazil. For example, our results demonstrate the floristic similarity between the arboreal caatinga and SDTFs from the Cerrado phytogeographical domain, suggesting they have equal conservation merit in terms of their flora. However, under the current Brazilian environmental law the arboreal caatinga receives full protection, whereas other SDTFs that are equally important, such as the Cerrado domain SDTFs, are unprotected and highly threatened.

### ACKNOWLEDGEMENTS

D.M.N. thanks the Brazilian government funding agency CAPES for supporting a 12-month study period at the Royal Botanic Garden Edinburgh (grant BEX 2415/11-9), where most of the ideas for this study were developed. D.M.N. and A.O.F. received additional funding from Conselho Nacional de Pesquisa Científica–Brazil (CNPq). D.M.N. and R.T.P. were supported by the National Environmental Research Council (grant NE/I028122/1). K.G.D. was funded by an NSF International Research Fellowship (OISE-1103573) during the time this research was completed. We are indebted to Ole Vetaas and the anonymous referees for their valuable contributions to the manuscript.



## REFERENCES

- de Andrade-Lima, D. (1954) *Contribution to the study of the flora of Pernambuco, Brazil*. MSc Thesis, State University of New York, New York.
- de Andrade-Lima, D. (1982) Present-day forest refuges in north-eastern Brazil. *Biological diversification in the tropics* (ed. by G.T. Prance), pp. 245–251. Plenum Press, New York.
- Behling, H. & Lichte, M. (1997) Evidence of dry and cold climatic conditions at glacial times in tropical southeastern Brazil. *Quaternary Research*, **48**, 348–358.
- Bivand, R. (2012) *spdep: spatial dependence. Weighting schemes, statistics and models*. R package version 0.5–51/r458. Available at: <http://R-Forge.R-project.org/projects/spdep/>.
- Borcard, D., Gillet, F. & Legendre, P. (2011) *Numerical ecology with R*. Springer, New York.
- ter Braak, C.J.F. (1986) Canonical correspondence analysis: a new eigen vector technique for multivariate direct gradient analysis. *Ecology*, **67**, 1167–1179.
- Bullock, S.H., Mooney, H.A. & Medina, E. (1995) *Seasonally dry tropical forests*. Cambridge University Press, Cambridge, UK.
- Burbridge, R.E., Mayle, F.E. & Killeen, T.J. (2004) Fifty-thousand-year vegetation and climate history of Noel Kempff Mercado National Park, Bolivian Amazon. *Quaternary Research*, **61**, 215–230.
- Caetano, S. & Naciri, Y. (2011) The biogeography of seasonally dry tropical forest in South America. *Seasonally dry tropical forests: ecology and conservation* (ed. by R. Dirzo, H.S. Young, H.A. Mooney and G. Ceballos), pp. 23–44. Island Press, Washington, DC.
- Castillo-Campos, G., Halffter, G. & Moreno, C.E. (2008) Primary and secondary vegetation patches as contributors to floristic diversity in a tropical deciduous forest landscape. *Biodiversity and Conservation*, **17**, 1701–1714.
- CGIAR-CSI (2006) *NASA Shuttle Radar Topographic Mission (SRTM) (data set available as 3 arc-second DEMs)*. Available at: <http://srtm.csi.cgiar.org/>.
- Collevatti, R.G., Terribile, L.C., Lima-Ribeiro, M.S., Nabout, J.C., Oliveira, G., Rangel, T.F., Rabelo, S.G. & Diniz-Filho, J.A.F. (2012) A coupled phylogeographic and species distribution modeling approach recovers the demographic history of a Neotropical seasonally dry forest tree species. *Molecular Ecology*, **21**, 5845–5863.
