







Global bioregionalization of warm drylands based on tree assemblages mined from occurrence big data

Manuel Cartereau^{1*} , Agathe Leriche¹ , Alex Baumel¹ ,
Ian Ondo² , Cyrille Chatelain³ , James Aronson^{4,5}  and
Frédéric Médail¹ 

¹ Aix Marseille Univ, Avignon Université, CNRS, IRD, IMBE, Aix-en-Provence, France

² Royal Botanic Gardens, Kew, Biodiversity Informatics and Spatial Analysis Department, Richmond, Surrey, TW9 3AB, UK

³ Conservatoire et Jardin Botaniques de la ville de Genève, Geneva, Switzerland

⁴ Center for Conservation and Sustainable Development, Missouri Botanical Garden, St. Louis, MO, USA

⁵ Ecological Health Network, Boston, MA, USA

*Correspondence: Manuel Cartereau, manuel.cartereau@imbe.fr

Abstract

Drylands represent about 41% of Earth's land area, host more than 1,500 tree species and support more than 20% of the world's human population. Trees are key to the functioning of numerous dryland ecosystems and contribute to goods and services for many local human communities, but many are threatened by global changes. From this perspective, mapping tree species assemblages of drylands can provide valuable information for conservation. To our knowledge, warm drylands, including hot deserts, have never been subject to a comprehensive tree biodiversity analysis independent of administrative boundaries or pre-defined regions. Our study aimed to address this gap by redefining warm drylands based on climate data and delineating bioregions using tree species assemblages at the global scale. We based the analyses on aridity and temperature data and a co-occurrence network approach using more than 1,000 tree species. Our data are mined from the Desert Trees of the World database, the Global Biodiversity Information Facility database, and the African Plant Database. This new delimitation of warm drylands reveals eight bioregions, covering about 19% of Earth's land area across all continents. These are: North America, two bioregions in South America, the southern Mediterranean Basin and Macaronesian islands, the Saharo-Sindian region and the Horn of Africa, Southern Africa, the Socotra archipelago, and Australia. These bioregions have very distinct tree species assemblages, as well as high rates of endemism. This original diversity is found under a wide range of aridity conditions both within and between bioregions, offering the opportunity to anticipate different responses of tree assemblages face to future climate change among the world's warm drylands. It will aid in conservation, restoration, and rehabilitation strategies involving the use of native trees among the most threatened regions worldwide.

Highlights

- Analysis of a dataset of 127,322 occurrence records corresponding to 1,092 tree species delimited eight bioregions in warm drylands of the world.
- These eight distinct bioregions cover about 19% of Earth's land area across all continents.
- Tree species diversity is not evenly distributed among bioregions (from 28 to 360 species) which host high proportions of endemic tree species (from 26 to 97%).
- Aridity conditions are highly variable within and among the eight bioregions, providing the opportunity to anticipate various challenging ecological conditions under future climate change.

Keywords: aridity, biodiversity, co-occurrence network, conservation, desert

Introduction

Drylands cover about 41% of Earth's land area and encompass a variety of ecosystem types such as deserts, savannas, steppes and other types of grasslands, tundra, and open woodlands (FAO 2019). They are defined as areas where the ratio of annual precipitation to mean annual potential evapotranspiration, referred as the Aridity Index (AI), is equal to or less than 0.65 (UNEP 1997). Despite their huge spatial extent, drylands' biodiversity remains poorly known (Davis et al. 2012) and the most recent biogeographical synthesis of global drylands offered no attempt at bioregionalization (Maestre et al. 2021).

In dry environments, water availability highly constrains plant productivity and growth, influencing plant diversity, ecosystem structure and composition. Despite limited access to water resources, many drylands are composed of mosaics of open forests, shrublands and grasslands. Numerous trees can even cope with hyper-arid climate in the Sahara (Médail and Quézel 2018, Brandt et al. 2020). In fact, drylands harbour a surprisingly high tree biodiversity, estimated by Aronson et al. (2019)¹ to be more than 1,500 species belonging to 423 genera and 100 families. The role of trees in vegetation dynamics in drylands has long been a matter of heated debate (Veldman 2016, Fagan 2020) and grassy ecosystems are sometimes misidentified as degraded forest (Kumar et al. 2020). However, wooded areas, even sparsely covered by trees, are a key component of numerous drylands ecosystems and they play a pivotal role in desertification mitigation (Smith et al. 2019). For example, Bastin et al. (2017) estimated that 21% of drylands were covered by wooded areas (tree canopy $\geq 10\%$).

Out of 1,126 species evaluated by the IUCN in Aronson et al.'s checklist, 208 (i.e., 18%) are threatened with extinction and 309 (i.e., 27%) have declining population trends (IUCN Red List 2022²). Besides, in such challenging environments, warmer and drier climate in the future, as a global trend (IPCC 2018), might have an important impact on dryland ecosystems (Mirzabaev et al. 2019). Also, anthropogenic climate change results in more frequent hotter droughts, leading to increasing stresses for tree vegetation (Hammond et al. 2022). We chose to focus this study on warm and very dry areas (i.e., hyper-arid and arid bioclimates, see below) because in these areas, some trees might already be close to their upper climatic tolerance limits (see Lancaster and Humphreys 2020, Santinella et al. 2020). Considering the importance of trees in dryland ecosystems, the potential threats they will face under future climate change, the essential role they play in sustaining human livelihoods, and the knowledge gaps on their spatial organization, quantifying and describing broad spatial pattern of tree species diversity and assemblages is a first step

to future research and action. Indeed, planning for ecosystem and biodiversity conservation, ecological restoration, and sustainable use, requires delimiting geographically broad units to organize, synthesize, and compare information from many regions of the world. For instance, the FAO's recent report (FAO 2019) evaluates "Trees, forests and land use in drylands", at the global scale, using the geoscheme of the United Nations Statistics Division (UN 2019³). However, because this zonation is based on administrative boundaries and not on any biogeographic pattern, it might not be relevant to tackle the diversity of woody ecosystems in global drylands and group them in a higher hierarchical level.

