



RESEARCH  
PAPER

# Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence

Fernanda P. Werneck<sup>1\*</sup>, Gabriel C. Costa<sup>2</sup>, Guarino R. Colli<sup>3</sup>,  
Darién E. Prado<sup>4</sup> and Jack W. Sites Jr<sup>1</sup>

<sup>1</sup>Department of Biology and Bean Life Science Museum, Brigham Young University, Provo, UT, 84602, USA, <sup>2</sup>Universidade Federal do Rio Grande do Norte, Centro de Biociências, Departamento de Botânica, Ecologia e Zoologia, Campus Universitário Lagoa Nova, CEP 59072-970, Natal, RN, Brazil, <sup>3</sup>Departamento de Zoologia, Universidade de Brasília, CEP 70910-900, Brasília, DF, Brazil, <sup>4</sup>Cátedra de Botánica, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, PO Box 14, S2125ZAA, Zavalla, SF, Argentina

## ABSTRACT

**Aim** To investigate the potential distribution of Seasonally Dry Tropical Forests (SDTFs) during the Quaternary climatic fluctuations; to reassess the formerly proposed ‘Pleistocenic arc hypothesis’ (PAH); and to identify historically stable and unstable areas of SDTF distributions in the light of palaeodistribution modelling.

**Location** SDTFs in lowland *cis*-Andean eastern-central South America.

**Methods** We first developed georeferenced maps depicting the current distributional extent of SDTFs under two distinct definitions (narrow and broad). We then generated occurrence datasets, which were used with current and past bioclimatic variables to predict SDTF occurrence by implementing the maximum entropy machine-learning algorithm. We obtained historical stability maps by overlapping the presence/absence projections of each of three climatic scenarios [current, 6 kyr BP during the Holocene, and 21 kyr BP during the Last Glacial Maximum (LGM)]. Finally, we checked the consistencies of the model prediction with qualitative comparisons of vegetation types inferred from available fossil pollen records.

**Results** The present-day SDTF distribution is disjunct, but we provide evidence that it was even more disjunct during the LGM. Reconstructions support a progressive southward and eastward expansion of SDTFs on a continental scale since the LGM. No significant expansion of SDTFs into the Amazon Basin was detected. Areas of presumed long-term stability are identified and confirmed (the three nuclear regions, Caatinga, Misiones and Piedmont, plus the Chiquitano region), and these possibly acted as current and historical refugial areas.

**Main conclusions** The LGM climate was probably too dry and cold to support large tracts of SDTF, which were restricted to climatically favourable areas relative to the present day (in contrast with the PAH, as it was originally conceived). Expansions of SDTFs are proposed to have occupied the southern portion of Caatinga nucleus more recently during the early–middle Holocene transition. We propose an alternative scenario amenable to further testing of an earlier SDTF expansion (either at the Lower Pleistocene or the Tertiary), followed by fragmentation in the LGM and secondary expansion in the Holocene. The stability maps were used to generate specific genetic predictions at both continental and regional scales (stable areas are expected to have higher genetic diversity and endemism levels than adjacent unstable areas) that can be used to direct field sampling to cover both stable (predicted refugia) and unstable (recently colonized) areas. Lastly, we discuss the possibility that SDTFs may experience future expansion under changing climate scenarios and that both stable and unstable areas should be prioritized by conservation initiatives.

## Keywords

Caatinga, genetic predictions, habitat stability, historical refugia, Holocene, palaeovegetation modelling, Pleistocene, Quaternary climatic fluctuations, Seasonally Dry Tropical Forest, South America, tropical forest.

\*Correspondence: Fernanda P. Werneck, Department of Biology and Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA.  
E-mail: fewerneck@gmail.com

## INTRODUCTION

Several lines of independent evidence (e.g. pollen records, climatological and genetic data) have been used to assess the importance of climatic fluctuations and vegetation shifts for biological diversification in the Neotropics (Moritz *et al.*, 2000; Mayle *et al.*, 2004; Carnaval & Moritz, 2008; Carnaval *et al.*, 2009). However, the extent to which such fluctuations influenced the evolution of distinct taxonomic groups in different geographic regions (arctic, temperate and tropical) is still under discussion (Hewitt, 2004). Currently, it is recognized that the importance of Quaternary climatic fluctuations might have been overestimated relative to older Tertiary events (Moritz *et al.*, 2000; Pennington *et al.*, 2004). In the Neotropical region, the model known as the 'Refuge theory' first focused on tropical rain forests, and suggested that successive climatic and vegetational cycles during the Pleistocene [particularly during the Last Glacial Maximum (LGM) 21,000 years before the present (21 kyr BP)] promoted either the expansion or retraction of species ranges with diversification driven by such forest fluctuations (Haffer, 1969; Brown & Ab'Sáber, 1979). In this model, speciation occurred in forest refuges during dry/cool periods (glacials), and in savanna refuges during wet/hot periods (interglacials). Some relictual formations [e.g. relict dune systems, Atlantic forest relicts (*brejos*) within the Caatinga region, savanna relicts within the Amazon, etc.] and proposed zones of endemism are often cited as congruent with the 'Refuge theory' (but see Colinvaux *et al.*, 2000, and Furley & Metcalfe, 2007).

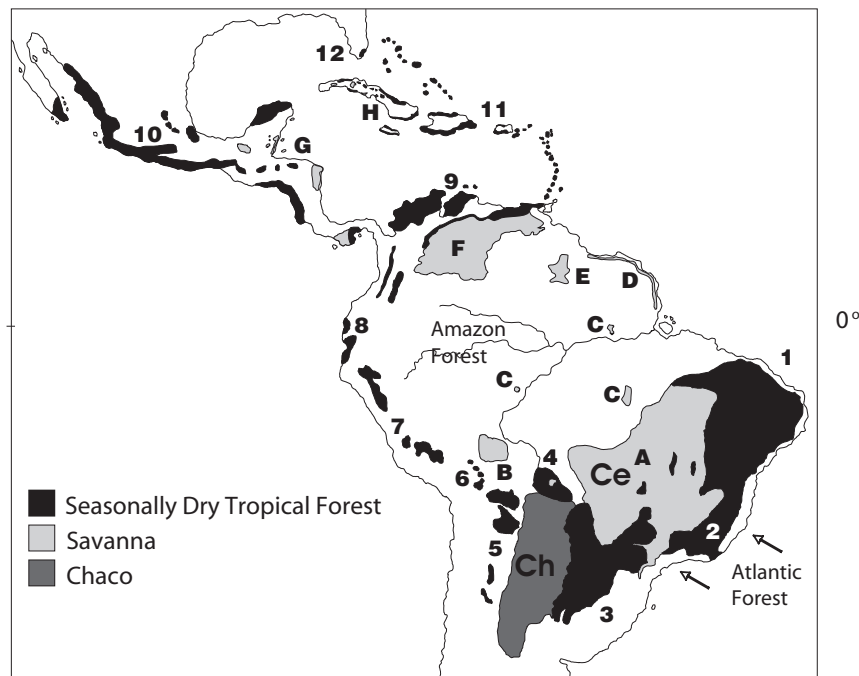
Nonetheless, an increasing body of evidence based on palaeoecological data (mainly the absence of savanna indicator taxa in pollen records) and dynamic vegetation model simulations challenges the 'Refuge theory'. First, regardless of what some generalizations might suggest, palaeoclimates were not uniform across South America, with LGM and Holocene periods showing patchy climates, both spatially and temporally. For an appreciation of the heterogeneity and complexity of Quaternary climate dynamics see the overview in Appendix S1 in the Supporting Information. Current data suggest past biome stability, with most of the Amazon Basin remaining forested at the LGM and with savannas more widespread just at the northern and southern Amazonian margins (Colinvaux *et al.*, 2000; Mayle & Beerling, 2004; Mayle *et al.*, 2004). These studies suggest that instead of a major savanna expansion into the Amazon during the LGM, there were significant changes in the forest-savanna boundaries and ecotonal dynamics, and that a clear savanna expansion in ecotonal areas took place only in the early-middle Holocene (Mayle & Beerling, 2004; Mayle *et al.*, 2004). Also, available genetic evidence suggests that, in several cases, the divergence of extant sister taxa pre-dates the Pleistocene (Moritz *et al.*, 2000) or, at least, an alternative conciliatory scenario in which species diversification was continuous since the Tertiary (late Eocene/early Oligocene, *c.* 39 Ma) until the Quaternary (Pleistocene) (Pennington *et al.*, 2004; Rull, 2008). Nonetheless, much of the debate regarding Quaternary vegetation cycles and their consequences focused on the Amazon and adopted a rainforest/savanna dichotomy, with the potential role of Season-

ally Dry Tropical Forests (SDTFs) largely ignored (Pennington *et al.*, 2000; Pennington *et al.*, 2004; Pennington *et al.*, 2006a).

SDTFs are broadly defined as forests occurring in tropical regions marked by prominent seasonal rainfall, with several months of severe drought (Mooney *et al.*, 1995). Decisive biophysical and climatic factors include soils with high nutrient content and moderate to high pH, frost-free areas where the highly seasonal rainfall is less than 1600 mm year<sup>-1</sup> and a period of at least 5 months of drought (Murphy & Lugo, 1986; Pennington *et al.*, 2006a). The vegetation is heterogeneous, including formations ranging from tall forests to cactus scrub, but is mostly tree-dominated and semi-deciduous to deciduous during the dry season (Murphy & Lugo, 1986; Pennington *et al.*, 2006a). SDTFs are among the most threatened tropical ecosystems, with the greatest annual rate of destruction (Janzen, 1988). On a global basis, Latin American SDTFs experienced the greatest deforestation rates (12% between 1980 and 2000) and virtually all remnants are at risk of extinction, highlighting their urgent priority for conservation (Miles *et al.*, 2006). General overviews concerning the definition, distribution, biogeography and conservation of SDTFs are available in Pennington *et al.* (2006b).