- Collevatti, R.G., Lima-Ribeiro, M.S., Diniz-Filho, J.A.F., Oliveira, G., Dobrovolski, R. & Terribile, L.C. (2013a) Stability of Brazilian seasonally dry forests under climate change: inferences for long-term conservation. *American Journal of Plant Sciences*, **4**, 792–805.
- Collevatti, R.G., Terribile, L.C., Oliveira, G., Lima-Ribeiro, M.S., Nabout, J.C., Rangel, T.F. & Diniz-Filho, J.A.F. (2013b) Drawbacks to palaeodistribution modelling: the case of South American seasonally dry forests. *Journal of Biogeography*, **40**, 345–358.
- Dirzo, R., Young, H.S., Mooney, H.A. & Ceballos, G. (2011) *Seasonally dry tropical forests*. Island Press, Washington, DC.
- Dray, S. (2010) *spacemakeR: spatial modelling. R package version 473 0.0–5/r101*. Available at: <http://R-Forge.R-project.org/projects/sedar/>.
- Dray, S., Pelissier, R., Couteron, P., Fortin, M.-J., Legendre, P., Peres-Neto, P.R., Bellier, E., Bivand, R., Blanchet, F.G., de Caceres, M., Dufour, A.-B., Heegaard, E., Jombart, T., Munoz, F., Oksanen, J., Thioulouse, J. & Wagner, H.H. (2012) Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs*, **82**, 257–275.
- EMBRAPA & IBGE (2003) *Mapa de solos do Brasil (escala 1:5.000.000)*. Centro nacional de pesquisas de solos da Empresa Brasileira de Pesquisa Agropecuária and Instituto Brasileiro de Geografia e Estatística, Brasília, Brazil.
- Gebhardt, A. (2009) *tripack: triangulation of irregularly spaced data*. R package version 1.3–4. Available at: <http://CRAN.R-project.org/package=tripack>.
- Gentry, A.H. (1995) Diversity and floristic composition of neotropical dry forests. *Seasonally dry tropical forests* (ed. by S.H. Bullock, H.A. Mooney and E. Medina), pp. 146–194. Cambridge University Press, Cambridge, UK.
- Gillespie, T.W., Grijalva, A. & Farris, C.N. (2000) Diversity, composition, and structure of tropical dry forests in Central America. *Polish Journal of Ecology*, **147**, 37–47.
- Guisan, A., Weiss, S.B. & Weiss, A.D. (1999) GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology*, **143**, 107–122.
- Haffer, J. (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131–137.
- 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.
- Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, NJ.
- IBGE (1993) *Mapa de vegetação do Brasil*. Fundação Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro, Brazil.
- Jones, P. & Harris, I. (2008) *CRU Time Series (TS): high resolution gridded datasets*. NCAS British Atmospheric Data Centre, London.
- Legendre, P., Mi, X., Ren, H., Ma, K., Yu, M., Sun, I. & He, F. (2009) Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology*, **90**, 663–674.
- Legendre, P., Borcard, D. & Roberts, D.W. (2012) Variation partitioning involving orthogonal spatial eigenfunction submodels. *Ecology*, **93**, 1234–1240.
- Lepš, J. & Šmilauer, J.P. (2003) *Multivariate analysis of ecological data using CANOCO*. Cambridge University Press, Cambridge, UK.
- Lewin-Koh, N.J. & Bivand, R. (2012) *maptools: tools for reading and handling spatial objects*. R package version 0.8–17/r238. Available at: <http://R-Forge.R-project.org/projects/maptools/>.