Bioregionalization, which aims to "reflect a range of biological, physical and ecological phenomena" (Mackey et al. 2007) is a prominent approach to biodiversity sciences such as biogeography, macroecology, and conservation biology, for summarizing the diversity of ecological and evolutionary patterns at large spatial scales (Mackey et al. 2007, Montalvo-Mancheno et al. 2020). Increased access to large-scale geospatial datasets (e.g., georeferenced species occurrence records available on the online repository GBIF, Global Biodiversity Information Facility: <https://www.gbif.org/>), and the advent of new statistical methods and tools that are user-friendly and readily accessible online (e.g., Infomap Bioregions: Edler et al. 2017, Yusefi et al. 2019) have allowed researchers to use bioregionalization to characterize species assemblages across spatial scales.

In this study, we aimed to delineate bioregions in warm drylands at the global scale, based on tree species assemblages and using large species occurrence databases. To achieve our objective, we spatially defined warm drylands based on climate data, then, we collated occurrences of tree species obtained from online databases and used the Infomap clustering algorithm (Edler et al. 2017), based on a co-occurrence network approach, to delineate bioregions.

Materials and Methods

Three steps were implemented to delineate bioregions in warm drylands based on trees occurrences: (i) defining and mapping warm drylands based on climate data, (ii) retrieving and curating tree species georeferenced occurrence records from three reliable online databases, within warm drylands and (iii) running a bioregionalization analysis based on tree species assemblages. Data manipulations and spatial analysis were performed with R v4.0.3 and QGIS v.3.4 software.

Defining warm drylands

Contrary to previous attempts to delineate warm drylands climatic envelope using temperature and precipitation data (Köppen 1900, Whittaker 1970), we based our climatic envelope on the Aridity

¹ Aronson, T., Le Floc'h, E., Aronson, J. Desert Trees of the World Database, Tropicos [online]. <http://legacy.tropicos.org/Project/Desert%20Trees>, accessed in 2019

² IUCN Red List [online]. <https://www.iucnredlist.org/>, accessed in 2022

³ UN Geographic regions [online]. <https://unstats.un.org/unsd/methodology/m49/>, accessed in 2019

Index because it considers both atmospheric water input (*via* precipitations) and output (*via* potential evapotranspiration). Thus, it gives a better estimation of the actual water availability for the vegetation to grow than variables based only on temperatures and precipitations. We defined drylands using a Global Aridity Index (AI) map (CGIARCSI, <https://cgiarcsi.community>; Trabucco and Zomer 2018) for deriving hyper-arid ($AI < 0.03$) and arid ($0.03 < AI < 0.2$) maps following the UNEP classification (UNEP 1997).

To discriminate warm from cold areas within drylands, we first removed areas at least partially covered by ice sheets where trees cannot grow (i.e., above 66°N and below 60°S latitudes). Then, to exclude cold biomes (e.g., tundra, high altitude steppes and deserts) we used the mean temperature of coldest quarter (BIO11; CHELSA: Karger et al. 2017) which showed a bimodal distribution in hyper arid and arid environments (Fig. S1), allowing to identify a threshold between warm and cold dry climates. We preferred using the mean temperature of coldest quarter (rather than the mean annual temperature) because it gives a better idea of minimal temperatures, and thus to cold stress, averaged on a relevant time scale for trees considering their life span (rather than minimum temperature of coldest month). To do so, we ran the k-means clustering algorithm (with the number of clusters k set to 2) on BIO11 in hyper-arid and arid areas. It revealed that warm and dry climate is situated above 3.8°C. Then, we removed the cold areas where $BIO11 \leq 3.8^\circ C$. All maps were obtained at a 30 arc-seconds resolution.

Tree occurrence data collection

We used the Desert Trees of the World Database (Aronson et al. 2019) as a species checklist. This database includes 1,574 species of trees, i.e., long-lived plants that develop one sturdy long-lived trunk equal or above one meter in height (Shreve and Wiggins 1951, Felger et al. 2001, Schatz 2001), growing in all drylands around the world. We obtained occurrence data for these species from GBIF⁴ (2020) and from the African Plant Database (2020⁵) to fill important data gaps in GBIF, especially for the Sahara Desert. Then, we cleaned the dataset using GBIF metadata, and CoordinateCleaner R package (Zizka et al. 2019) (see Appendix 1). Occurrence records falling outside of warm drylands area (see 2.1.) were removed.

Bioregionalization

To delimitate bioregions, we used the Infomap Bioregions algorithm (Edler et al. 2017). This interactive web application uses an adaptive spatial grid to collate species occurrences into assemblages, i.e., a spatial grid whose local resolution adapts to records spatial

density to make grid cells contain a comparable amount of occurrence records (larger cell size when occurrence records density is low, smaller cell size when occurrence records density is high). This is advantageous when data are unevenly distributed, as in the case of warm drylands trees. Then, the infomap algorithm uses these assemblages to define bioregions: it builds a bipartite network between species and grid cells by connecting grid cells through common species, which is then used to cluster grid cells in bioregions. In other terms, grid cells sharing similar species assemblages are assigned to the same bioregion and an indicative species score is calculated (Edler et al. 2017) to inform the most indicative taxa in each bioregion. Infomap does not implement any constraint for spatial contiguity. This method has proven to outperform methods based on species composition similarities between grid-cells because being less sensitive to species sampling biases (Vilhena and Antonelli 2015).