The current distribution of SDTFs in South America is discontinuous, with large areas extending from the Caatinga in north-eastern Brazil to the valley of the Uruguay River, and smaller areas ringing the Amazon Basin and extending into Central America, Mexico, and the Caribbean (Fig. 1). The largest areas referred to as 'nuclei' by Prado & Gibbs (1993) and Prado (2000) are found in north-eastern Brazil (the 'Caatingas nucleus', 1), along the Paraguay-Paraná rivers system (the 'Misiones nucleus', 2), and in south-western Bolivia and north-western Argentina (the 'Subandean Piedmont nucleus', 3) (Fig. 1). Significant portions of SDTF also occur in the Bolivian Chiquitano forest and along the Caribbean coast of Colombia and Venezuela, whereas smaller fragments occur in dry valleys in the Andes of Bolivia, Peru, Ecuador and Colombia, and also in coastal areas of Ecuador and northern Peru. They also appear scattered in areas of favourable edaphic conditions throughout the Cerrado (Ce) savanna biome, in the central Brazil Plateau (Fig. 1).

Debates over Neotropical biogeography were ignited by the suggestion that the current disjunct distributions of SDTFs should be considered a unique relictual biome left over from a formerly extensive uninterrupted formation that reached its maximum extension during a dry-cool period in the LGM, during the late Pleistocene [the 'Pleistocenic arc hypothesis' (PAH); Prado & Gibbs, 1993; Pennington *et al.*, 2000; Prado, 2000]. Also, awareness that the SDTF and Amazonian pollen spectra are nearly indistinguishable by the presence/absence of individual taxa below the genus level (e.g. both are dominated by legumes; Mayle & Beerling, 2004; Mayle *et al.*, 2004; Gosling *et al.*, 2009) raised the possibility that part of the Amazon Basin remained forested during dry glacial periods but may have been occupied by SDTFs rather than rainforests (Pennington *et al.*, 2000; Mayle *et al.*, 2004). Indeed, pollen signatures from modern rainforests, SDTF and Cerrado forest formations (*cerradão*) are



**Figure 1** Distribution of rain forests (Amazon and Atlantic), Seasonally Dry Tropical Forests (SDTFs) and other South American dry vegetation formations. SDTFs: 1, north-east Brazil (Caatinga); 2, south-east Brazilian seasonal forests; 3, Misiones nucleus; 4, Bolivian Chiquitano region; 5, Piedmont nucleus; 6, Bolivian inter-Andean valleys; 7, Peruvian and Ecuadorian inter-Andean valleys; 8, Pacific coastal Peru and Ecuador; 9, Caribbean coast of Colombia and Venezuela; 10, Mexico and Central America; 11, Caribbean Islands (small islands coloured black are not necessarily covered by SDTFs); 12, Florida. Savannas: A, Cerrado; B, Bolivian savannas; C, Amazonian savannas (smaller areas not represented); D, coastal (Amapá, Brazil to Guyana); E, Rio Branco-Rupununi; F, Llanos; G, Mexico and Central America; H, Cuba. Ch: Chaco. Modified from Pennington *et al.* (2006a), with permission.

not easily distinguished by the presence/absence of individual taxa (with *Anadenanthera* being a possible exception for SDTFs; Fabaceae family, common name 'Angico'), and an assemblage approach would be more suitable as indicative of a particular vegetation type (Gosling *et al.*, 2009).

Recent studies have both corroborated (Zanella, 2000; López *et al.*, 2006; Werneck & Colli, 2006; Caetano *et al.*, 2008) and refuted (Mayle, 2004) the PAH as an explanation of the biogeographic history of SDTFs. Contrasting evidence from dynamic vegetation model simulations suggested that, rather than being relictual and refugial populations, the current distribution of SDTFs is Holocene in age, having arisen from forests spreading southwards from Amazon since the LGM (Mayle, 2004; Mayle *et al.*, 2004). However, those simulations were primarily based on a rainforest perspective, mostly concerned with investigating responses of Amazonian ecosystems and carbon storage to climatic changes. Results demonstrated that savannas did not cover the Amazon Basin at the LGM, and SDTFs were then claimed as a possible alternative explanation. From this standpoint, a further investigation to clarify the distribution of SDTFs across the climatic fluctuations from the LGM to the present day is warranted. Preferably, such an investigation should include an alternative methodological approach that incorporates data directly associated with SDTFs (e.g. occurrence/distribution data into a palaeovegetation distribution modelling framework) to identify possible areas of historical stability (and instability),

which can then be tested with independent datasets and approaches (e.g. palynological, phylogeographical and evidence from patterns of endemism; Graham *et al.*, 2006).

The relevance and justification for determining the actual distribution of SDTFs across LGM and Holocene climatic fluctuations are many and go far beyond pure biogeographical interest. First, rainforest, SDTF and savanna have very different values of net carbon storage (Amazon > SDTF > Cerrado savanna), with corresponding distinct impacts on the global carbon budget (Mayle & Beerling, 2004). Consequently, the comprehension of the temporal vegetation dynamics and the interplay between these different biomes may provide interesting clues about the impacts of ongoing global climate change. Second, persuasive arguments stress the importance of considering probable ecosystem responses to climate change for biodiversity conservation strategies (Hannah *et al.*, 2002). Ecosystems are not stationary, and conservation efforts should not be implemented as if they were (Hannah *et al.*, 2002). Third, spatial distributional patterns, species richness, endemism and levels of genetic diversity can be strongly shaped by historical habitat stability (Graham *et al.*, 2006; Carnaval & Moritz, 2008). Spatially explicit modelling combining current and palaeoclimatic models can identify the existence and extent of such stable areas (potential refugia) across past climatic fluctuations, which can in turn be cross-validated by geological, palaeoenvironmental (pollen records) and genetic diversity data (Hugall *et al.*, 2002;

Richards *et al.*, 2007; Carnaval *et al.*, 2009). For example, within the South American Atlantic forest region, a concordant pattern between LGM refugia inferred from palaeoclimate modelling and genetic data was revealed, which permitted the identification of a large stable area through the late Pleistocene–Holocene, the so-called Bahia refuge (Carnaval & Moritz, 2008; Carnaval *et al.*, 2009). However, such an approach has not yet been applied to any of the South American dry vegetation formations. In this case, we would expect a reverse pattern for South American SDTFs when compared with the Atlantic forest: while the Atlantic forest experienced a period of retraction in mostly dry periods of the Pleistocene, followed by expansion from the interglacial periods (such as the Holocene) to the present (pre-industrial), SDTFs are expected to experience an expansion period during the LGM, with subsequent retraction to their current relictual distributions (see Appendix S1 for an overview of the Quaternary climatic change across the Neotropics). A precise recognition of current SDTF refugia depends upon the identification of regions of historical stability. However, the extent and timing of SDTF expansions is still controversial (Mayle, 2004; Mayle & Beerling, 2004), and the hypothesis that they formed a continuous LGM 'Pleistocenic arc' needs further investigation.

The evolution of the South American biota was probably influenced by the relative importance of precipitation, atmospheric CO<sub>2</sub> concentration and temperature during the LGM and Holocene vegetation dynamics (Burbridge *et al.*, 2004). Accordingly, the time frame investigated here is justified by the lack of consensus regarding how the distribution of SDTFs changed from the LGM through the Holocene. Here we apply a distribution modelling (DM) approach to investigate the potential historical distribution and to identify historically stable areas of SDTFs during Quaternary climatic fluctuations. We consider three climatic scenarios (see Appendix S1): current, mid-Holocene (drier period, beginning of precipitation increase) and late Pleistocene at LGM (overall colder and drier period), and then check our palaeomodels for consistency with vegetation types inferred from available palaeoenvironmental data (fossil pollen records). Finally, we generate testable genetic predictions at both continental and regional scales and discuss some conservation implications.

## MATERIALS AND METHODS

### Determining the distributional extent of SDTFs in eastern-central South America

One point evident in the debate over the definition and biogeography of SDTFs is that even the extent of their distribution is not fully described. The primary consequence is that there is no single georeferenced map corresponding to the complete distribution of SDTFs, but only incomplete maps from a variety of sources. For example, the most reproduced SDTF distribution map (Fig. 1) represents only a schematic approximation derived from a variety of sources and an author's personal knowledge following the input of regional specialists, with no associated

geographic information (R. T. Pennington, personal communication). Aiming to analyse the conservation status of SDTFs on a world-wide scale, Miles *et al.* (2006) estimated a global distribution of these forests with associated spatial data. They adopted as a primary source the global biogeographic classification of Olson *et al.* (2001), based on terrestrial ecoregions. However, for South America they made critical mistakes by including the Chaco biome within the range of SDTFs, and by limiting their investigation to the tropical zone. This interpretation excludes important subtropical SDTF areas, such as parts of the Piedmont nucleus of northern Argentina and parts of the Misiones nucleus in south-eastern Paraguay and north-eastern Argentina. Their final map is incomplete, with only three recognized areas of SDTF within all of South America (Appendix S2). A further approach to produce a land-cover map based on remotely sensed satellite data is becoming accepted as the vegetation map for South America, notwithstanding the fact that it also confounds the Chaco and SDTF ranges (Eva *et al.*, 2004). For investigations on a global scale (e.g. Miles *et al.*, 2006) such limitations might not be too serious, but for finer-scale DM these datasets are not appropriate. Accordingly, here we opt to model the spatial range of SDTFs in lowland *cis*-Andean South America (eastern-central) based on two geographical map definitions we developed, representing narrower and broader distributional definitions, respectively, for the SDTFs (Appendix S3). For a complete description of the map's development see Appendix S4.

### Environmental data and distribution modelling

Bioclimatic variables have proved to be good predictors for demonstrating the impacts of climate change on biodiversity (Pearson & Dawson, 2003). Past distributions of organisms were demonstrated to be significantly different from present distributions largely in response to dramatic climate and landscape changes, with potential Pleistocene refugia identified in both temperate (Waltari *et al.*, 2007) and tropical zones (Hugall *et al.*, 2002; Carnaval & Moritz, 2008; Carnaval *et al.*, 2009). Despite not yet being applied to South American dry vegetation formations in a historical context, we expect bioclimatic layers to be good predictors for these biomes as well, especially considering the strong climatic associations described previously (Mooney *et al.*, 1995).

Current climatic variables were downloaded from the WorldClim project at a 30'' (1 km<sup>2</sup>) spatial resolution (available on-line at: <http://www.worldclim.org/>; Hijmans *et al.*, 2005). Past climate data for the LGM (21 kyr BP) and for the Holocene (6 kyr BP) were used following the ECHAM3 atmospheric general circulation model (DKRZ, 1992), available at the Palaeoclimatic Modelling Intercomparison Project webpage (PMIP; <http://pmip.lscce.ipsl.fr/>). See Appendix S1 for the justification for selecting ECHAM3 within the context of the heterogeneous Quaternary climatic fluctuations. The bioclimatic layers were cropped to span from latitude 12°47' N to 34°46' S and longitude 78°31' W to 35° W; this represents a larger spatial range than the estimated distribution of SDTFs (Appendix S3) that

**Table 1** Environmental variables used in the distribution modelling under both definitions of Seasonally Dry Tropical Forests.

