ELSEVIER

A synthesis of botanical informatics for vascular plants in Africa

Hong Qian [a,*], Yadong Zhou [b,c], Jian Zhang [d], Yi Jin [e], Tao Deng [f], Shunsheng Cheng [g]

[a] *Research and Collections Center, Illinois State Museum, 1011 East Ash Street, Springfield, IL 62703, USA*
[b] *Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China*
[c] *Sino-Africa Joint Research Center, Chinese Academy of Sciences, Wuhan 430074, China*
[d] *Tiantong National Forest Ecosystem Observation and Research Station & School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241, China*
[e] *State Key Laboratory of Plant Physiology and Development in Guizhou Province, School of Life Sciences, Guizhou Normal University, Guiyang 550001, China*
[f] *CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China*
[g] *State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China*

ARTICLE INFO

ABSTRACT

We synthesize published botanical information for the flora of Africa to address multiple questions, including the following: How many species of vascular plants are there in Africa? How are the species distributed among different phylogenetic clades or taxonomic groups of vascular plants and different geographic regions across Africa? How complete are species lists derived from open-access online sources at various spatial scales? And at what spatial scales might species lists derived from open-access online sources be used to address macro-ecological and biogeographic questions for the flora of Africa? We used eight open-access online botanical databases and 188 regional, national, and local floras and checklists of vascular plants to generate species lists for sampling areas at various spatial scales examined in this study. Africa harbours at least 65,414 native species of vascular plants. Using the online botanical databases examined in this study to generate country-level or larger species lists, the completeness of the resulting species lists will be typically >90%. The completeness of species lists derived from geo-referenced occurrence records available online was 36.6, 53.1 and 72.1%, respectively, at the spatial scales of 100 km × 100 km, 200 km × 200 km, and 300 km × 300 km.

1. Introduction

Africa covers 30.4 million km^2 of land, which is about 20% of the land area of the Earth and almost symmetrically astride the equator. Africa is rich in both species diversity and ecosystem diversity (Linder, 2014; White, 1983). Six of the 25 mega-biodiversity hotspots of the world are located in Africa (Myers et al., 2000). Of the 235 centers of plant diversity recognized worldwide, 84 (36%) are located in Africa (Davis et al., 1994). There are at least 25 families of vascular plants that are endemic to Africa (Klopper et al., 2007). Africa possesses three main biomes: subtropical desert, tropical savanna, and tropical forest (White, 1983). At the northern end of Africa lies the Sahara, which is the largest desert and arguably the largest extremely species-poor area in the world (Linder, 2014). At the southern end of the continent lies the Cape flora, which is arguably the most species-rich temperate flora on the planet (Linder, 2014). Between these two ends lie the extensive species-poor Sahelian semi-desert of West Africa, the world's richest flora of

succulent plants (the Succulent Karoo) in South Africa and Namibia, and the species-rich rainforests of West and Central Africa (White, 1983). Furthermore, one of the world's six floristic kingdoms, the Cape Floristic Kingdom, is located in the southernmost part of Africa (Good, 1974; Takhtajan, 1986).

The flora of Africa has attracted attention of many botanists for more than 400 years. Land plants in South Africa were first collected no later than 1600 (Victor et al., 2016). The Cape of South Africa was one of the first areas outside Europe to be explored botanically (Goldblatt, 1978). Based on plant specimens collected from the Cape, Carl Linnaeus published the first flora for the Cape, *Flora Capensis* (Linnaeus, 1759), which consists of 502 plant names. During the past few decades, much effort from a great number of botanists has been devoted to compiling species checklists and floras of vascular plants for each of the countries in Africa and its broad geographical regions (e.g. southern Africa). Country-level floras or checklists of vascular plants have been compiled for nearly all the countries in Africa (see Table S1 in Supporting Information), either