The tree occurrence dataset was analysed a first time with the Infomap Bioregions algorithm to assign species to bioregion(s). Then, we manually checked whether each species was native in each bioregion to which it was assigned by the algorithm, by consulting the Kew Plant of the World Online portal⁶ (2021), and additional literature (Ghazanfar 2003-2017, Miller and Morris 2004, Arbonnier 2009, African Plant Database 2020, Euro+Med Plant Base 2020⁷). We filtered out records occurring outside of the native area of the species to perform the bioregionalization design only on occurrences of native species. Twenty-four species had no native occurrence record in our dataset and thus were removed from our study list.

The Infomap Bioregions algorithm was then run on the native occurrence dataset. After preliminary analysis, the algorithm was parametrized to allow a cell size from 2 to 4° and prioritize a number of occurrence records from 10 to 100 and a number of species from 2 to 100 per grid cell. Also, we calculated the percentage of endemic species for each bioregion, i.e., the percentage of species occurring only in the given bioregion

Depending on data amount and spatial density, as well as species diversity distribution, several solutions, differing by the number and the size of bioregions, can be obtained. Infomap handles this problem by applying a cost parameter when clustering the bipartite network, which constrains the number of resulting bioregions. We ran the algorithm setting the cost for the number of clusters from 1 (high number of clusters, i.e., bioregions, preferred) to 3 (low number of clusters preferred) with a 0.25 increment. The number of bioregions stopped varying from a cost of 2.5. As the main objective is the delimitation of bioregions at a global scale, we chose to present here the results

4 Global Biodiversity Information Facility [online]. <https://www.gbif.org/>, accessed 29 July 2020

5 African Plant Database v3.4.0 [online]. Conservatoire et Jardin botaniques de la Ville de Genève and South African National Biodiversity Institute, Pretoria. <http://africanplantdatabase.ch>, accessed in 2020

6 Plants of the World Online [online]. Facilitated by the Royal Botanic Gardens, Kew. <http://www.plantsoftheworldonline.org/>, accessed in 2021

7 Euro+Med PlantBase - the information resource for Euro-Mediterranean plant diversity [online]. <http://www2.bgbm.org/EuroPlusMed/>, accessed in 2020

with a cost of 2.5. However, to appreciate consistency of the bioregions obtained with the Infomap Bioregions algorithm using a cost of 2.5, we also present results from the Infomap bioregionalization with cost set to 1 and 3 in Appendix 2, together with a bioregionalization using a hierarchical clustering method based on species composition dissimilarities between Infomap grid cells (using the `hclust` and `dist` binary R function with Jaccard index). We truncated the dendrogram to obtain 8 bioregions to ensure comparison with the results from Infomap presented here (Appendix 2).

Results

World warm drylands

Selecting areas where AI ranges from 0 to 0.2 and BIO11 is above 3.8°C (with a maximum of 29.1°C), allows to clearly identify the warm drylands on all continents (except Antarctica). The total area of the warm drylands is 27,519,770 km² of which 63% and 37% are arid and hyper-arid, respectively. This area represents 19% of global land surfaces (including Antarctica) (Fig. 1).

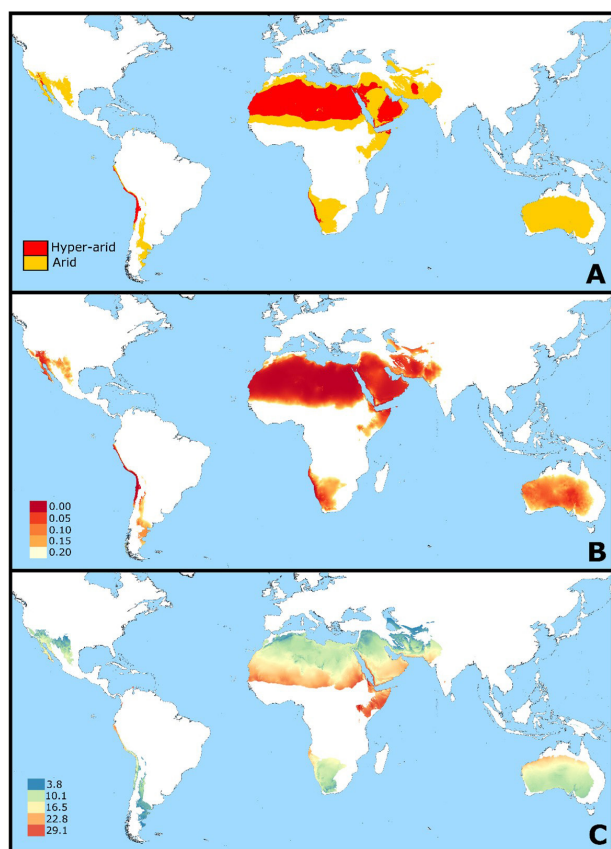


Figure 1. Maps (CRS EPSG 4326) of warm hyper-arid and arid areas (A), Aridity Index unitless (B, low values mean drier conditions), and mean temperature of coldest quarter in °C (C).

Bioregions

The final dataset (i.e., native occurrences of tree species) included 127,322 occurrence records corresponding to 1,092 species ((min) mean \pm sd (max): (1) 117 \pm 220 (3,182) records per species). Of the 1,574 tree species initially present on the checklist (Aronson et al. 2019), 482 were excluded either because they do not grow in warm and hyper-arid or arid areas ($n=377$), because they were not informed by any valid occurrence record, considering our cleaning procedure ($n=59$), or because they were synonyms ($n=46$). Of the 1,092 species remaining, 8 species were not attributed to any bioregion by the algorithm because their occurrence records belong to grid cells with less than 10 occurrence records and/or less than 2 species. The number of records (from 10 to 6,001; Fig. 2A) and the number of species (from 2 to 99; Fig. 2B), are presented in each grid cell generated by the Infomap Bioregions algorithm.