Variable	Description	Percentage contribution for narrower definition	Percentage contribution for broader definition (order of relative contribution)
BIO16	Precipitation of wettest quarter	34.1	30 (2)
BIO4	Temperature seasonality	28.1	33 (1)
BIO11	Mean temperature of coldest quarter	16.2	17.2 (3)
BIO15	Precipitation seasonality	8.1	5.9 (5)
BIO7	Temperature annual range	3.8	5.9 (4)
BIO3	Isothermality	3	1.3 (9)
Alt	Altitude	2.8	1.8 (7)
BIO10	Mean temperature of warmest quarter	1.7	2.3 (6)
BIO17	Precipitation of driest quarter	1.4	1.4 (8)
BIO14	Precipitation of driest month	0.9	1.2 (10)

Variables, taken from the WorldClim database except for altitude, are ordered according to the heuristic estimates of their relative contributions to the Maxent model (Phillips & Dudík, 2008) under the narrower definition.

also includes tropical and subtropical zones with bioclimatic conditions compatible with the occurrence of SDTFs. To avoid overparameterization of our DMs with redundant climatic variables, correlation levels for the set of current climatic variables were investigated using a correlation matrix built with ArcGIS version 9.1 (ESRI). Highly correlated variables ( $r > 0.9$ ) were kept in the model on the basis of their biological relevance, following a similar procedure described by Rissler & Apodaca (2007). Ultimately, we selected 9 of the 19 initial environmental variables available at the WorldClim database, plus altitude (Table 1).

We modelled SDTF vegetation from the LGM to the present, implementing the maximum entropy machine-learning algorithm Maxent (Phillips & Dudík, 2008). Maxent has been shown to out-perform other modelling algorithms traditionally used to generate distribution map predictions (Elith *et al.*, 2006). We trained models based on the set of current selected bioclimatic variables that were then projected onto the same palaeoclimatic variables to generate the map predictions. The logistic output format generated by Maxent is an intuitive way to interpret the model's results, in which the conditional probability of occurrence (under current or past climates) is estimated from a set of environmental layers (Phillips & Dudík, 2008). The occurrence datasets for the DM were generated from 1000 presence points randomly selected from the entire distribution of the SDTFs as determined by both definitions we developed (broad and narrow), following recently proposed protocols for other poorly known Neotropical biomes (Carnaval & Moritz, 2008). We then trained Maxent models under current climatic scenarios based on 750 presence records randomly selected from the original 1000 presence points and used the remaining 250 presence records to test each model.

To identify historically stable areas of SDTF distribution, we converted the continuous outputs into presence/absence maps by selecting threshold values for which sensitivity (proportion of true positive predictions versus the number of actual positive sites) is equal to specificity (the proportion of true negative

predictions versus the number of actual negative sites). This approach maximizes agreement between observed and modelled distributions, balancing the cost arising from an incorrect prediction against the benefit gained from a correct prediction (Manel *et al.*, 2001; Pearson *et al.*, 2006). Next, we obtained the historical stability maps representing the predicted refugia where SDTFs have potentially occurred since the LGM, by overlapping the presence/absence projections of each climatic scenario (current, 6 kyr BP and 21 kyr BP) under both narrower and broader definitions (Hugall *et al.*, 2002; Carnaval & Moritz, 2008).

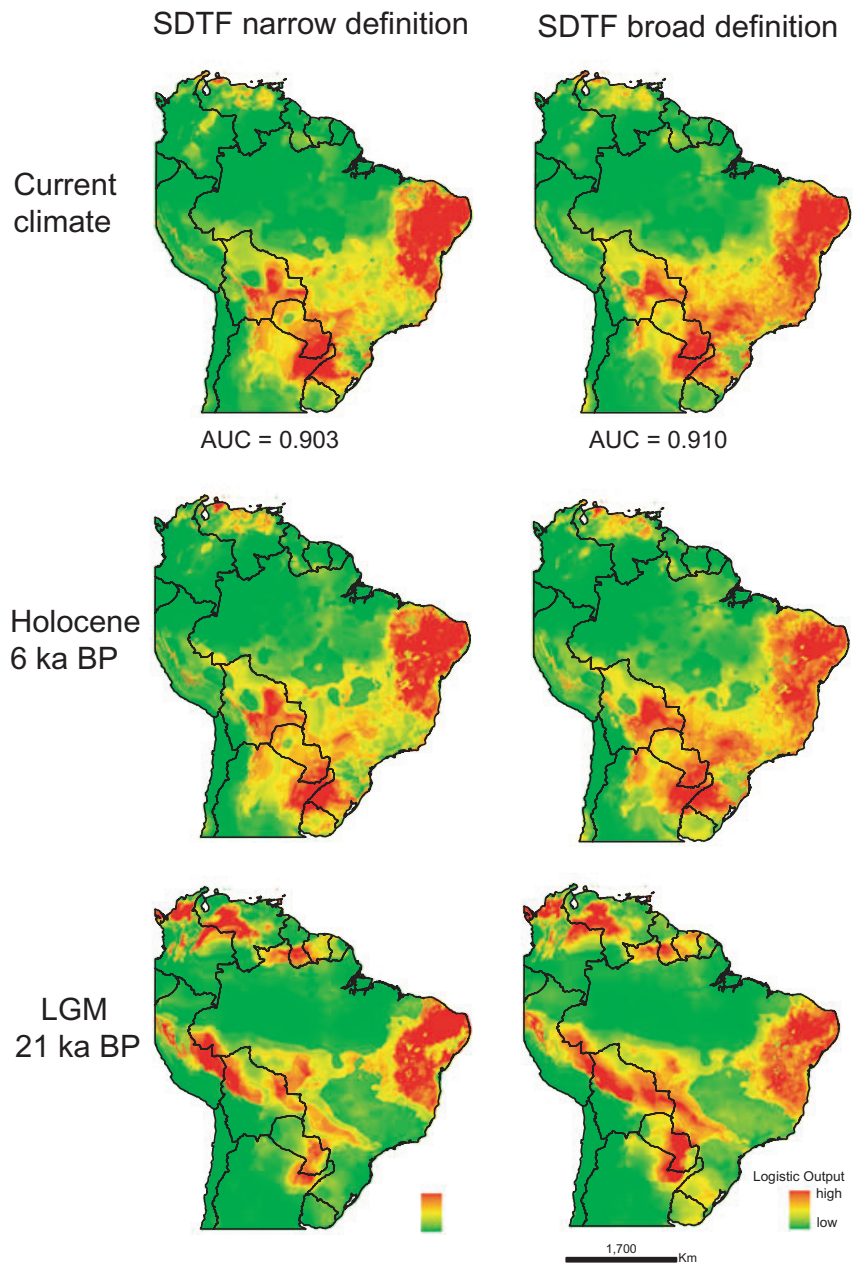
### Palaeomodelling validation based on palynological evidence

We checked the consistency of our model predictions by qualitatively comparing the predicted SDTF distributions under the more conservative definition (narrow) with the vegetation types inferred from available palaeoenvironmental data (fossil pollen records). We included sites located within current SDTFs and predicted ranges, and near their distributional boundaries, excluding sites located in the core areas of the Amazon and Atlantic rainforests.

## RESULTS

### Potential distribution modelling and historical stability

The predicted SDTF distributions modelled under current climatic conditions were very similar under both definitions (narrow and broad), except for an overpredicted occurrence of SDTFs in central Brazil, Cerrado areas for the broader definition modelling (Fig. 2). Both models achieved comparable and satisfactory performances, with a slightly better performance by the broader definition modelling, as evidenced by the area under the receiver operating characteristic curve (AUC) values (Fig. 2).



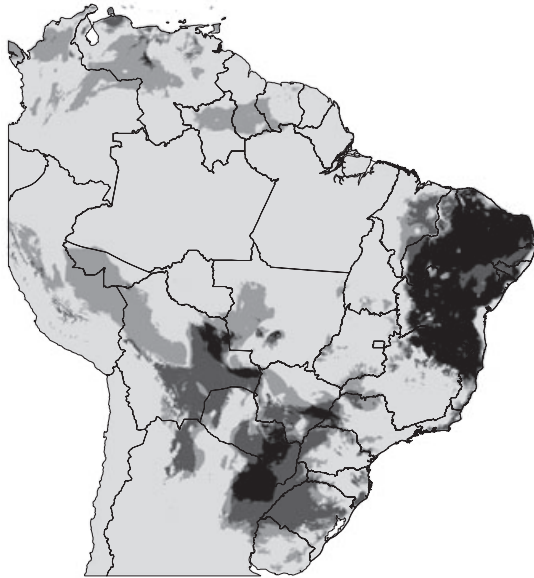
**Figure 2** Modelled ranges of Seasonally Dry Tropical Forests (SDTFs) under narrow and broad distributional definitions for current, 6 kyr BP and 21 kyr BP climatic scenarios. Warmer colours of the logistic output format correspond to regions with a higher probability of SDTF occurrence. AUC refers to area under the receiver operating characteristic curve values, an indicator of model performance.

The AUC is considered an effective indicator of model performance independent of the threshold probability at which the occurrence is accepted (Manel *et al.*, 2001), and is often used to assess the consistency in model predictions of species distributions (Pearson *et al.*, 2006). Generally, AUC values  $> 0.9$  (as is the case for both the SDTF model predictions) are considered indicative of high-accuracy models (Swets, 1988).