---

for individual countries or for a group of countries within a broad region (e.g. plant checklists or floras for Tanzania, Uganda, and Kenya were included in the *Flora of Tropical East Africa*; Flora of Tropical East Africa Editors, 1952–2012; Zhou et al., 2017). Furthermore, for most of the countries in Africa, regional and local species checklists of vascular plants within a country have been compiled and published (see Table S1 for examples). Thus, country-level checklists of vascular plants have been well documented in Africa, from which several mega-regional checklists have been produced (e.g. southern Africa, Germishuizen, 2006; Sub-Saharan Africa, Klopper et al., 2006, 2007). However, the botanical information published in individual national and regional floras and checklists has not been synthesized for all of Africa. As a result, unlike many other continents or continent-like regions whose species richness of vascular plants at a continental level has been well documented based on continent-level species checklists (e.g. Europe, Tutin et al., 1964-1980; North America and South America, Ulloa et al., 2017; Australia, Australian Plant Census at www.anbg.gov.au/citation.html), it remains unknown how many described species of vascular plants occur in Africa, which is a basic and important question to biogeography, macroecology, and conservation biology. Previous studies have estimated the species richness of vascular plants in Africa, but estimates vary greatly between different studies. For example, Davis et al. (1994) estimated that Africa has 40,000–50,000 species of vascular plants, possibly as many as 60,000 species, while Govaerts (2001) estimated that for seed plants alone Africa harbours 74,232 species. The first goal of the present study is to answer the questions of how many native species of vascular plants have been documented in the literature and reliable plant databases for Africa and how the species are distributed among major clades (e.g. families, orders and clades above the ordinal level) across the phylogeny of vascular plants. We address these questions based on nearly all, if not all, available inventory-based floristic data pertinent to the flora of Africa. Answers to these questions will fill an important knowledge gap.

Species richness of plants varies greatly across Africa, from the species-poor Sahara Desert of northern Africa through the species-rich tropical rainforests of central Africa to the species-rich temperate flora of southern Africa, as noted above. Disentangling ecological and evolutionary causes of differences in species richness and floristic relationships between regions requires complete or nearly complete species lists for each geographical region. Previous studies have frequently used species lists of country-level geographical units (e.g. Sandel et al., 2020; van Kleunen et al., 2015; Zhang et al., 2018) or state- or province-level geographical units of large countries (e.g. China, Qian et al., 2019; Canada and the USA, Qian, 2009) to address biogeographic and macroecological questions. Much effort has been devoted to making country-level distributional records and geo-referenced occurrence points of vascular plants in Africa available online [e.g. African Plant Database (APD), www.ville-ge.ch/musinfo/bd/cjb/africa/recherche.php; Plants of the World online (POWO), www.plantsoftheworldonline.org/; Global Biodiversity Information Facility (GBIF), www.gbif.org; GlobalTreeSearch, tools.bgci.org/global_tree_search.php; tropical African vascular plants (RAINBIO), https://gdauby.github.io/rainbio/index.html; West African Plants initiative, Asase et al., 2020]. If complete or nearly complete country-level species lists of vascular plants can be derived from all available online botanical sources, many biogeographic and ecological questions may be addressed based on species lists derived from the online sources. The second goal of the present study is to assess the completeness of country-level species lists of vascular plants derived from online sources, by comparing them with country-level species lists published in the literature. It is possible that species lists published in the literature may not be complete to some degree. Thus, the completeness of a species list derived from open-access online data sources for an area is relative to the published flora of the area. The result from this analysis can guide future use of country-level species lists derived from open-access online data in biogeographic and ecological studies.

Addressing many ecological questions requires species distributional

data at spatial resolutions much finer than most countries. Many studies have used species lists at spatial scales ranging from about 100 km × 100 km grid cells (e.g. Fritz and Rahbek, 2012; McKnight et al., 2007) to about 300 km × 300 km grid cells (e.g. Linder, 2001). Species lists for grid cells are commonly generated using geo-referenced occurrence points (Asase and Peterson, 2016; Koffi et al., 2015; Stropp et al., 2016). Downloadable geo-referenced occurrence points for vascular plants of Africa are available from GBIF and RAINBIO; thus data with these sources may be used to generate species lists for grid cells across Africa at the 10,000-km^2 scale or larger. GBIF is the largest global data portal of species occurrence records. However, previous studies have shown that species lists derived from GBIF may be substantially incomplete in some regions and at some spatial scales. For example, Qian et al. (2018) showed that the completeness of species lists of vascular plants derived from GBIF is 37.2% at the province scale (~343,000 km^2) and 12.7% at the county scale (~8200 km^2) for China. Stropp et al. (2016) assessed the completeness of species lists of flowering plants derived from GBIF for 25 km × 25 km grid cells for Africa, and they demonstrated that only 0.6% of the grid cells of Africa have a completeness of ≥50%. The RAINBIO database was built to document geo-referenced occurrences for vascular plants in tropical Africa. The third goal of the present study is to assess the completeness of species lists of vascular plants derived from both GBIF and RAINBIO at three spatial scales (100 km × 100 km, 200 km × 200 km, 300 km × 300 km); data at these spatial scales are commonly used in biogeographic and macroecological studies. The aim of this analysis is to provide a guideline for the use of geo-referenced occurrence data with GBIF and RAINBIO in future studies addressing biogeographic and ecological questions at spatial scales of <100,000 km^2 in Africa.