- Linares-Palomino, R. (2006) Phytogeography and floristics of seasonally dry tropical forests in Peru. *Neotropical savannas and seasonally dry forests: plant biodiversity, biogeography and conservation* (ed. by R.T. Pennington, J.A. Ratter and G.P. Lewis), pp. 249–272. CRC Press, Boca Raton, FL.
- Linares-Palomino, R., Oliveira-Filho, A.T. & Pennington, R.T. (2011) Neotropical seasonally dry forests: diversity, endemism and biogeography of wood plants. *Seasonally dry tropical forests: ecology and conservation* (ed. by R. Dirzo, H.S. Young, H.A. Mooney and G. Ceballos), pp. 3–21. Island Press, Washington, DC.
- Lott, E. & Atkinson, T. (2006) Mexican and Central American seasonally dry tropical forests: Chamela-Cuixmala, Jalisco, as a focal point for comparison. *Neotropical savannas and seasonally dry forests: plant biodiversity, biogeography and conservation* (ed. by R.T. Pennington, J.A. Ratter and G.P. Lewis), pp. 315–342. CRC Press, Boca Raton, FL.
- McCune, B. & Grace, J.B. (2002) *Analysis of ecological communities*. MjM Software Design, Gleneden Beach, OR.
- Mooney, H.A., Bullock, S.H. & Medina, E. (1995) Introduction. *Seasonally dry tropical forests* (ed. by S.H. Bullock, H.A. Mooney and E. Medina), pp. 1–8. Cambridge University Press, Cambridge, UK.
- 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. (2012) *vegan: community ecology package*. R package version 2.0–3. Available at: <http://CRAN.R-project.org/package=vegan>.
- Oliveira-Filho, A.T., Jarenkow, J.A. & Rodal, M.J.N. (2006) Floristic relationships of seasonally dry forests of eastern South America based on tree species distribution patterns. *Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation* (ed. by R.T. Pennington, G.P. Lewis and J.A. Ratter), pp. 159–192. CRC Press, Boca Raton, FL.
- Oliveira-Filho, A.T., Budke, J.C., Jarenkow, J.A., Eisenlohr, P.V. & Neves, D.M. (2013) Delving into the variations in tree species composition and richness across South American subtropical Atlantic and Pampean forests. *Journal of Plant Ecology*, **6**, 1–23.
- Pennington, R.T., Prado, D.E. & Pendry, C.A. (2000) Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, **27**, 261–273.
- Pennington, R.T., Ratter, J.A. & Lewis, G.P. (2006) An overview of the plant diversity, biogeography and conservation of neotropical savannas and seasonally dry forests. *Neotropical savannas and seasonally dry forests: plant biodiversity, biogeography and conservation* (ed. by R.T. Pennington, J.A. Ratter and G.P. Lewis), pp. 1–29. CRC Press, Boca Raton, FL.
- Pennington, R.T., Lavin, M. & Oliveira-Filho, A.T. (2009) Woody plant diversity, evolution, and ecology in the tropics: perspectives from seasonally dry tropical forests. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 437–457.
- Prado, D.E. (1993a) What is the Gran Chaco vegetation in South America? I. A review. Contribution to the study of flora and vegetation of the Chaco. V. *Candollea*, **48**, 145–172.
- Prado, D.E. (1993b) What is the Gran Chaco vegetation in South America? II. A redefinition. Contribution to the study of flora and vegetation of the Chaco. VII. *Candollea*, **48**, 615–629.
- Prado, D.E. & Gibbs, P.E. (1993) Patterns of species distribution in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden*, **80**, 902–927.
- Punyasena, S.W., Eshel, G. & McElwain, J.C. (2008) The influence of climate on the spatial patterning of Neotropical plant families. *Journal of Biogeography*, **35**, 117–130.
- Queiroz, L.P. (2006) The Brazilian Caatinga: phytogeographical patterns inferred from distribution data of the Leguminosae. *Neotropical savannas and seasonally dry forests: plant biodiversity, biogeography and conservation* (ed. by R.T. Pennington, J.A. Ratter and G.P. Lewis), pp. 121–157. CRC Press, Boca Raton, FL.
- Quinn, G.P. & Keough, M.J. (2002) *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge, UK.
- R Core Team (2015) *R: a language and environment for statistical computing*. Version 3.1.0. R Foundation for Statistical Computing, Vienna. Available at: <http://www.R-project.org/>.
- Ratter, J.A., Pott, A., Pott, V.J., da Cunha, C.N. & Haridasan, M. (1988) Observations on woody vegetation types in the Pantanal and at Corumba, Brazil. *Notes of the Royal Botanical Garden Edinburgh*, **45**, 503–525.
- Santos, R.M., Oliveira-Filho, A.T., Eisenlohr, P.V., Queiroz, L.P., Cardoso, D.B.O.S. & Rodal, M.J.N. (2012) Identity and relationships of the arboreal Caatinga among other floristic units of seasonally dry tropical forests (SDTFs) of north-eastern and central Brazil. *Ecology and Evolution*, **2**, 409–428.
- Sarkinen, T., Iganci, J.R.V., Linares-Palomino, R., Simon, M.F. & Prado, D.E. (2011) Forgotten forests: issues and prospects in biome mapping using seasonally dry tropical forests as a case study. *BMC Ecology*, **11**, 27.
- Shimodaira, H. (2004) Approximately unbiased tests of regions using multistep multiscale bootstrap resampling. *Annals of Statistics*, **32**, 2616–2641.
- Soininen, J. (2014) A quantitative analysis of species sorting across organisms and ecosystems. *Ecology*, **95**, 3284–3292.
- ter Steege, H., Pitman, N.C.A., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J., Prevoist, M., Spichiger, R., Castellanos, H., von Hildebrand, P. & Vasquez, R. (2006) Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, **443**, 444–447.
- Suzuki, R. & Shimodaira, H. (2011) *pvclust: hierarchical clustering with P-values via multiscale bootstrap resampling*. R package version 1.2–2. Available at: <http://CRAN.R-project.org/package=pvclust>.
- Titeux, N., Dufrene, M., Jacob, J.P., Paquay, M. & Defourny, P. (2004) Multivariate analysis of a fine-scale breeding bird