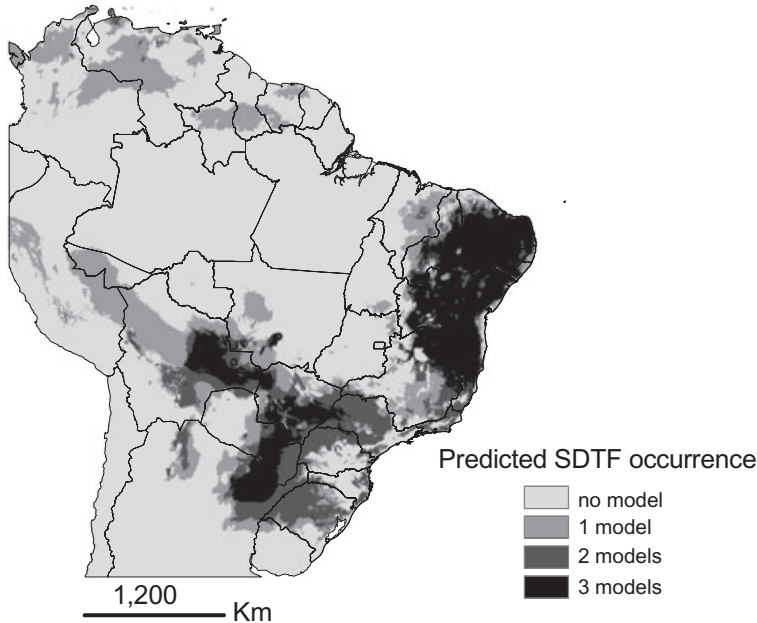
In general, datasets under both definitions recovered very similar potential SDTF palaeodistributions under each investigated climatic scenario, but with the broader definition outputs attaining modestly higher probabilities of occurrence. The range of SDTF predicted distributions under the Holocene climatic scenario was similar to the current climate prediction and potentially more widespread than at the LGM climatic scenario,

with potential SDTF occurrences along the southern edge of the Amazon, in parts of the Brazilian states of Pará, Rondônia and Mato Grosso, albeit with low probabilities (Fig. 2). Interestingly, the models predicted a few major gaps in central Brazil (the states of Goiás and Mato Grosso) where SDTFs were consistently predicted not to occur during the Holocene (Fig. 2). Conversely, the predicted SDTF distributions recovered under the LGM climatic scenario were significantly more restricted in central-eastern Brazil, and the predicted north-western Bolivian vegetation block extended into extreme western Brazil and further north into Peru (adjacent to the Andes). Also, SDTFs were predicted to be more extensive in northern South America during the LGM, potentially covering parts of Colombia, Venezuela and the Guiana highlands. In summary, three main

### SDTF Refugial areas under narrow definition



### SDTF Refugial areas under broad definition

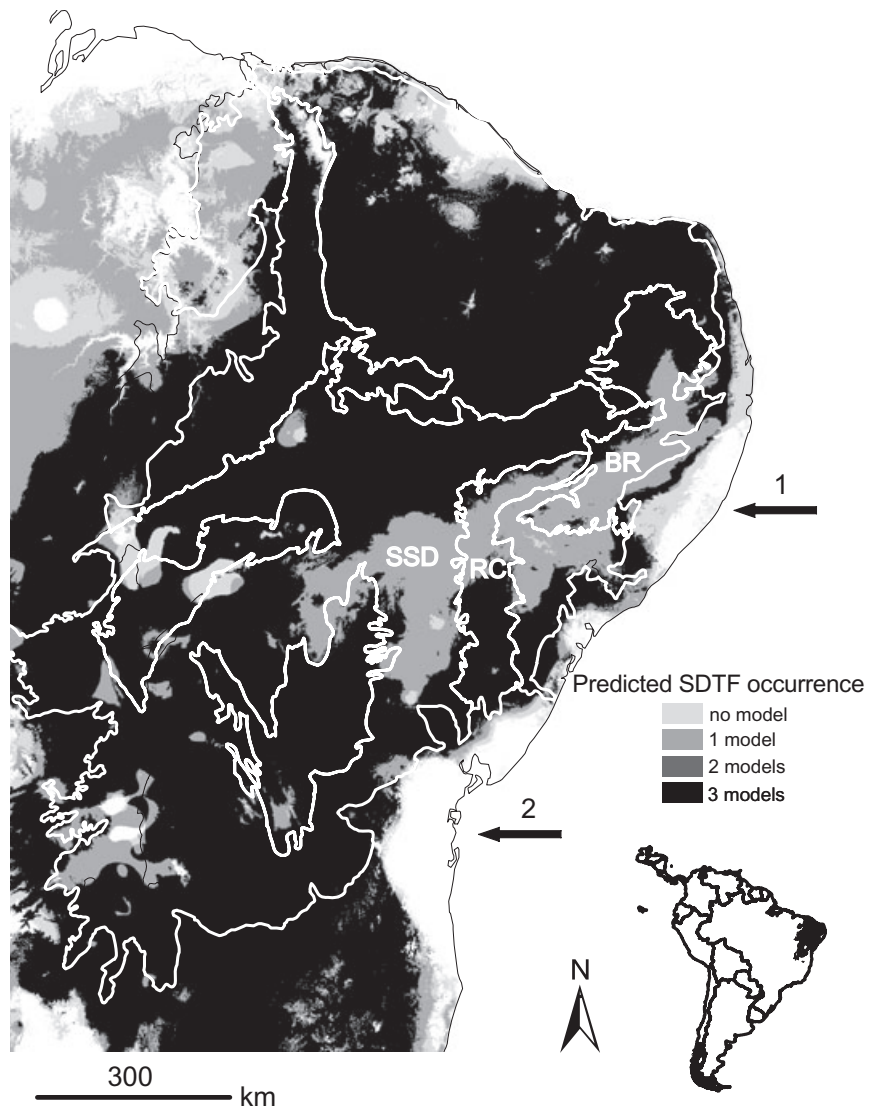


**Figure 3** Summary maps of historically stable and unstable Seasonally Dry Tropical Forest (SDTF) areas under narrow and broad definitions representing the predicted refugia where SDTFs have potentially occurred since the Last Glacial Maximum (LGM). Maps were obtained by overlapping the predicted logistic outputs (i.e. the presence/absence projections) under current, 6 kyr BP and 21 kyr BP ECHAM3 climatic scenarios.

and largely isolated SDTF refugia can be identified during the LGM under both definitions: (1) a northern Amazon block, including the Guiana Shield countries and parts of Colombia and Venezuela; (2) a north-eastern Brazilian block, corresponding to the Caatinga nucleus; and (3) a south-western South American block, extending from south-west Brazil to the Andes, and including the Misiones and Piedmont nuclei and the Bolivian Chiquitano forests. An important feature is the potential narrow corridor just south of the Amazon, connecting the eastern and western extremes of the predicted diagonal distribution (refugia 2 and 3 defined above; Fig. 2).

Regarding long-term stability, four potential refugial areas were predicted (Fig. 3). Two of the potential refugia are located in

south-central South America, one corresponding to the Misiones nucleus in north-eastern Argentina, and continuous with the Paraguay–Paraná rivers system in south-eastern Paraguay, and parts of the Brazilian state of Mato Grosso do Sul (the Serra da Bodoquena region). Another refugium corresponds to the Chiquitano region of eastern Bolivia, which presents some continuance with the Piedmont nucleus, and extending north-west to the Andean region. A small refugium is displaced to the east in southern area of the Brazilian state of Mato Grosso (Fig. 3). The largest refugium corresponds to the core area of the Caatinga in north-eastern Brazil, including northern Minas Gerais state and particularly surrounding the Atlantic forest refugia (Bahia and Pernambuco Refugia) recently identified by Carnaval & Moritz



**Figure 4** Caatinga stable (black) and unstable areas (various shades of grey) across climatic fluctuations under the narrow Seasonally Dry Tropical Forest (SDF) definition, with ecoregion limits in white (Velloso *et al.*, 2002). Arrows identify major areas where none of our models predicted occurrence of SDFs, and are coincident with historically stable areas for the Atlantic forest identified by Carnaval & Moritz (2008). 1, Pernambuco refuge; 2, Bahia Atlantic forest refuge. RC, Raso da Catarina; BR, Borborema Plateau; SSD, Southern Sertaneja Depression.

(2008) (Figs 3 & 4). This large refugium covers most of the current Caatinga distribution, although less stable areas that partially match three of the recognized Caatinga ecoregions can be identified, namely Raso da Catarina, the Borborema Plateau and the Southern Sertaneja Depression (RC, BR and SSD, respectively, in Fig. 4). Finally, it is interesting to note that the region currently occupied by the dry Chaco in north-western Paraguay was never predicted to harbour stable SDF areas (Fig. 3).

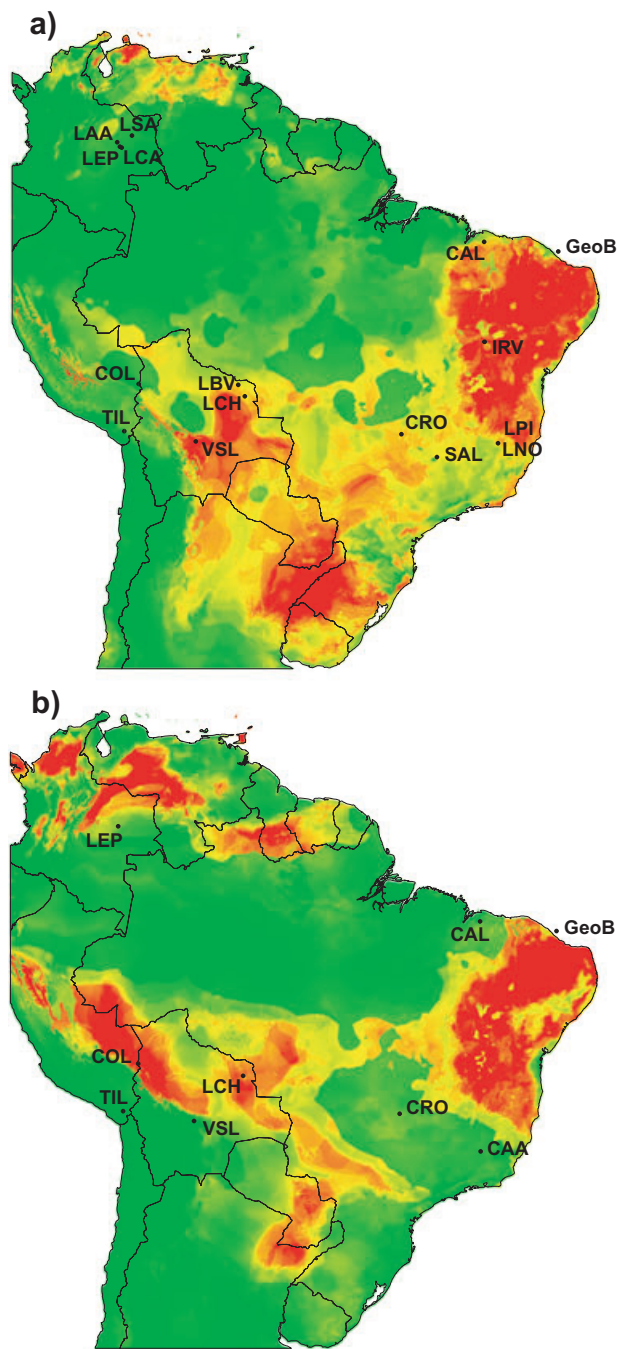
**Palynological validation**

Fossil pollen records were available from 16 localities for the 6 kyr BP Holocene period and from 9 localities for the 21 kyr BP LGM period (Fig. 5). The SDF palaeodistribution modelling predictions were largely congruent with the palynological records from both 6 and 21 kyr BP. In general, the vegetation types predicted from fossil pollen profiles cross-validated the model outputs and confirmed slightly more restricted SDF distributions at the LGM, later followed by early-middle

Holocene expansions, with the mitigation of the drier climatic conditions (Appendix S5, Fig. 5). Validation comes from two classes of evidence, either by providing total concordance (when both the DM and the palynological record suggest the occurrence of SDFs) or partial support (when an area located outside the predicted SDF occurrence was predicted to harbour a vegetation type different from SDF, not contradicting the palaeoclimatic model outputs).