Overall, the goals of this study are as follows: (1) to determine how many native species of vascular plants have been documented in the literature and reliable plant databases for Africa and how the species are distributed among major clades across the phylogeny of vascular plants, (2) to assess the completeness of country-level species lists of vascular plants derived from online sources, and (3) to assess the completeness of species lists of vascular plants derived from both GBIF and RAINBIO at three spatial scales (100 km × 100 km, 200 km × 200 km, 300 km × 300 km).

2. Materials and methods

2.1. Data collection for regional floras

Africa in this study is defined to include continental Africa, Madagascar and main surrounding islands (i.e. Canary Islands, Madeira, Cape Verde, and São Tomé and Príncipe, which are commonly considered as part of Africa). Plant distribution data for Africa are commonly documented for political countries. Because species richness increases with increasing area (i.e. species-area relationship) and because difference in area among the countries in Africa is very large (by about 210 times; e.g. 10,689 km^2 for Gambia versus 2,381,741 km^2 for Algeria), it makes little sense to compare species richness among geographic units with such a big variation in area. To minimize variation in area and increase comparability in species richness among geographic units across Africa, we grouped two or more smaller countries into a single geographic unit. Specifically, we divided Africa into 27 regions, which we termed "botanical regions", based on country boundaries (Fig. S1). Fourteen botanical regions each included only one large administrative (political) country while the remaining botanical regions each included two or more smaller administrative countries (Fig. S1). We documented species composition of vascular plants for each of the 27 botanical regions and for each of the islands primarily based on floristic information in the following eight online data sources: African Plant Database (APD, http://www.ville-ge.ch/musinfo/bd/cjb/africa/recherche.php), Plants of the World Online database (POWO, http://www.plantsoftheworldonline.org/), tropical African vascular plant database (RAINBIO, http

s://gdauby.github.io/rainbio/index.html), Botanical Database of Southern Africa (BODATSA, http://posa.sanbi.org/), Global Tree Search database (GlobalTreeSearch, https://tools.bgci.org/global_tree_search. php), Catalogue of the Plants of Madagascar (http://www.tropicos.org /Project/Madagascar), the plant database of Missouri Botanical Garden (TROPICOS, https://www.tropicos.org/home), and e-floras for Africa with the World Flora Online (WFO, www.worldfloraonline.org). In addition, we used 188 regional, national and local floras and checklists of vascular plants in the literature or online sources (Table S1). These data sources collectively allow generating complete or nearly complete lists of the vascular plant species that have been known to science for Africa as a whole and for each of the botanical regions and islands examined in this study. Botanical nomenclature of species was standardized according to WPO, which is a successor of The Plant List (www. theplantlist.org), a global plant database commonly used in standardizing botanical nomenclature (e.g. Cayuela et al., 2012; Zhou et al., 2021). We generated a species list including all plant names in the data sources used in this study, used WFO database to determine which names are accepted names and which names are synonyms in the African species list, and used the nomenclature information to standardize nomenclature for plant names in each of the individual data sources. Infraspecific taxa were combined with their respective species. Non-native distributions were excluded, based on the information about the nativity status of each species in each country or larger region available in APD, POWO, RAINBIO, GlobalTreeSearch, BODATSA, the Global Naturalized Alien Flora database (van Kleunen et al., 2019), World Economic Plants (Wiersema and León, 2013), and numerous regional and national floras and checklists published in the literature (Table S1). For a region larger than a country or Africa, if a species is native to any country within the region or Africa, it was considered to be native to the region or Africa.

We grouped the genera of seed plants into families and orders based on the Angiosperm Phylogeny Website (http://www.mobot.org/M OBOT/research/APweb), and grouped the genera of pteridophytes into families and orders based on the Pteridophyte Phylogeny Group (2016), respectively. Delineations of families and orders for angiosperms are consistent with those of APG IV (Angiosperm Phylogeny Group, 2016). We displayed the numbers of families, genera and species in each order based on an order-level phylogenetic tree, which was extracted from the mega-tree reported in Jin and Qian (2019).

We investigated the completeness of species lists derived from the above-mentioned online sources at various spatial scales. Samples at the largest spatial scale examined included the 27 botanical regions (Fig. S1). At this scale, we used botanical data extracted from all the above mentioned online sources to generate regional species lists and compared these species lists with those derived from the online sources plus species lists for regional, national and local floras and checklists published in the literature (Table S1). This comparison allows us to determine whether the completeness of species lists derived from online botanical sources at this spatial scale is sufficiently high for macro-ecological and biogeographical studies.