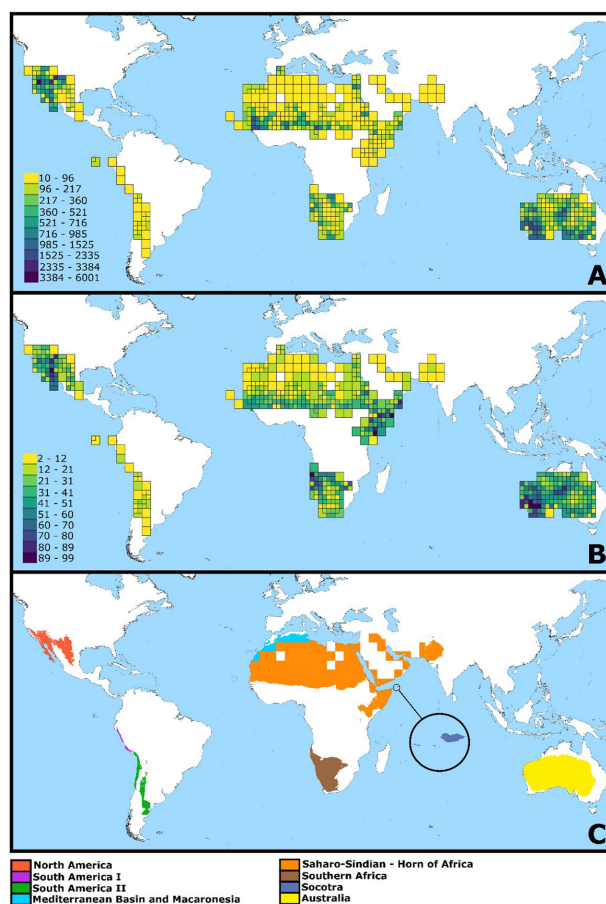


Figure 2. Maps (CRS EPSG 4326) of tree occurrence records number (A), tree species richness (B), and bioregions based on tree species assemblages in warm drylands (C), with a zoom on the Socotra bioregion for better visualization. The maps (A) and (B) are based on the adaptive resolution grid generated by Infomap Bioregions. The map (C) is based on the adaptive resolution grid generated by Infomap Bioregions, masked to warm drylands.

With a cluster cost set to 2.5, our study reveals the existence of 8 bioregions within the warm drylands area: North America (including Great Basin, Mojave, Sonora and Chihuahua Deserts), two regions in South America (I: Galapagos archipelago and the Sechura Desert, and II: the Atacama and Monte Deserts), the southern Mediterranean Basin (along with Andalusia) and Macaronesian islands, the Saharo-Sindian region and the Horn of Africa, Southern Africa (including Namib, Namaqualand, Kalahari, and Karroo Deserts), the Socotra archipelago, and Australia (including all desertic areas) (Fig. 2C, Appendix 3 and 4). When using a cost for the number of clusters set to 1, we found in addition to the 8 bioregions above, 5 small subdivisions, in the Middle East, the Horn of Africa and North America. With

a cost for the number of clusters set to 3, bioregions are grouped in 4 large divisions corresponding to the New World, Old World, Australia, and Socotra archipelago (Appendix 2). Despite some boundaries inconsistencies, the bioregionalization resulting from the dissimilarity matrix method revealed similar bioregions except for: Australia in which two bioregions are found (West versus East), the Southern Africa bioregion with a small extension in the South of the Horn of Africa, and the West-East split of the Mediterranean Basin – Macaronesia and the Saharo-Sindian – Horn of Africa (including Socotra) bioregions (Appendix 2).

The 8 retained bioregions have a very differentiated tree species composition even when considering only their 5 most abundant taxa (Table 1). Among the

Table 1. Tree species richness, number (and percentage) of endemic tree species of the 8 bioregions obtained for warm drylands, top 5 of the most abundant (n = number of records) and top 5 of the most indicative species (indicative score) of each bioregion.