However, in some areas the pollen record suggests the occurrence of the Cerrado savanna and grassland vegetation or cloud forest, while our palaeoclimatic models identified them as areas of potential SDF distribution. This overprediction of SDFs in areas more likely to be covered by Cerrado vegetation occurred for the Holocene climatic scenario at Village Siberia Lake (VSL), at Caçó Lake (CAL) and at the Crominia sites (CRO). In addition, overprediction occurred in areas probably covered by cloud forest very close to the predicted border at Consuelo Lake (COL) for the LGM climatic scenario (Appendix S5, Fig. 5).





**Figure 5** Predicted distributions of Seasonally Dry Tropical Forest under narrow definition and the localities with published palynological data at the Holocene, 6 kyr BP (a) and at the Last Glacial Maximum, 21 kyr BP (b). Locality acronyms follow from Appendix S5.

## DISCUSSION

### Potential historical distribution of SDTF based on palaeodistribution modelling and palynological evidence

The modelled historical distributions of SDTF revealed consistent performance as illustrated by: (1) high AUC values (> 0.9);

(2) congruence with predicted historically stable areas identified for other South American biomes (Carnaval & Moritz, 2008); and (3) congruence with palynological records available for the Quaternary period investigated here (Appendix S5). Also, we interpreted as additional support for model consistency the non-prediction of SDTF occurrence over large areas in central Brazil and in smaller isolated areas within the large Caatinga region, regions more likely to be covered by Cerrado vegetation and by expansions of the Atlantic forest (which subsequently left present-day rain forest natural relicts embedded within the Caatinga vegetation, the so-called *brejos*), respectively. We take these lines of evidence as collective confirmation of the general reliability of the predictions discussed here. We acknowledge some potential limitations, like the absence of edaphic factors in the models. The distribution of calcareous soils might be the overriding control on the distribution of SDTF across South America, rather than climate. However, there are no palaeosoil georeferenced layers available, and LGM soils had only minor changes (such as water and organic contents), with major soil types relatively constant. For example, in the tropics a terrestrial surface cooling of 0.45 °C was produced at the LGM by vegetation and soil feedbacks, but without changing the major soil type (Jiang, 2008). Hence, we do rely upon our simulations of past SDTF distributions even with this possible shortcoming.

According to the palaeodistribution models and palynological validation, the potential historical distribution of SDTF can be summarized as follows. SDTFs had a more restricted distribution during the LGM south of the Amazon, with three main isolated vegetation blocks organized diagonally: an east–west block north of the Amazon; the large Caatinga nucleus of north-eastern Brazil; and a south-east–north-west block south of the Amazon and extending south through eastern Paraguay and into north-eastern Argentina. At that time, a single east–west narrow corridor located in central Brazil (across parts of the states of Goiás, Tocantins and Mato Grosso) potentially connected two of these distributional extremes. This SDTF retraction phase during the supposedly dry LGM period was followed by an expansion starting in the early Holocene, with a major gap predicted in central Brazil, in the present-day core region of Cerrado (Figs 2 & 3). After this expansion phase during the early–middle Holocene transition, the distribution of SDTF experienced only minor fluctuations until achieving its current state. The present-day predicted distribution is more extensive than during glacial times of the Pleistocene, with the exception of northern Amazonian portion and the northern Peruvian margin of the Andes, where SDTFs were potentially more widespread at the LGM. Conflicting with our initial expectations, these patterns mirror to a lesser extent the model described for the historical distribution of the Atlantic forest: forest retraction in dry periods of the Pleistocene followed by expansion from refugia during interglacial periods (such as the Holocene) to the present (Carnaval & Moritz, 2008).

The severe climatic conditions of the LGM could be too dry and too cold to sustain the establishment of large SDTF formations as previously supposed. In fact, SDTFs by definition can tolerate several months of drought but also require some

minimum level of humidity during at least a few months of highly seasonal rainfall (Mooney *et al.*, 1995). Reflecting the magnitude of the aridity, most fossil pollen records are from lakes, and the frequent sedimentation gaps during the LGM (see Appendix S5) suggest dry climatic conditions unfavourable for the persistence of lakes and bogs and subsequent pollen preservation (Ledru *et al.*, 1998). Moreover, the climate was also cold for frost-sensitive deciduous tree species, and SDTFs were possibly restricted to regions where frosts were absent or less frequent (Behling & Lichte, 1997). This explains the absence of predicted occurrence during the LGM in south-eastern Brazilian regions that are currently covered by SDTFs, for example at the Catas Altas palynological site (CAA, Fig. 5). Greater LGM SDTF expansion in the northern Amazonian block (Colombia/Venezuela and the Guiana highlands) could have been due to moderately lower precipitation than present, whilst reduction in SDTF at the LGM in northern Argentina/southern Brazil was more likely due to limiting low temperature at these much higher latitudes. As a consequence, the full-glacial time of the Pleistocene probably supported a more restricted SDTF distribution south of the Amazon, which only recently expanded to occupy the southern portion of the Caatinga nucleus (Mayle, 2004).

### The 'Pleistocenic arc hypothesis'

Our reconstructions contrast with the proposed PAH (Prado & Gibbs, 1993; Pennington *et al.*, 2000; Prado, 2000) that predicted a larger extent of SDTF during dry periods of the Pleistocene at the LGM (even extending into areas now largely covered by Amazon rainforest and Cerrado). According to this hypothesis, the supposed SDTF extensions during the LGM went through subsequent fragmentation (i.e. vicariance) during the Last Glacial–Holocene transition (marked by increased temperature and precipitation levels), which resulted in the present-day disjunct distributions of SDTF, with some nuclear regions acting as current refugia (Prado & Gibbs, 1993). As this hypothesis was based on disjunct floral distributions and patterns of endemism, with no consideration of any temporal or historical context, it may be in error with respect to the period(s) in which SDTF expansion(s) really occurred (more on this topic below). Even though the proposed timing of our findings and of the PAH is different, we did find support for a former slightly more extensive distribution. While the PAH suggests a larger extent during the Pleistocene (Prado & Gibbs, 1993; Pennington *et al.*, 2000; Prado, 2000), our results are more aligned with other studies proposing more recent expansions of SDTF during the Holocene (Mayle, 2004; Mayle *et al.*, 2004). Based on dynamic vegetation model simulations, Mayle (2004) also did not detect a greater SDTF expansion at the LGM. Instead, he proposed a population migration and long-distance dispersal scenario marked by progressive southward expansion of SDTFs over the last 21,000 yr BP, reaching the Chiquitano region (eastern Bolivia) only in the early Holocene (Mayle, 2004). Our results show that the predicted distribution at the LGM included a potential narrow east–west corridor connecting SDTFs in

north-eastern Brazil and at south-western South America extremes; this connection could serve as a biogeographical link sheltering floristic and faunistic dispersal routes. As a consequence, the long-distance dispersal hypothesis of SDTFs in central South America previously proposed by Mayle (2004) remains a scenario that still needs independent testing.

Pennington *et al.* (2000) hypothesized that during dry Quaternary glacial periods, part of the Amazon Basin could have been covered by SDTFs rather than rain forest, arguing that many of the woody species pollen taxa indicated by previous articles could equally come from rain forest and SDTF species alike. This hypothesis was later criticized for overemphasizing the shared taxa between habitats relative to the quantitative aspects of the pollen record (Bush *et al.*, 2004). We found no significant detectable expansion of SDTFs into the Amazon Basin; rather, our palaeomodelling results suggest only a very limited Holocene expansion of SDTFs just at the south-eastern margin of the Amazon, with low probabilities of occurrence that do not demonstrate historical stability (Fig. 3). Our palaeomodels agree with studies showing that the Amazon Basin remained forested during most of the Quaternary (Mayle & Beerling, 2004; Mayle *et al.*, 2004) and that the interpretations of geomorphological features requiring large arid land surfaces for Amazon Basin glacial ages require reassessment (Colinvaux & de Oliveira, 2000; Colinvaux *et al.*, 2000). The lack of calcareous soils across the Amazon basin might also be the reason for the absence of SDTFs there under drier conditions of the LGM. However, Pennington *et al.* (2000) suggest that fertile várzea soils along the main Amazon river channels could have been suitable for SDTF expansion at times of low sea level, in which case they could have grown throughout the basin, but in a dendritic distribution along river systems. Also, it is conceivable that the unique climatic, atmospheric and boundary conditions of the LGM could have permitted SDTF taxa to mix with rain forest taxa, forming novel communities with no modern-day analogue, which would not be captured by our modelling approach. Further investigation is needed to explicitly test these possibilities.