2.2. Data collection for grid cells (10,000–100,000 km^2)

Most online botanical sources mentioned above provide plant distribution information at much larger spatial scales (e.g. country-level distributions with POWO and GlobalTreeSearch) than grid cells commonly used in ecological and biogeographic studies (10,000–100,000 km^2, e.g. McKnight et al., 2007, Linder, 2001). Thus, they can not be used to generate species lists for grid cells at the three spatial scales that we examined in this study (i.e. 100 km × 100 km, 200 km × 200 km, and 300 km × 300 km). RAINBIO and GBIF are two major online sources with geo-referenced plant distribution data that can be downloaded and used to generate species lists at the three spatial scales for Africa. Accordingly, we used data obtained from these two online sources to generate species lists of vascular plants at the three spatial

scales. We used World Flora Online to standardize botanical nomenclature, and we combined infraspecific taxa with their respective species for the data obtained from RAINBIO and GBIF. RAINBIO separated native distributions from introduced distributions, with 614,016 occurrence records for native distributions, each of which possesses values of latitude and longitude. GBIF included 7.64 million occurrence records for African vascular plants (accessed in March 2020). Because GBIF does not distinguish between native and non-native distributions, we used the above-mentioned regional native species lists to determine native versus non-native occurrences for each botanical region. We excluded those occurrence records which either do not have data for latitude and longitude or have data for latitude and longitude but locations indicated by the latitude and longitude data are outside of the administrative countries indicated in GBIF. These data cleaning processes resulted in 2.49 million occurrence records that can be assigned to grid cells based on their latitudes and longitudes. We combined the RAINBIO and GBIF geo-referenced occurrence records and used them to generate species lists for each grid cell at the three spatial scales. We used Albers equal-area projection to divide Africa into 100 km × 100 km grid cells and then grouped them into grid cells of 200 km × 200 km and 300 km × 300 km.

2.3. Data analysis

One way to evaluate the completeness of species lists derived from RAINBIO and GBIF at the three spatial scales is to compare species richness of these species lists with that derived from inventory-based species lists at the same spatial scale under investigation. However, complete inventory-based species lists for Africa at the three spatial scales that we investigated (i.e. 100 km × 100 km, 200 km × 200 km, and 300 km × 300 km, or for areas of 10,000, 40,000 and 90,000 km^2) do not exist. We used an alternative approach, which is based on the species-area relationship (SAR), to assess the completeness of species lists derived from RAINBIO and GBIF at the three spatial scales. A commonly used SAR is the power-law model (Arrhenius, 1921): $S = cA^z$, where S represents species richness, A represents sampling area, and c and z are constants. Numerous studies (e.g. Ricklefs and Lovette, 1999; Rosenzweig, 1995) have shown that when both S and A are logarithm transformed, logS tends to be linearly correlated with logA. Accordingly, we built the log-log SAR model and used it to estimate S for a given A. Previous studies have shown that z-value can vary substantially between latitudes and regions (Qian et al., 2007). Therefore, we built different SAR models for different regions across Africa. Specifically, we identified six regions in Africa (Fig. S2) that have sufficient inventory-based species lists at different spatial scales to build robust SARs. We used 382 species lists to build six SARs for the six regions. These species lists were either species lists derived from individual literature sources presented in Table S1 or generated by combining regional or national species lists for those geographic areas that are larger than countries. Botanical nomenclature for species in these species lists were standardized according to World Flora Online, infraspecific taxa were combined with their respective species, and introduced species in each species list were excluded. For each of the six regions, the smallest sampling area was smaller than 10,000 km^2, and the largest sampling area was greater than 90,000 km^2. Thus, using the SAR models built in this study to estimate species richness in grid cells at the three spatial scales that we examined is free from the problem of extrapolation.

We conducted correlation and regression analyses to assess the relationships among variables. For correlation analyses, we used the Pearson correlation coefficient to assess the relationship between pairwise variables. Because spatial autocorrelation commonly occurs in broad-scale ecological data, which inflates test of statistical significance, we followed previous authors (e.g., Fritz and Rahbek, 2012; Hawkins et al., 2011; Qian et al., 2019) not to report *P*-values for correlation coefficients. Instead, we evaluated the strength of each correlation coefficient by effect size. Specifically, we considered a correlation to be

strong for $|r| > 0.66$, moderate for $0.66 \geq |r| > 0.33$, or weak for $|r| \leq 0.33$ (Qian et al., 2019). We used the packages SYSTAT (Wilkinson et al., 1992) to conduct statistical analyses.