Bioregion	Species richness	Endemic species richness	Top 5 most abundant species (n)		Top 5 most indicative species (score)	
North America	232	211 (91%)	<i>Malosma laurina</i>	(571)	<i>Yucca faxoniana</i>	(5.6)
			<i>Platanus racemosa</i>	(537)	<i>Castela emoryi</i>	(5.6)
			<i>Psoralea argemone</i>	(495)	<i>Malosma laurina</i>	(5.6)
			<i>Oleña tesota</i>	(485)	<i>Platanus racemosa</i>	(5.6)
			<i>Cylindropuntia bigelovii</i>	(484)	<i>Psoralea argemone</i>	(5.6)
South America I	28	9 (32%)	<i>Cordia lutea</i>	(75)	<i>Myrcianthes ferreyrae</i>	(42.4)
			<i>Opuntia echinos</i>	(65)	<i>Ceiba trischistandra</i>	(42.4)
			<i>Jasminocereus thouarsii</i>	(31)	<i>Escallonia resinosa</i>	(42.4)
			<i>Bursera graveolens</i>	(16)	<i>Erythrina velutina</i>	(42.4)
			<i>Neoraimondia arequipensis</i>	(14)	<i>Jasminocereus thouarsii</i>	(42.4)
South America II	52	34 (65%)	<i>Geoffroea decorticans</i>	(77)	<i>Escallonia angustifolia</i>	(41.3)
			<i>Zuccagnia punctata</i>	(66)	<i>Senegalia gilliesii</i>	(41.3)
			<i>Prosopis flexuosa</i>	(63)	<i>Cordia decandra</i>	(41.3)
			<i>Parkinsonia praecox</i>	(62)	<i>Trichocereus atacamensis</i>	(41.3)
			<i>Cordia decandra</i>	(52)	<i>Eulychnia acida</i>	(41.3)
Mediterranean Basin - Macaronesia	43	11 (26%)	<i>Ziziphus lotus</i>	(235)	<i>Dracaena tamaranae</i>	(13.5)
			<i>Dracaena draco</i>	(156)	<i>Sideroxylon canariense</i>	(13.5)
			<i>Tamarix canariensis</i>	(125)	<i>Vachellia gummifera</i>	(13.5)
			<i>Phoenix dactylifera</i>	(87)	<i>Juniperus turbinata</i>	(13.5)
			<i>Argania spinosa</i>	(81)	<i>Tamarix ramosissima</i>	(13.5)
Saharo-Sindian - Horn of Africa	278	177 (64%)	<i>Vachellia tortilis</i>	(3077)	<i>Blepharispernum hirtum</i>	(1)
			<i>Balanites aegyptiaca</i>	(2957)	<i>Boswellia frereana</i>	(1)
			<i>Maerua crassifolia</i>	(1812)	<i>Cadaba baccarinii</i>	(1)
			<i>Boscia senegalensis</i>	(1743)	<i>Prosopis cineraria</i>	(1)
			<i>Commiphora africana</i>	(1425)	<i>Commiphora rostrata</i>	(1)
Southern Africa	171	106 (62%)	<i>Vachellia erioloba</i>	(423)	<i>Diospyros lycioides</i>	(7.5)
			<i>Aloidendron dichotomum</i>	(394)	<i>Vachellia hebeclada</i>	(7.5)
			<i>Boscia albitrunca</i>	(356)	<i>Vachellia luederitzii</i>	(7.5)
			<i>Nymania capensis</i>	(321)	<i>Commiphora glaucescens</i>	(7.5)
			<i>Combretum imberbe</i>	(276)	<i>Moringa ovalifolia</i>	(7.5)

Table 1. Continued...

Bioregion	Species richness	Endemic species richness	Top 5 most abundant species (n)		Top 5 most indicative species (score)	
Socotra	74	56 (76%)	<i>Boswellia socotrana</i>	(24)	<i>Boswellia socotrana</i>	(132.6)
			<i>Euclea divinorum</i>	(21)	<i>Croton socotranus</i>	(132.6)
			<i>Boswellia elongata</i>	(19)	<i>Acridocarpus socotranus</i>	(132.6)
			<i>Croton socotranus</i>	(18)	<i>Jatropha uncostata</i>	(132.6)
			<i>Acridocarpus socotranus</i>	(17)	<i>Dirachma socotrana</i>	(132.6)
Australia	360	350 (97%)	<i>Eucalyptus largiflorens</i>	(498)	<i>Acacia pruinocarpa</i>	(6.4)
			<i>Acacia papyrocarpa</i>	(477)	<i>Flindersia maculosa</i>	(6.4)
			<i>Acacia burkittii</i>	(459)	<i>Archidendropsis basaltica</i>	(6.4)
			<i>Corymbia opaca</i>	(447)	<i>Acacia marramamba</i>	(6.4)
			<i>Eucalyptus salubris</i>	(445)	<i>Eucalyptus salubris</i>	(6.4)

8 bioregions, Australia hosts the highest number of tree species (360) and the highest percentage of endemic tree species (97%). The lowest species richness is found in South America I (28), whereas the Mediterranean-Macaronesian region is characterized by the lowest proportion of endemic species (26%, i.e., 11 species). The Socotra archipelago is remarkable for its high number of tree species (74) for such a small area (the smallest bioregion), with a high rate of endemism of 76%. The Saharo-Sindian - Horn of Africa bioregion, the largest, is the second species richest bioregion (278) with a rate of endemism of 64% (Table 1).

Also, aridity conditions appear to be highly variable within and among the eight bioregions (Fig. 3). Indeed, while some bioregions are more arid than others, most of them show a wide aridity gradient, sometimes with a bimodal distribution of the aridity index, such as the South America II bioregion where both hyper-arid (Atacama Desert) and arid (Monte Desert) climates occur.

Discussion

To our knowledge, this is the first attempt to carry out a tree species-based bioregionalization of warm drylands at the global scale. We focused this approach on warm to hot arid and hyper-arid areas and found that our warm drylands distribution is comparable to the bioclimate distribution “BWh: Arid, desert, hot” in the Kopper-Geiger classification, which is based on temperature and precipitation data, with the exception of South America II bioregion, classified as “BWk: Arid, desert, cold”, i.e., with a mean annual temperature < 18°C (Beck et al. 2018). Also, some areas often referred as warm drylands do not appear on our maps such as the Southwestern part of Madagascar classified as “semi-arid” based on the aridity index. The extreme Northern parts of Colombia and Venezuela have relatively small areas which do fall in the warm drylands climatic envelope but are very poorly documented in occurrence records of our species list. However, all the main warm drylands commonly recognized are represented in our study.

We recognize that gathering occurrence data from large online databases may introduce various sources of uncertainty together with taxonomic and geographic biases inherited from the opportunistic nature of the data collection (Meyer et al. 2015). For instance, more data are available for North America and Australia compared to South America, the Sahara, or Asian deserts (fig. 2A). Whether a low number of occurrence records for a given species reflects a rarity of the species may be uncertain, therefore a low number of tree species within a grid cell may not reflect a low species richness. Nevertheless, it seems that occurrence records sampling size is not necessarily associated with species richness in our study. For example, both the Horn of Africa and the central area of Sahara have scarce data while cell grids show a relatively high species richness for the former region, and a low species richness for the latter (fig. 2A and B). Indeed, it is well known that the Sahara Desert, which constitutes the largest hyper-arid continuous area in the world is represented by a relatively low number of occurrence records. However, its strong inaccessibility resulted in few botanical explorations, and most of the data are old and still included in grey literature (Médail and Quézel 2018), thus any plant species richness estimates for the Sahara Desert must be considered with caution.