Interestingly, SDTF distributions in inter-Andean dry valleys were predicted to be more extensive during the LGM than during both Holocene and current climatic conditions. The Bolivian SDTFs are scattered across the small and isolated northern and southern Andean dry valleys and at the eastern lowlands, with the largest natural relict in the Chiquitano region in the eastern state of Santa Cruz. This distributional pattern is often interpreted as consistent with a broader past SDTF distribution under the PAH, when colonization of the northern dry valleys by lowland species possibly took place (Herzog & Kessler, 2002). Conversely, species-based distribution modelling found a more restricted SDTF distribution in the Andes during the LGM (modelled by the authors as a cold, wet period; Soria-Auza, 2009) and vegetation simulations (with very coarse spatial resolution) suggested that SDTFs only covered the Chiquitano region since the early Holocene (Mayle, 2004). Our results are consistent with the first scenario: an earlier colonization of the northern valleys by SDTF species during the Pleistocene LGM

period. Nevertheless, the floristic composition of these Pleistocene SDTF communities was probably quite different from those of modern SDTFs, as indicated by the lack of *Anadenanthera* (a key SDTF indicator taxon) in the fossil pollen record of the LGM in the Chiquitano region (Appendix S5; Burbridge *et al.*, 2004). As evidenced by the discordance between the palaeorecords at Consuelo Lake (COL) and our models, extensive LGM SDTF distributions in the Andean region could represent an overprediction of our models. There is a consensus that the extent of Andean SDTFs fluctuated significantly, with some favourable periods enabling connections between lower- and higher-elevation areas. However, additional high-resolution regional studies regarding the timing of such connections are clearly needed.

Despite not fully corroborating the PAH as originally conceived (Prado & Gibbs, 1993; Pennington *et al.*, 2000; Prado, 2000), the historical distribution of SDTFs predicted here does not conflict with most of the evidence often used to support it. Disjunct distributions and the occurrence of many plant endemics in SDTFs (Prado & Gibbs, 1993; Pennington *et al.*, 2000; Prado, 2003; Queiroz, 2006) and some bee (Zanella, 2000) and lizard species (Werneck & Colli, 2006) have been suggested as consistent with the PAH. Nonetheless, instead of originating through vicariance after the fragmentation of a hypothesized continuous LGM distribution, these disjunctions and endemism patterns could have originated either due to ancient dispersal events or at an earlier expansion period, which fragmented before the LGM. Long-distance ancient dispersal events of unrelated taxonomic groups are less likely than vicariant associations (*sensu* Myers & Giller, 1988), and there is substantial support that SDTF is a highly dispersal-limited biome as evidenced by phylogenetic niche conservatism and strong geographically structured phylogenies of constituent woody plant clades (Pennington *et al.*, 2009). So, we support the earlier expansion scenario to be submitted to further testing. Two alternative approaches can be used to test this scenario: (1) exploration of the historical distributions of SDTF during the climatic conditions of the Lower Pleistocene (Gelasian; 1.8–2.5 Ma) and earlier Tertiary period (limited by the general lack of past climate layers for these periods); and (2) molecular phylogenetic/phylogeographic studies dating the divergence times between SDTF endemic species and/or populations and their close relatives. If this earlier expansion scenario (lower Pleistocene or Tertiary) is plausible, we would expect older divergence times and strong phylogeographic structure between the SDTF nuclear regions (refugia) and genetic evidence of recent population expansion and secondary contact (especially in the unstable regions) during the Holocene expansion period.

### Relations of SDTFs with other dry biomes

The mid-Holocene is often described as the period when the final establishment of Cerrado took place, with its present-day distribution stabilizing around 7 kyr BP, even though it is a biome of much older age (Ledru, 2002). This pattern is consistent with the SDTF distributions modelled under the Holocene

climatic scenario, which predicted an expansion of SDTFs during this period, but with major gaps of unsuitable areas located in central Brazil, very probably covered by Cerrado vegetation (Fig. 2). In fact, the higher and more spatially continuous plateaus in central Brazil (e.g. the Central Goiás and Guimarães plateaus) are hypothesized to have formed a single large Cerrado refugium during the Late Pleistocene/early Holocene, while lower plateaus and peripheral depressions were much drier than today and dominated by more xeric-adapted vegetation, derived either as SDTF expansions or from colder dry biomes in southern South America (e.g. Pampa and Monte) (Ab'Sáber, 1983). Amelioration of the climate towards current conditions allowed the Cerrado to spread from this large refugium, while other vegetation types retracted their distributions within the Cerrado biome to some isolated patches with favourable edaphic conditions (derived from limestones), mainly in peripheral depressions (Prado & Gibbs, 1993). Conversely, our models also overpredicted the occurrence of SDTF in some central Brazilian areas, particularly under the broader definition modelling (Fig. 2). Such areas are suitable in terms of climatic envelopes, but probably lack other factors suitable for the occurrence of SDTF (not included in our models), such as fertile soils. These two classes of results confirm the long-term complex nature of the highly dynamic and fluctuating savanna and dry forest boundaries (Furley & Metcalfe, 2007).

The Chaco is a dry vegetation formation of lowland alluvial plains located primarily in northern Argentina, western Paraguay, south-eastern Bolivia and the extreme western edge of Mato Grosso do Sul state in Brazil (Prado, 1993b; Pennington *et al.*, 2000). Like the SDTF and Cerrado, Chaco climate is marked by strong seasonality, but with more severe summers and winter frosts. Historical connections between the Chaco and the Caatinga have been proposed (Andrade-Lima, 1954; Vanzolini, 1974), but more recently these links were demonstrated to be negligible, and the Chaco was excluded from the definition of SDTF both floristically and biophysically in terms of climate, soils and topographic conditions (Prado, 1993a,b; Pennington *et al.*, 2000). Our models never predicted the occurrence of stable SDTF areas in north-western Paraguay (currently occupied by the dry Chaco), providing additional evidence to support such exclusion, and endorse the view that both biomes are determined by different bioclimatic factors and experienced distinct evolutionary histories.

However, specific modelling for each of the other dry biomes is necessary to explore the fit of the predicted distributions and to assess if their vegetation dynamics are complementary to the historical distributions of SDTF proposed here. Further investigation incorporating the possible effects of finer-scale factors (e.g. soils, fire and flood dynamics) for the distributional differences between these biomes (SDTF, Cerrado and Chaco) represent interesting research agendas, with major impacts on the development of improved phylogeographic predictions for broadly distributed species and strategies for biodiversity conservation. For example, climatically favourable areas can lack other biophysical determinants for the occurrence of a given biome (e.g. fertile soils for SDTFs and plant biomass fuel in the

dry season for Cerrado), and we caution that geographically expansive modelling such as this study will require follow-up studies.

### Historical stability, biodiversity and genetic predictions

The predicted areas of long-term stability under both definitions (narrow and broad; Fig. 3) match the proposed nuclei representing the largest present-day areas of SDTF (Prado & Gibbs, 1993), namely the 'Caatinga nucleus' in north-eastern Brazil, the 'Misiones nucleus' along the Paraguay–Paraná rivers system (eastern Paraguay and north-eastern Argentina) and the 'subandean Piedmont nucleus' in south-western Bolivia and north-western Argentina (Prado, 2000). This last nucleus is approximately continuous with the Chiquitano region (eastern Bolivia), which also demonstrates historical stability and should eventually also be considered as a nuclear region of South American SDTF. In other words, the SDTF nuclei were demonstrated to have historical stability, potentially acting as long-term and current refugial areas. Accordingly, Pennington *et al.* (2009) argue that the fragmented SDTF distribution (including their nuclei) has persisted over sufficiently long evolutionary time-scales to have shaped patterns of woody plant species distributions, and phylogenetic and population genetic structure.

The Caatinga is the largest refugium but it is not a continuous stable region because it also includes relatively large zones of low stability, mostly confined to three of the eight previously documented Caatinga ecoregions (Velloso *et al.*, 2002): Raso da Catarina, the Borborema Plateau and the Southern Sertaneja Depression (Fig. 4). These within-Caatinga unstable regions coincide with the locality of many present-day montane forest refugia, the so-called *brejos de altitude* (Porto *et al.*, 2004). In addition, the large Caatinga refugium is in strict proximity with the Atlantic forest refugia predicted from palaeomodelling (Carnaval & Moritz, 2008). In these areas (Pernambuco and Bahia refugia; areas 1 and 2, respectively, in Fig. 4), climate is dramatically distinct and SDTFs were never predicted to occur. Hence, the unstable regions identified within Caatinga were presumably susceptible to many vegetational shifts and seemingly accommodated Atlantic forest expansions across the climatic fluctuations. It is interesting to remark that although presenting long-term stability, the Caatinga refugia vegetation may have differed structurally across Quaternary fluctuations, as suggested by leaf macrofossil travertine deposits showing that LGM climates in the region enabled growth of semi-deciduous forest in areas that today support only cactus thorn–scrub (Auler *et al.*, 2004).

The Quaternary climatic oscillations influenced organismal evolutionary histories in many ways, including causing extinctions and repeated changes in the distributions of surviving species (Hewitt, 2004). In such a context, historical habitat stability can predict species and genetic diversity as well as distributional and endemism patterns (Hugall *et al.*, 2002; Graham *et al.*, 2006; Carnaval *et al.*, 2009). As a consequence, habitat shifts due to past climate changes are expected to produce popu-

lation genetic signatures increasingly detectable with the development of sophisticated molecular phylogeographic analyses (Avice, 2009). Our predicted historical modelling maps (Figs 2–4) can be used to formulate GIS-based, spatially explicit endemism and genetic diversity hypotheses that to be tested require sampling across stable and unstable regions (for a comprehensive appraisal of all phases involved in such 'biodiversity prediction' approach see Carnaval *et al.*, 2009). First, temporally unstable regions (e.g. southern Caatinga in the Brazilian states of Minas Gerais and São Paulo; the above-noted within-Caatinga unstable regions; the large region around the 'Misiones nucleus' along the Paraguay–Paraná rivers system; the Andean region from Bolivia to Peru; light colours in Figs 3 & 4), are expected to have been colonized only more recently and consequently to retain genetic signatures of population expansions, and display overall lower or shallower levels of within-species genetic diversity when compared with stable areas (Hewitt, 2004). Other testable expectations are that the unstable areas will reveal a lack of isolation-by-distance patterns if colonization was too recent to permit stabilization of migration–drift equilibrium, as well as lower levels of endemism (Carnaval & Moritz, 2008; Carnaval *et al.*, 2009). On the other hand, temporally stable areas (darker colours in Figs 3 & 4) probably had higher persistence and sheltered reasonable population sizes through climatic fluctuations, and are expected to retain higher genetic diversity than those in unstable recently colonized areas and strong phylogeographic structure between refugia (Hewitt, 2004; Carnaval *et al.*, 2009). Despite not aiming to test these specific predictions, Caetano *et al.* (2008) provided support for our model when revealing the higher genetic diversity of the tree *Astronium urundeuva* (Anacardiaceae) from the south-western region (coincident with the Misiones refugial area), and evidence of secondary contact at the southern limit of the Caatinga SDTF nucleus (here recovered as an unstable region prone to population expansion; see below).