3. Results

3.1. Species richness of native plants derived from species lists

There were 65,414 native species of vascular plants, belonging to 311 families and 5099 genera, in Africa, including continental Africa, Madagascar, Canary Islands, Madeira, Cape Verde, and São Tomé and Príncipe, but excluding Sinai Peninsula, which belongs to Asia. When different major taxonomic groups of vascular plants were considered separately, Africa as a whole had 1614, 120, and 63,680 species of pteridophytes, gymnosperms and angiosperms, respectively. The numbers of families, genera and species in each order and each of more basal clades (i.e. clades above the ordinal level) within each of these three major groups of vascular plants were presented in Fig. 1. The five largest families in Africa are Fabaceae (5669 species), Asteraceae (5617), Rubiaceae (3397), Orchidaceae (2953) and Poaceae (2714). These five families accounted for 31.1% of the species of vascular plants in Africa. Of the 311 families of vascular plants in Africa, the 50 largest
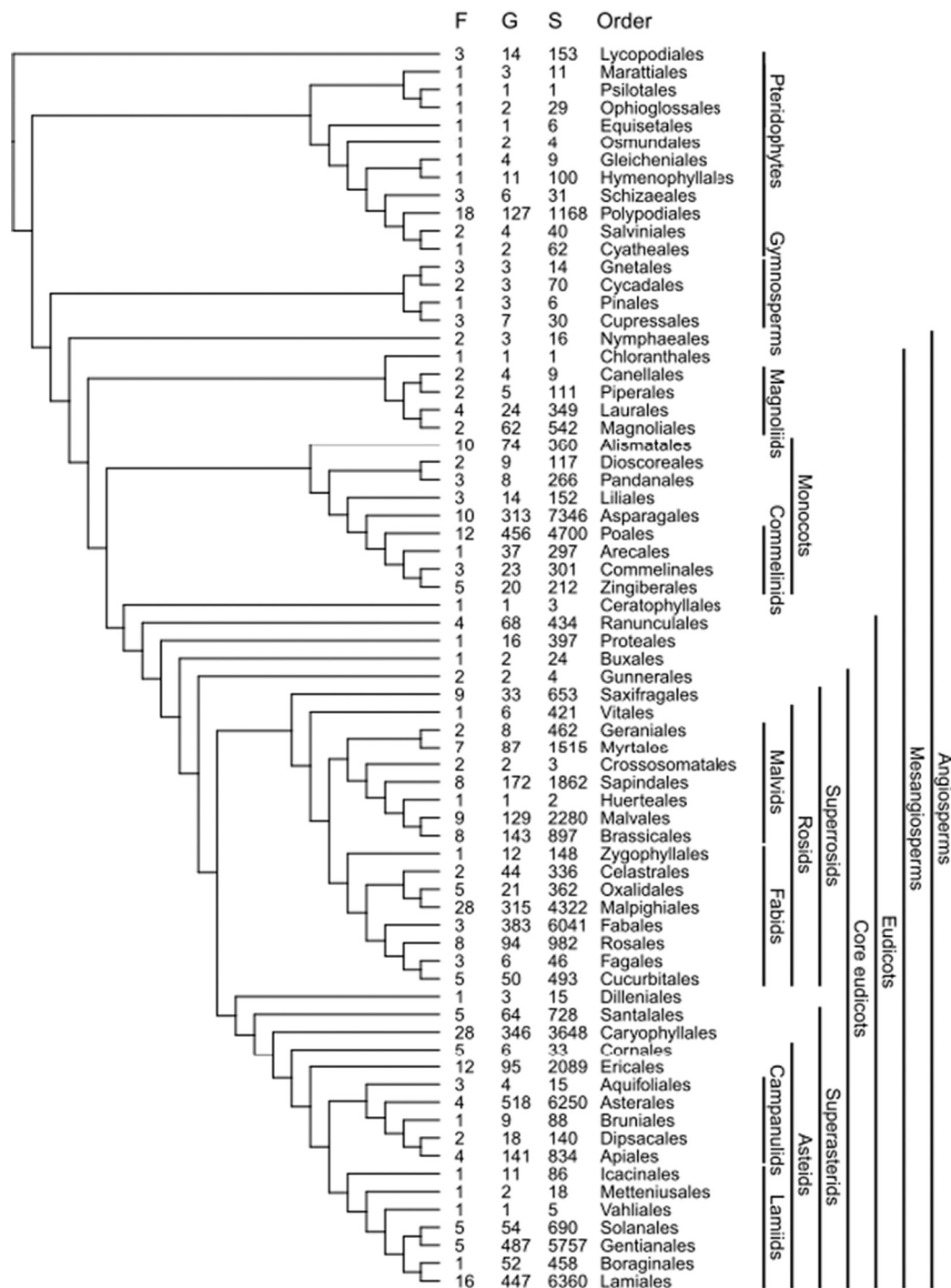

Fig. 1. Summary of the numbers of families (F), genera (G) and species (S) in each of the orders of vascular plants in Africa. The topology of the phylogenetic tree was extracted from the megaphylogeny reported in Jin and Qian (2019).

families accounted for 80.2% of the species of vascular plants in Africa (Fig. 2).