- atlas using a geographical information system and canonical correspondence analysis: environmental and spatial effects. *Journal of Biogeography*, **31**, 1841–1856.
- Trejo, I. & Dirzo, R. (2002) Floristic diversity of Mexican seasonally dry tropical forests. *Biodiversity and Conservation*, **11**, 2063–2084.
- Vanzolini, P.E. (1963) Problemas faunísticos do Cerrado. *Simpósio sobre o Cerrado* (ed. by M. Ferri), pp. 305–321. Editora da Universidade de São Paulo, São Paulo, Brazil.
- Vetaas, O.R. & Ferrer-Castán, D. (2008) Patterns of woody plant species richness in the Iberian Peninsula: environmental range and spatial scale. *Journal of Biogeography*, **35**, 1863–1878.
- Walter, H. (1985) *Vegetation of the earth and ecological systems of the geo-biosphere*, 3rd edn. Springer-Verlag, Berlin.
- Werneck, F.P., Costa, G.C., Colli, G.R., Prado, D.E. & Sites, J.W., Jr (2011) Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence. *Global Ecology and Biogeography*, **20**, 272–288.
- Werneck, F.P., Nogueira, C., Colli, G.R., Sites, J.W., Jr & Costa, G.C. (2012a) Climatic stability in the Brazilian Cerrado: implications for biogeographical connections of South American savannas, species richness and conservation in a biodiversity hotspot. *Journal of Biogeography*, **39**, 1695–1706.
- Werneck, F.P., Gamble, T., Colli, G.R., Rodrigues, M.T. & Sites, J.W., Jr (2012b) Deep diversification and long term persistence in the South American ‘dry diagonal’: integrating continent-wide phylogeography and distribution modeling of geckos. *Evolution*, **66**, 3014–3034.
- Whitney, B.S., Mayle, F.E., Punyasena, S.W., Fitzpatrick, K.A., Burn, M.J., Guillen, R., Chavez, E., Mann, D., Pennington, R.T. & Metcalfe, S.E. (2011) A 45 kyr palaeoclimate record from the lowland interior of tropical South America. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **307**, 177–192.
- Whitney, B.S., Mayle, F.E., Burn, M.J., Guillen, R., Chavez, E. & Pennington, R.T. (2013) Sensitivity of Bolivian seasonally-dry tropical forest to precipitation and temperature changes over glacial–interglacial timescales. *Vegetation History and Archaeobotany*, **23**, 1–14.
- Zomer, R.J., Bossio, D.A., Trabucco, A., Yuanjie, L., Gupta, D.C. & Singh, V.P. (2007) *Trees and water: smallholder agroforestry on irrigated lands in northern India*. IWMI Research Report 122. International Water Management Institute, Colombo, Sri Lanka.
- Zomer, R.J., Trabucco, A., Bossio, D.A., van Straaten, O. & Verchot, L.V. (2008) Climate change mitigation: a spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture Ecosystems and Environment*, **126**, 67–80.

## DATA ACCESSIBILITY

A full description with details of data accessibility for Neo-TropTree can be found at <http://prof.icb.ufmg.br/treetatlas>

## BIOSKETCH

**Danilo M. Neves** is a Brazilian Research Fellow at the Royal Botanic Garden Edinburgh, UK. He is interested in tropical ecology and evolutionary biology of plants, with an emphasis on niche evolution of Neotropical trees.

Author contributions: D.M.N., K.G.D. and A.O.F. designed the paper; D.M.N., A.O.F. and M.L.B. assembled the database; D.M.N. and K.G.D. analysed the data; A.O.F. and M.L.B. commented on earlier versions of the manuscript; D.M.N., K.G.D. and R.T.P. led the writing. All authors read and approved the final manuscript.

---

Editor: Ole Vetaas