At the community level, historically stable areas should permit more species to arise and persist, resulting in high species diversity and endemism (Graham *et al.*, 2006), as supported by the large number of endemic plant species that have accumulated in the biggest refugium, the Brazilian Caatinga (Queiroz, 2006). Finally, above the species level, phylogenetic signal concordant with the long-term stability scenario proposed here include: the occurrence of clades confined to SDTFs (evidence of phylogenetic niche conservatism); endemic species restricted to a single nucleus expected to be monophyletic and relatively old; sister species predicted to occupy the same SDTF nucleus (Pennington *et al.*, 2009).

With Holocene population expansions, the SDTF biota could experience secondary contact between refugia, and here 'suture zones' might stabilize in 'troughs' of suboptimal habitat and be characterized by historical or ongoing hybridization (Hugall *et al.*, 2002; Swenson, 2008). One well-defined suture zone has recently been quantified for multiple pairs of sister species/sister haploclades within species, between stable refugia in the Australian wet tropics (Moritz *et al.*, 2009). For species that are restricted to a given stable area of SDTF, the levels of genetic

diversity could also be investigated across the distributional range, to test several alternative hypotheses, including: (1) decreasing genetic diversity towards the range periphery; (2) increasing diversity towards the range periphery; and (3) a recolonizing hypothesis that for many species genetic diversity may be expected to peak in the subperipheral regions of a distribution, rather than in the extreme periphery or the distributional core (Kark *et al.*, 2008).

Despite higher levels of biodiversity, knowledge of Neotropical phylogeographic and genetic diversity patterns is extremely deficient when compared with that of temperate regions (Hewitt, 2004), especially with respect to South American dry vegetation formations (Beheregaray, 2008). Furthermore, within the Neotropics most phylogeographic studies are concentrated in rain forest regions, and these repeatedly report a variety of complex historical patterns in response to climatic fluctuations (for some examples see Moritz *et al.*, 2000, Noonan & Gaucher, 2006, and Carnaval *et al.*, 2009). With a single exception (Caetano *et al.*, 2008), SDTFs have not been represented in these studies, which are urgently needed to test some of the hypotheses proposed here. Our stability maps generate genetic predictions that can drive field sampling to cover both stable (predicted refugia) and unstable (recently colonized) areas, as well as formerly unsampled areas. Comparative phylogeographic approaches based on coalescent methods (Richards *et al.*, 2007; Carnaval *et al.*, 2009) can then provide rigorous statistical tests of these predictions.

### Conservation implications

Knowledge about past vegetation dynamics is essential to comprehend how vegetational boundaries might shift in response to future climate change. Taking into account that the main goal of conservation biology is to preserve biodiversity in the context of maintaining long-term ecological and evolutionary processes (Moritz, 2002), and that ecosystems are not static, temporal vegetation dynamics should be considered in conservation efforts (Hannah *et al.*, 2002). In the case of SDTFs, one of the most threatened tropical ecosystems (Janzen, 1988), urgent and effective conservation strategies are especially important.

Insights from mid-Holocene palaeodata, coupled with modelling studies, suggest that there will be progressive replacement of rainforest by SDTFs and savanna in the Bolivia's Noel Kempff Mercado National Park (NKMNP) over the 21st century, in response to the increased drought and warming predicted by general circulation models (Mayle *et al.*, 2007). Our results revealed a similar general expansion trend, starting at early-middle Holocene, for SDTFs in general and also in the Bolivian Chiquitano region. Considering the occurrence of many threatened species and the endangered status of savannas and SDTFs (all under greater anthropogenic pressure than the rain forests), these could be interpreted as positive outcomes from a conservation perspective given that future climate change could actually favour the expansion of these biomes and the associated biota (Mayle *et al.*, 2007). However, most biologists would not advocate a passive response to such a rain forest–dry forest shift,

but rather the protection of large natural areas of both formations and the ecotones between them. Without the refugia, for example, the scenario of potential expansion of SDTF could be compromised by a general lack of source populations. For example, the Caatinga region has only 11 strictly protected areas, representing less than 1% of the region; the smallest protected area of any major Brazilian biome (Leal *et al.*, 2005). Finally, the unstable regions can also play an important role in the conservation of SDTFs. Some of these areas are in close contact with other vegetation formations, forming huge ecotonal mosaics and environmental gradients. These areas are thought to select for adaptive variation and may have high speciation and conservation potential, even though the transitional vegetation zones have been neglected in recent conservation initiatives (Smith *et al.*, 2001; Mayle *et al.*, 2007).

### MAIN CONCLUSIONS

In summary, our models were in general agreement with available palynological evidence, and where this was not the case discrepancies could be explained by well-known phenomena (deficient pollen records due to an absence of LGM bedding planes in lakes, restriction by the absence of other determinant environmental factors, etc.). The distribution of SDTF was predicted to be more restricted in southern Amazon under arid climates of the LGM, with a potential narrow corridor in central Brazil connecting the extremes of the predicted distributional range. Subsequently, SDTFs experienced a gradual southward expansion starting in the early Holocene, with a major gap predicted in central Brazil where Cerrado vegetation was established. After this expansion phase, SDTFs experienced only minor fluctuations until attainment of their current distribution. LGM climate was probably too dry and cold to support large tracts of SDTF, results that contradict the timing of the PAH. We propose an alternative scenario amenable to further testing, either by molecular phylogeographic studies or palaeodistribution modelling, of a Lower Pleistocene or earlier Tertiary SDTF expansion, followed by fragmentation during the LGM and secondary expansion in the Holocene.

The SDTF nuclei and the Chiquitano region are predicted to represent areas of long-term stability and very likely served as refugial areas during the LGM. These refugia are consequently expected to have higher levels of genetic diversity and endemism than adjacent unstable areas. This scenario could be used to direct geographic sampling for molecular studies aiming to cross-validate our palaeomodelling results, and test the spatially explicit phylogeographic hypotheses proposed here. Finally, both stable and unstable areas of SDTFs should be prioritized by conservation initiatives, so these forests can fulfil their evolutionary potential, which may include a period of future expansion under changing climate scenarios.

### ACKNOWLEDGEMENTS

This work was submitted in partial fulfilment by F.P.W. for the PhD degree at Brigham Young University (BYU). F.P.W. and

G.C.C. were supported by fellowships from Fulbright/CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; 15073722-2697/06-8 and 15053155-2018/04-7, respectively) and F.P.W. by a graduate research assistantship from the BYU Department of Biology. F.P.W. received additional funding from the National Geographic Society (NGS), Society of Systematic Biologists (SSB), Neotropical Grassland Conservancy (NGC), Society for the Study of Amphibians and Reptiles (SSAR) and Idea Wild. G.R.C. is supported by grants from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and D.E.P. thanks CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina). We thank Ana Carolina Carnaval and the Museum of Vertebrate Zoology (MVZ) GIS Portal staff for assistance with the palaeoclimatic layers; the UNEP-WCMC Spatial Analysis Unit for access to their estimated tropical dry forest dataset; the Systematic Association for permission to reproduce Fig. 1, F. Pareyn for access to the Caatinga ecoregions dataset; S. K. Herzog, M. Kessler, R. Soria-Auza and L. P. Queiroz for helpful bibliography exchange; and the two referees, Francis E. Mayle and R. Toby Pennington, for their detailed comments and suggestions which substantially improved the paper.

## REFERENCES

- Ab'Sáber, A.N. (1983) O domínio dos cerrados: introdução ao conhecimento. *Revista do Serviço Público*, **111**, 41–55.
- Andrade-Lima, D. (1954) *Contribution to the study of the flora of Pernambuco, Brazil. Monografias*. Universidade Rural de Pernambuco, Recife.
- Auler, A.S., Wang, X., Edwards, R.L., Cheng, H., Cristalli, P.S., Smart, P.L. & Richards, D.A. (2004) Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. *Journal of Quaternary Science*, **19**, 693–701.
- Avise, J.C. (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36**, 3–15.
- Beheregaray, L.B. (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, **17**, 3754–3774.
- Behling, H. & Lichte, M. (1997) Evidence of dry and cold climatic conditions at glacial times in tropical southeastern Brazil. *Quaternary Research*, **48**, 348–358.
- Brown, K.S. & Ab'Sáber, A.N. (1979) Ice-age forest refuges and evolution in Neotropics: correlation of paleoclimatological, geomorphological and pedological data with biological endemism. *Paleoclimas*, **5**, 1–30.
- 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.
- Bush, M.B., de Oliveira, P.E., Colinvaux, P.A., Miller, M.C. & Moreno, J.E. (2004) Amazonian paleoecological histories: one hill, three watersheds. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **214**, 359–393.
- Caetano, S., Prado, D., Pennington, R.T., Beck, S., Oliveira-Filho, A.T., Spichiger, R. & Naciri, Y. (2008) The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Molecular Ecology*, **17**, 3147–3159.
- Carnaval, A.C. & Moritz, C. (2008) Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.
- Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T. & Moritz, C. (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, **323**, 785–789.
- Colinvaux, P.A. & de Oliveira, P.E. (2000) Palaeoecology and climate of the Amazon basin during the last glacial cycle. *Journal of Quaternary Science*, **15**, 347–356.
- Colinvaux, P.A., de Oliveira, P.E. & Bush, M.B. (2000) Amazonian and neotropical plant communities on glacial timescales: the failure of the aridity and refuge hypotheses. *Quaternary Science Reviews*, **19**, 141–169.
- DKRZ (1992) The ECHAM3 atmospheric general circulation model. *DKRZ Technical Report No. 6*. Deutsches Klimarechenzentrum, Hamburg, Germany.
- Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Eva, H.D., Belward, A.S., De Miranda, E.E., Di Bella, C.M., Gond, V., Huber, O., Jones, S., Sgrenzaroli, M. & Fritz, S. (2004) A land cover map of South America. *Global Change Biology*, **10**, 731–744.
- Furley, P.A. & Metcalfe, S.E. (2007) Dynamic changes in savanna and seasonally dry vegetation through time. *Progress in Physical Geography*, **31**, 633–642.
- Gosling, W.D., Mayle, F.E., Tate, N.J. & Killeen, T.J. (2009) Differentiation between Neotropical rainforest, dry forest, and savannah ecosystems by their modern pollen spectra and implications for the fossil pollen record. *Review of Palaeobotany and Palynology*, **153**, 70–85.
- Graham, C.H., Moritz, C. & Williams, S.E. (2006) Habitat history improves prediction of biodiversity in rainforest fauna. *Proceedings of the National Academy of Sciences USA*, **103**, 632–636.
- Haffer, J. (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131–137.
- Hannah, L., Midgley, G.F., Lovejoy, T., Bond, W.J., Bush, M., Lovett, J.C., Scott, D. & Woodward, F.I. (2002) Conservation of biodiversity in a changing climate. *Conservation Biology*, **16**, 264–268.
- Herzog, S.K. & Kessler, M. (2002) Biogeography and composition of dry forest bird communities in Bolivia. *Journal für Ornithologie*, **143**, 171–204.
- Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 183–195.
- 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.