When continental Africa and Madagascar were considered together, they included 63,147 native species of vascular plants, belonging to 310 families and 4964 genera. When they were considered separately, continental Africa had 52,801 native species of vascular plants in 300 families and 4467 genera, and Madagascar had 12,351 native species of vascular plants in 236 families and 1714 genera. When continental Africa was divided into the three broad latitudinal zones (i.e. Southern Africa, Tropical Africa, and Northern Africa; Fig. S1), as frequently adopted in previous studies (e.g. Klopper et al., 2007), there were 22,026, 31,323 and 7052 native species of vascular plants in the three latitudinal zones, respectively. Sub-Saharan Africa (i.e. the combination of Southern Africa and Tropical Africa) harbours 47,714 native species of vascular plants. The number of native vascular plant species shared by all possible pairs of Southern Africa, Tropical Africa, Northern Africa, and Madagascar ranged from 400 to 5635 (Table 1).

When species richness of vascular plants was compared among the 27 botanical regions across Africa (Fig. S1), the botanical region including South Africa, Lesotho and Swaziland had the highest species richness (~20,000 native species; Fig. 3), which was followed by Madagascar (~12,000 native species). In contrast, Mauritania had the lowest species richness (~1100 native species) among all 27 botanical regions.

3.2. Completeness of botanical-region-level species lists derived from online botanical sources

The number of species of vascular plants in the species list of a botanical region derived from the eight online data sources (see Methods) was greater than that based on the complete native species list of the botanical region. This is because the species list derived from the online data sources included both native and introduced species. On average, 23.3% of the species list of a botanical region derived from the online data sources were non-native species to the botanical region (Fig. 3b). Of the 27 botanical regions (Fig. S1), Egypt had the highest percentage (~40%) of non-native species (Fig. 3b). The number of non-native species was strongly correlated with that of native species among the 27 botanical regions (Pearson correlation coefficient $r = 0.929$). Species richness was nearly perfectly correlated (Pearson's correlation coefficient $r = 0.999$ in all the three cases reported in Fig. 3) with species density, which was calculated by dividing species richness of a botanical region by logarithm-transformed geographic area of the botanical region, a method commonly used in the literature (e.g. Fridley et al., 2006; Qian, 1998; Vetaas and Grytnes, 2002). Species density of the botanical regions was presented in Fig. 4.

Table 1
The numbers of species of vascular plants shared among the four broad regions in Africa (as shown in Fig. S1).

	Southern Africa	Tropical Africa	Northern Africa
Tropical Africa	5635		
Northern Africa	779	1876	
Madagascar	1163	1954	400

When native species in botanical-region-level species lists derived from the online data sources were considered, the numbers of native species in these species lists were perfectly correlated with those of complete species lists ($r = 1.000$, $n = 27$). This result was consistent with the high completeness of native species lists derived from the online data sources (Fig. 3b). When the eight online data sources were used to generate botanical-region-level species lists of native vascular plants, the resulting species list of a botanical region would include, on average, 97.0% of native species in the botanical region, ranging from 93.2 to 99.8% (Fig. 5). However, when the two online data sources that had geo-referenced occurrence records (i.e. GBIF and RAINBIO) were used alone to generate botanical-region-level species lists, the average completeness of species lists for native plants was 79.0%, ranging from 45.6 to 93.9% (Fig. 5).

3.3. Completeness of species lists derived from online botanical sources for grid cells

Based on our grid system, 2786 (87%) of the 3199 grid cells in Africa at the spatial scale of 100 km × 100 km contained at least one plant occurrence record. For the GBIF data set, the number of occurrence records was strongly correlated with that of species ($r = 0.687$). Thus, the variation in occurrence records per cell across Africa was highly congruent with that for species (Fig. S3). On average, each 100 km × 100 km grid cell had 881.9 ($\pm$ 2967.6 SD) occurrence records and 274.2 ($\pm$ 480.4) species of native vascular plants. The number of species occurrences in the 2786 grid cells was 763,929. When data from GBIF and RAINBIO were considered together, the number of species occurrences in the 2786 grid cells increased to 803,730, and the average of species in each 100 km × 100 km grid cell was 288.5 ($\pm$ 493.7).