Our results revealed a clear geographic structure with no bioregion split between disjunct areas or intertwined with any other one. This is the result of distinctive species compositions among these bioregions. For instance, despite its proximity with the very large Saharo-Sindian – Horn of Africa region and its continental origin (Miller and Morris 2004), the Socotra archipelago exhibits a remarkably high rate of endemism, i.e., 76% (37% for the total vascular flora, Banfield et al. 2011), and appears as a single and coherent bioregion. Its tree species composition is so particular that Socotra still constitutes a single bioregion when a strong constraint on the number of resulting bioregions is implemented, and its top five most common and indicative species are all Socotran endemic trees, apart from *Euclea divinorum* Hiern.

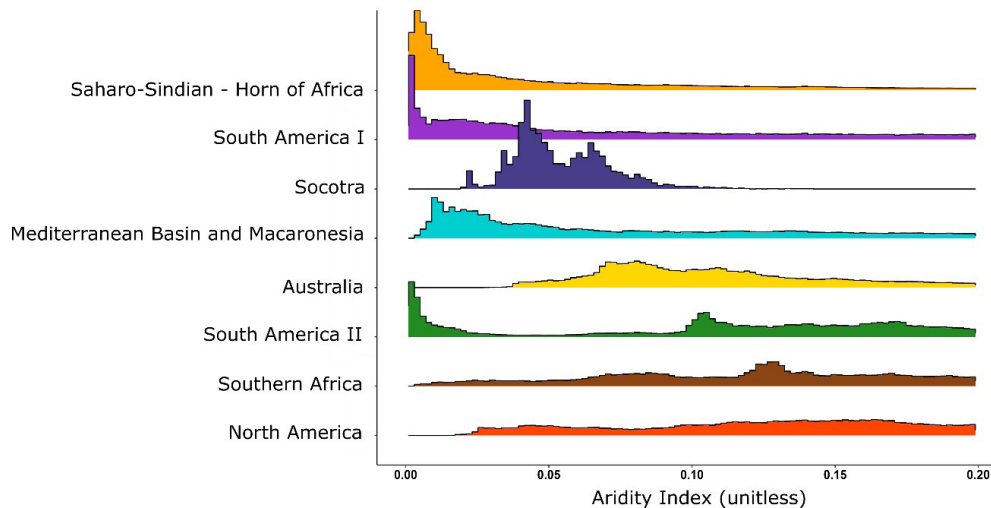


Figure 3. Aridity Index distribution of the eight tree assemblage-based bioregions in warm drylands ordered by (from top to bottom) the median value of the Aridity Index (left to right: driest to less dry conditions).

Among its endemic flora, several tree species display outstanding adaptations to aridity such as succulence, bottle tree, pachycaul and dracoid growth forms. Socotra's flora has multiple biogeographic affinities, show contrasted evolutionary patterns (i.e., *in situ* radiations, anagenesis and anacladogenesis) but the vicariance *versus* dispersal paradigm regarding its endemic flora remains widely unsolved (Banfield et al. 2011).

Although the broad biogeographic pattern we present here can be recognized in the bioregionalization resulting from the dissimilarity matrix method, some inconsistencies persist. However, it is known that bipartite network approaches outperform approaches based on dissimilarity matrix for the aim of bioregionalization (Vilhena and Antonelli 2015). The bioregions we present here provides a comparative framework to address the complex evolutionary and ecological history of tree assemblages in warm drylands. Still, bioregionalization can vary depending on the group of organisms, the environments considered, and whether phylogenetic relationships are considered. For example, in succulent plants, transcontinental geographic disjunctions among closely related taxa are frequent and are likely to be caused by long distance dispersal events (Gagnon et al. 2019). Although there is a strong biome conservatism in plant evolution, several dispersal events occur on the long term from sclerophyll to arid biomes, suggesting that drylands play a strong sink role (Crisp et al. 2009). At finer phylogenetic scales, studies also support the colonization of dry ecosystems by mesic lineages (Zizka et al. 2020) or sclerophyllous lineages pre-adapted to xeric conditions (Vásquez-Cruz and Sosa 2019), followed by *in situ* diversification mainly from the mid-late Miocene to the Early Pliocene (Wu et al. 2018).

Besides, this new bioregionalization presented here shows some discrepancies with the *Terrestrial ecoregions of the World* map (Olson et al. 2001): the very large Saharo-Sindian – Horn of Africa

region overlaps with no less than three different biogeographic realms, supposed to include units from lower hierarchical levels: Palearctic, Afrotropic, and Indo-Malay. We acknowledge that, at a finer scale, our delineation of bioregions might be debatable. This is mainly caused by reduced local data availability or the occurrence of non-native species records difficult to spot in large datasets. For example, the Cape Verde archipelago belongs to both the Mediterranean-Macaronesian and the Saharo-Sindian – Horn of Africa regions whereas it constitutes a single ecoregion in Olson et al. 's map (2001), the “Cape Verde Islands dry forests”. Nonetheless, the impact of such issues on downstream analysis (e.g., ancestral states reconstruction in phylogenies, comparison of various biological variables across bioregions etc.) highly depends on the research question and spatial scale. Furthermore, our new bioregionalization scheme fairly reflects the autochthonous distribution areas of the studied species (see Table S1).