- Hugall, A., Moritz, C., Moussalli, A. & Stanisci, J. (2002) Reconciling paleodistribution models and comparative phylogeography in the wet tropics rainforest land snail *Gnarosiphia bellendenkerensis* (Brazier 1875). *Proceedings of the National Academy of Sciences USA*, **99**, 6112–6117.
- Janzen, D.H. (1988) Tropical dry forests: the most endangered major tropical ecosystem. *Biodiversity* (ed. by E.O. Wilson), pp. 130–137. National Academy Press, Washington, DC.
- Jiang, D. (2008) Vegetation and soil feedbacks at the Last Glacial Maximum. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **268**, 39–46.
- Kark, S., Hadany, L., Safriel, U.N., Noy-Meir, I., Eldredge, N., Tabarroni, C. & Randi, E. (2008) How does genetic diversity change towards the range periphery? An empirical and theoretical test. *Evolutionary Ecology Research*, **10**, 391–414.
- Leal, I.R., Silva, J.M.C., Tabarelli, M. & Lacher, T.E. Jr (2005) Changing the course of biodiversity conservation in the Caatinga of northeastern Brazil. *Conservation Biology*, **19**, 701–706.
- Ledru, M.-P. (2002) Late Quaternary history and evolution of the Cerradões as revealed by palynological records. *The Cerrados of Brazil: ecology and natural history of a Neotropical savanna* (ed. by P.S. Oliveira and R.J. Marquis), pp. 33–50. Columbia University Press, New York.
- Ledru, M.-P., Salgado-Labouriau, M.L. & Lorscheitter, M.L. (1998) Vegetation dynamics in southern and central Brazil during the last 10,000 yr B.P. *Review of Palaeobotany and Palynology*, **99**, 131–142.
- López, R.P., Larrea-Alcázar, D. & Macía, M.J. (2006) The arid and dry plant formations of South America and their floristic connections: new data, new interpretation? *Darwiniana*, **44**, 18–31.
- Manel, S., Williams, H.C. & Ormerod, S.J. (2001) Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, **38**, 921–931.
- Mayle, F.E. (2004) Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. *Journal of Quaternary Science*, **19**, 713–720.
- Mayle, F.E. & Beerling, D.J. (2004) Late Quaternary changes in Amazonian ecosystems and their implications for global carbon cycling. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **214**, 11–25.
- Mayle, F.E., Beerling, D.J., Gosling, W.D. & Bush, M.B. (2004) Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 499–514.
- Mayle, F.E., Langstroth, R.P., Fisher, R.A. & Meir, P. (2007) Long-term forest–savannah dynamics in the Bolivian Amazon: implications for conservation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362**, 291–307.
- Miles, L., Newton, A.C., DeFries, R.S., Ravilious, C., May, I., Blyth, S., Kapos, V. & Gordon, J.E. (2006) A global overview of the conservation status of tropical dry forests. *Journal of Biogeography*, **33**, 491–505.
- 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.
- Moritz, C. (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.
- Moritz, C., Patton, J.L., Schneider, C.J. & Smith, T.B. (2000) Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533–563.
- Moritz, C., Hoskin, C.J., MacKenzie, J.B., Phillips, B.L., Tonione, M., Silva, N., VanDerWal, J., Williams, S.E. & Graham, C.H. (2009) Identification and dynamics of a cryptic suture zone in tropical rainforest. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1235–1244.
- Murphy, F.G. & Lugo, A.E. (1986) Ecology of tropical dry forest. *Annual Review of Ecology and Systematics*, **17**, 67–88.
- Myers, A.A. & Giller, P.S. (1988) *Analytical biogeography: an integrated approach to the study of animal and plant distributions*. Chapman and Hall, London.
- Noonan, B.P. & Gaucher, P. (2006) Refugial isolation and secondary contact in the dyeing poison frog *Dendrobates tinctorius*. *Molecular Ecology*, **15**, 4425–4435.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D’Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. (2001) Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience*, **51**, 933–938.
- Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.
- Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T.P. & Lees, D.C. (2006) Model-based uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704–1711.
- 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., Lavin, M., Prado, D.E., Pendry, C.A., Pell, S.K. & Butterworth, C.A. (2004) Historical climate change and speciation: Neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 515–538.
- Pennington, R.T., Lewis, G.P. & Ratter, J.A. (2006a) An overview of the plant diversity, biogeography and conservation of Neotropical savannas and seasonally dry forests. *Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation* (ed. by R.T. Pennington, G.P. Lewis and J.A. Ratter), pp. 1–29. CRC Press, Boca Raton, FL.
- Pennington, R.T., Lewis, G.P. & Ratter, J.A. (eds) (2006b) *Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation*. CRC Press Taylor, 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 and Systematics*, **40**, 437–457.
- Phillips, S.J. & Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.
- Porto, K.C., Cabral, J.J.P. & Tabarelli, M. (2004) *Brejos de altitude em Pernambuco e Paraíba: história natural, ecologia, e conservação*. Ministério do Meio Ambiente, Brasília, DF.
- Prado, D. (1993a) What is the Gran Chaco vegetation in South America? I. A review. Contribution to the study of flora and vegetation of the Chaco. *Candollea*, **48**, 145–172.
- Prado, D. (1993b) What is the Gran Chaco vegetation in South America? II. A redefinition. Contribution to the study of flora and vegetation of the Chaco. *Candollea*, **48**, 615–629.
- Prado, D.E. (2000) Seasonally dry forests of tropical South America: from forgotten ecosystems to a new phyto-geographic unit. *Edinburgh Journal of Botany*, **57**, 437–461.
- Prado, D.E. (2003) As Caatingas da América do Sul. *Ecologia e conservação da Caatinga* (ed. by I.R. Leal, M. Tabarelli and J.M.C. Silva), pp. 3–73. Editora Universitária UFPE, Recife.
- Prado, D.E. & Gibbs, P.E. (1993) Patterns of species distributions in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden*, **80**, 902–927.
- Queiroz, L.P. (2006) The Brazilian Caatinga: phytogeographical patterns inferred from distribution data of the Leguminosae. *Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation* (ed. by R.T. Pennington, G.P. Lewis and J.A. Ratter), pp. 121–157. CRC Press, Boca Raton, FL.
- Richards, C.L., Carstens, B.C. & Knowles, L.L. (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *Journal of Biogeography*, **34**, 1833–1845.
- Rissler, L.J. & Apodaca, J.J. (2007) Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Systematic Biology*, **56**, 924–942.
- Rull, V. (2008) Speciation timing and Neotropical biodiversity: the Tertiary–Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology*, **17**, 2722–2729.
- Smith, T.B., Kark, S., Schneider, C.J., Wayne, R.K. & Moritz, C. (2001) Biodiversity hotspots and beyond: the need for preserving environmental transitions. *Trends in Ecology and Evolution*, **16**, 431.
- Soria-Auza, R.W. (2009) *Diversity and biogeography of ferns and birds in Bolivia: applications of GIS based modelling approaches*, p. 129. Faculties of Mathematics and Naturewissenschaften, Georg-August Universität Göttingen, Göttingen.
- Swenson, N.G. (2008) The past and future influence of geographic information systems on hybrid zone, phylogeographic and speciation research. *Journal of Evolutionary Biology*, **21**, 421–434.
- Swets, J.A. (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293.
- Vanzolini, P.E. (1974) Ecological and geographical distribution of lizards in Pernambuco, northeastern Brazil (Sauria). *Papéis Avulsos de Zoologia, São Paulo*, **28**, 61–90.
- Velloso, A.L., Sampaio, E.V.S.B. & Pareyn, F.G.C. (2002) *Ecorregiões propostas para o bioma Caatinga*. Associação Plantas do Nordeste – Instituto de Conservação Ambiental – The Nature Conservancy do Brasil, Recife.
- Waltari, E., Hijmans, R.J., Peterson, A.T., Nyári, A.S., Perkins, S.L. & Guralnick, R.P. (2007) Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE*, **7**, e563.
- Werneck, F.P. & Colli, G.R. (2006) The lizard assemblage from seasonally dry tropical forest enclaves in the Cerrado biome, Brazil, and its association with the Pleistocenic arc. *Journal of Biogeography*, **33**, 1983–1992.
- Zanella, F.C. (2000) The bees of the Caatinga (Hymenoptera, Apoidea, Apiformes): a species list and comparative notes regarding their distribution. *Apidologie*, **31**, 579–592.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Overview of Quaternary climatic change across the Neotropics.

**Appendix S2** South American Seasonally Dry Tropical Forest distribution according to Miles *et al.* (2006) (in green).

**Appendix S3** Estimated distribution of South American Seasonally Dry Tropical Forests under (a) narrower and (b) broader definitions (see text for details regarding the development of the maps).

**Appendix S4** Description of methods used to determine the distribution extent of Seasonally Dry Tropical Forests in eastern-central South America.

**Appendix S5** Tables depicting comparisons between published pollen records and model predictions for Seasonally Dry Tropical Forest occurrence at 6 kyr BP and 21 kyr BP, respectively.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.



## BIOSKETCH

**Fernanda P. Werneck** is a Brazilian PhD student at Brigham Young University. She is interested in the diversification, biogeography, phylogeography, niche modelling and conservation genetics of the South American dry biomes, with emphasis on associated reptiles.

Editor: Thomas Gillespie