The species-area relationships for the six selected regions across Africa (Fig. S2) were all strong (Table 2), which explained on average 86.2% of the variation in species richness among sampling areas in different sizes. The slope (z-value) of a species-area relationship varied from 0.30 to 0.42 among the six regions (Table 2). When the average species richness derived from occurrence records in GBIF and RAINBIO at the spatial scale of 100 km × 100 km was compared with that estimated by the species-area relationship at the same spatial scale for a
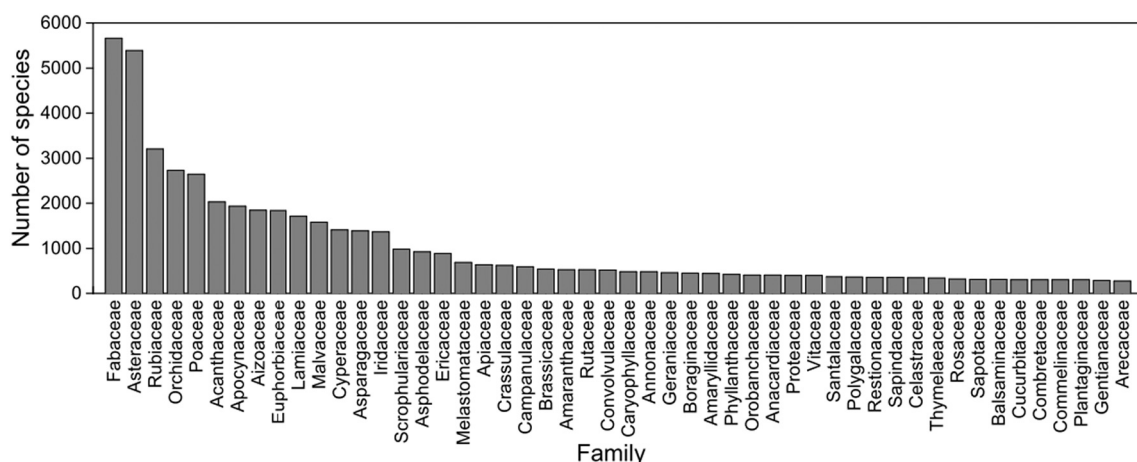
Fig. 2. Species richness of the 50 largest families of vascular plants in Africa.
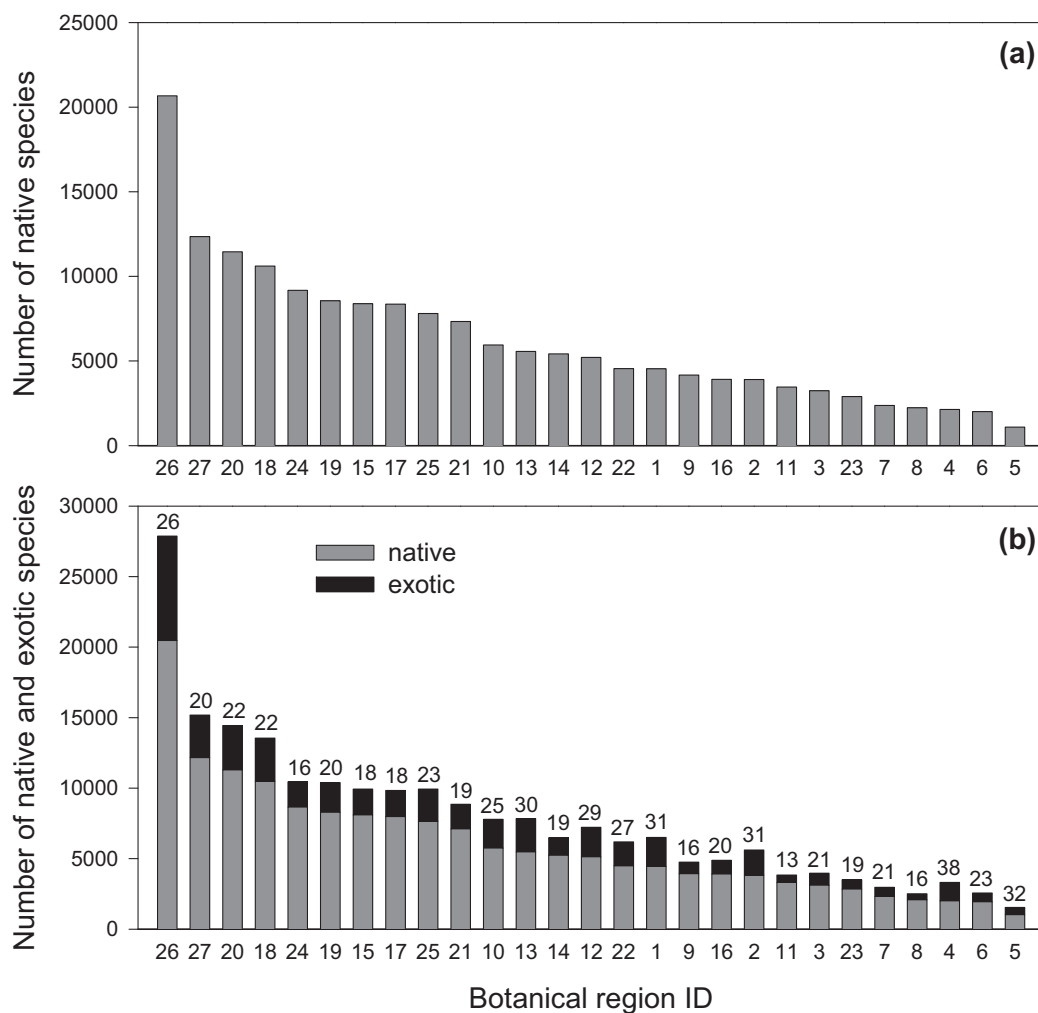
Fig. 3. Species richness of vascular plants for each of the 27 botanical regions in Africa, which were ordered based on species richness of native plants from richest to poorest. (a) The number of native plant species in each species list derived from all literature and online sources, (b) the number of native and exotic plant species and percentage of exotic plant species (above each bar) in each species list derived from the four online sources (GBIF, APD, GTS and RAINBIO). See Fig. 4 for the locations of the botanical regions, and species density (i.e. species richness divided by $\log_{10}$-transformed area in square kilometers) for the botanical regions.