As biogeographic processes impact differently across spatial scales and groups of organisms, delineating bioregions is an important first step for the study of biodiversity spatial patterns and species assemblages responses to global changes. A wide range of aridity conditions is observed within and between bioregions, providing the opportunity to anticipate different responses of tree assemblages to future climate change. Our approach makes use of available geographical data, i.e., occurrence records and climate spatial grids from online databases, and a user-friendly bioregionalization web application, to delimit and identify tree biodiversity patterns associated with warm drylands at the global scale. Such an approach will offer the possibility to undertake biogeographic comparative analyses on the diversity and ecology of warm drylands trees. This bioregionalization can also improve our understanding of the biogeographic processes and species' responses that shaped the distribution of tree diversity at global scale in challenging ecological conditions. Furthermore, this

new map of tree assemblages in warm drylands could help inform and prioritize management and conservation actions in these environments particularly threatened by global changes by providing relevant biogeographic units.

Acknowledgments

We are grateful to the two anonymous reviewers for their comments on the earlier versions of the manuscript.

Author Contributions

Manuel Cartereau: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration.

Agathe Leriche: Conceptualization, Methodology, Writing - Review & Editing, Supervision.

Alex Baume: Conceptualization, Validation, Writing - Review & Editing, Supervision.

Ian Ondo: Methodology, Software, Writing - Review & Editing.

Cyrille Chatelain: Data Curation, Writing - Review & Editing.

James Aronson: Validation, Data Curation, Writing - Review & Editing.

Frédéric Médail: Conceptualization, Validation, Writing - Review & Editing, Project administration, Supervision.

Supplementary Material

The following materials are available as part of the online article at <https://escholarship.org/uc/fb>

Figure S1. Distribution of the mean temperature of the coldest quarter of the year within hyper-arid and arid areas

Appendix 1. Retrieving and cleaning procedure of occurrence data

Appendix 2. Alternative bioregionalizations of warm drylands based on tree species assemblages

Appendix 3. Bioregions shapefile

Appendix 4. Bioregions masked to warm drylands raster layer

Table S1. Species/Bioregions matrix

References

- Arbonnier, M. (2009) Arbres, arbustes et lianes des zones sèches d'Afrique de l'Ouest. 3rd edn. Quae & MNHN, Paris.
- Banfield, L.M., Van Damme, K., Miller, A.G. (2011) Evolution and biogeography of the flora of the Socotra archipelago (Yemen). In: The biology of island floras (ed. by D. Bramwell and J. Caujapé-Castells). Cambridge University Press, New York.
- Bastin, J.-F., Berrahmouni, N., Grainger, A., et al. (2017) The extent of forest in dryland biomes. *Science*, 6338, 635-638. <https://doi.org/10.1126/science.aam6527>
- Beck, H.E., Zimmermann, N.E., McVicar, T.R., et al. (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, 5, 180214. <https://doi.org/10.1038/sdata.2018.214>
- Brandt, M., Tucker, C.J., Kariryaa, A., et al. (2020) An unexpectedly large count of trees in the West African Sahara and Sahel. *Nature*, 587, 78-82. <https://doi.org/10.1038/s41586-020-2824-5>
- Crisp, M.D., Arroyo, M.T.K., Cook, L.G., et al. (2009) Phylogenetic biome conservatism on a global scale. *Nature*, 458, 754-756. <https://doi.org/10.1038/nature07764>
- Davies, J., Poulsen, L., Schulte-Herbrüggen, B., et al. (2012) Conserving Dryland Biodiversity, IUCN, Nairobi.
- Edler, D., Guedes, T., Zizka, A., et al. (2017) Infomap Bioregions: interactive mapping of biogeographical regions from species distributions. *Systematic Biology*, 66, 197-204. <https://doi.org/10.1093/sysbio/syw087>
- Fagan, M.E. (2020) A lesson unlearned? Underestimating tree cover in drylands biases global restoration maps. *Global Change Biology*, 9, 4679-4690. <https://doi.org/10.1111/gcb.15187>
- FAO (2019) Trees, forests and land use in drylands: the first global assessment - Full report, FAO Forestry Paper No. 184, Rome.
- Felger, R.S., Johnson, M.B., Wilson, M.F. (2001) The trees of Sonora, Mexico. Oxford University Press, New York.
- Gagnon, E., Ringelberg, J.J., Bruneau, A., et al. (2019) Global succulent biome phylogenetic conservatism across the pantropical

- Caesalpinia Group (Leguminosae). *New Phytologist*, 222, 1994-2008. <https://doi.org/10.1111/nph.15633>
- Ghazanfar, S.A. (2003-2017) *Flora of Oman*. 4 volumes. National Botanic Garden of Belgium, Meise.
- Hammond, W.M., Williams, A.P., Abatzoglou, J.T., Adams, H.D., Klein, T., López, R., Sáenz-Romero, C., Hartmann, H., Breshears, D.D., & Allen, C.D. (2022) Global field observations of tree die-off reveal hotter-drought fingerprint for Earth's forests. *Nature Communications*, 13, 1-11. <https://doi.org/10.1038/s41467-022-29289-2>
- IPCC (2018) *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*. Cambridge University Press, Cambridge, UK and New York, NY, USA, 616 pp., <https://doi.org/10.1017/9781009157940>
- Karger, D.N., Conrad, O., Böhner, J., et al. (2017) Climatologies at high resolution for the Earth land surface areas. *Scientific Data*, 4, 170122. <https://doi.org/10.1038/sdata.2017.122>
- Köppen, W. (1900) Versuch einer Klassifikation der Klimate, vorzugsweise nach ihren Beziehungen zur Pflanzenwelt. *Geographische Zeitschrift*, 6, 593-611.
- Kumar, D., Pfeiffer, M., Gaillard, C., et al. (2020) Misinterpretation of Asian savannas as degraded forest can mislead management and conservation policy under climate change. *Biological Conservation*, 241, 108293. <https://doi.org/10.1016/j.biocon.2019.108293>
- Lancaster, L.T. & Humphreys, A.M. (2020) Global variation in the thermal tolerances of plants. *Proceedings of the National Academy of Sciences USA*, 117, 13580-13587. <https://doi.org/10.1073/pnas.1918162117>
- Mackey, B.G., Berry, S.L. & Brown, T. (2007) Reconciling approaches to biogeographical regionalization: a systematic and generic framework examined with a case study of the Australian continent. *Journal of Biogeography*, 35, 213-229. <https://doi.org/10.1111/j.1365-2699.2007.01822.x>
- Maestre, F.T., Benito, B.M., Berdugo, M., et al. (2021) Biogeography of global drylands. *New Phytologist*, 231, 540-558. <https://doi.org/10.1111/nph.17395>
- Médail, F. & Quézel, P. (2018) *Biogéographie de la flore du Sahara. Une biodiversité en situation extrême*. IRD Editions, Conservatoire et Jardin botaniques de la Ville de Genève, Marseille.
- Meyer, C., Kreft, H., Guralnick, R., et al. (2015) Global priorities for an effective information basis of biodiversity distributions. *Nature Communications*, 6, 8221. <https://doi.org/10.1038/ncomms9221>
- Miller, A.G. & Morris, M. (2004) *Ethnoflora of the Soqotra Archipelago*. Royal Botanic Garden, Edinburgh.
- Mirzabaev, A., Wu, J., Evans, J., et al. (2019) Desertification. In: *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems* (ed. by P.R. Shukla, J. Skea, E. Calvo Buendia, et al.), pp. 249-343. In-Press. Available at <https://www.ipcc.ch/srcl/>
- Montalvo-Mancheno, C.S., Ondeï, S., Brook, B.W., et al. (2020) Bioregionalization approaches for conservation: methods, biases, and their implications for Australian biodiversity. *Biodiversity and Conservation*, 29, 1-17. <https://doi.org/10.1007/s10531-019-01913-6>
- Olson, D., M., Dinerstein, E., Wikramanayake, E.D., et al. (2001) Terrestrial ecoregions of the world: a new map of life on earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, 51, 933-938. [https://doi.org/10.1641/0006-3568\(2001\)051\[0933:TEOTWA\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2)
- Schatz, G.E. (2001) *Generic tree flora of Madagascar*. Royal Botanic Gardens, Kew and Missouri Botanical Garden, USA.
- Sentinella, A.T., Warton, D.I., Sherwin, W.B., et al. (2020) Tropical plants do not have narrower temperature tolerances, but are more at risk from warming because they are close to their upper thermal limits. *Global Ecology and Biogeography*, 8, 1387-1398. <https://doi.org/10.1111/geb.13117>

- Shreve, F. & Wiggins, I.L. (1951) Vegetation and flora of the Sonoran Desert. Carnegie Institute of Science, Washington, D.C.
- Smith, P., Nkem, J., Calvin, K., et al. (2019) Interlinkages between desertification, land degradation, food security and greenhouse gas fluxes: synergies, trade-offs and integrated response options. In: Climate change and land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (ed. by P.R. Shukla, J. Skea, E. Calvo Buendia, et al.), pp. 551-652. In press. Available at <https://www.ipcc.ch/srccl/>
- Trabucco, A. & Zomer, R.J. (2018) Global aridity index and potential evapo-transpiration (ET0) Climate Database v2. CGIAR Consortium for Spatial Information (CGIAR-CSI). Available at <https://cgiarcsi.community>
- UNEP (1997) World atlas of desertification, 2nd edn. Arnold, London.
- Vásquez-Cruz, M. & Sosa, V. (2019) Assembly and origin of the flora of the Chihuahuan Desert: the case of sclerophyllous Rosaceae. *Journal of Biogeography*, 47, 445-459. <https://doi.org/10.1111/jbi.13745>
- Veldman, J.W. (2016) Clarifying the confusion: old-growth savannahs and tropical ecosystem degradation. *Philosophical Transactions of the Royal Society B*, 371, 20150306. <https://doi.org/10.1098/rstb.2015.0306>
- Vilhena, D.A. & Antonelli A. (2015) A network approach for identifying and delimiting biogeographical regions. *Nature Communications*, 6, 6848. <https://doi.org/10.1038/ncomms7848>
- Whittaker, R.H. (1970) *Communities and Ecosystems*. Macmillan, New York.
- Wu, S.-D., Zhang, L.-J., Lin, L., et al. (2018) Insights into the historical assembly of global dryland floras: the diversification of Zygophyllaceae. *BMC Evolutionary Biology*, 18, 166. <https://doi.org/10.1186/s12862-018-1277-z>
- Yusefi, G.H., Safi, K., Brito, J.C. (2019) Network- and distance-based methods in bioregionalization processes at regional scale: an application to the terrestrial mammals of Iran. *Journal of Biogeography*, 46, 2433-2443. <https://doi.org/10.1111/jbi.13694>
- Zizka, A., Silvestro, D., Andermann, T., et al. (2019) CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, 10, 744-751. <https://doi.org/10.1111/2041-210X.13152>
- Zizka, A., Carvalho-Sobrinho, J.G., Pennington, R.T., et al. (2020) Transitions between biomes are common and directional in Bombacoideae (Malvaceae). *Journal of Biogeography*, 47, 1310-1321. <https://doi.org/10.1111/jbi.13815>

Submitted: 11 February 2022

First decision: 6 April 2022

Accepted: 27 June 2022

Edited by Janet Franklin