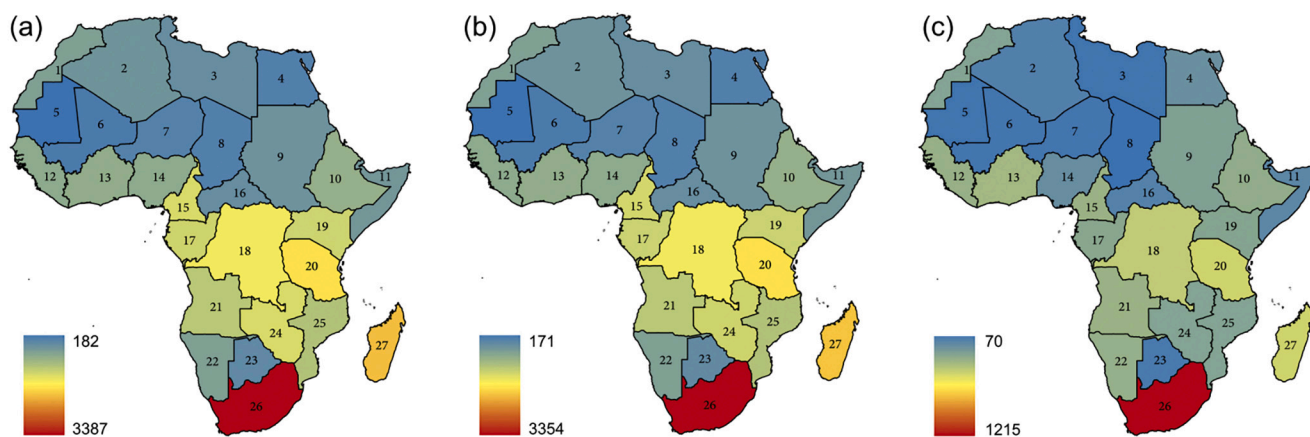

Fig. 4. Species density (i.e. species richness divided by $\log_{10}$-transformed area in square kilometers) of vascular plants for each of the 27 botanical regions in Africa. (a) Species density for native plant species in each species list derived from all literature and online sources, (b) species density for native plant species in each species list derived from the four online sources (GBIF, APD, GTS and RAINBIO), (c) species density for exotic plant species in each species list derived from the four online sources (GBIF, APD, GTS and RAINBIO). Species richness used to calculate species density in (a), (b), and (c) for each botanical region was shown in Fig. 3.

given region, species richness derived from GBIF and RAINBIO was, on average, only 36.6% of that estimated by the species-area relationship (Fig. 6). When the spatial scale increased to 200 km × 200 km and 300 km × 300 km, species richness derived from GBIF and RAINBIO was, on average, 53.1% and 72.1%, respectively, of those estimated by species-area relationships (Fig. 6).
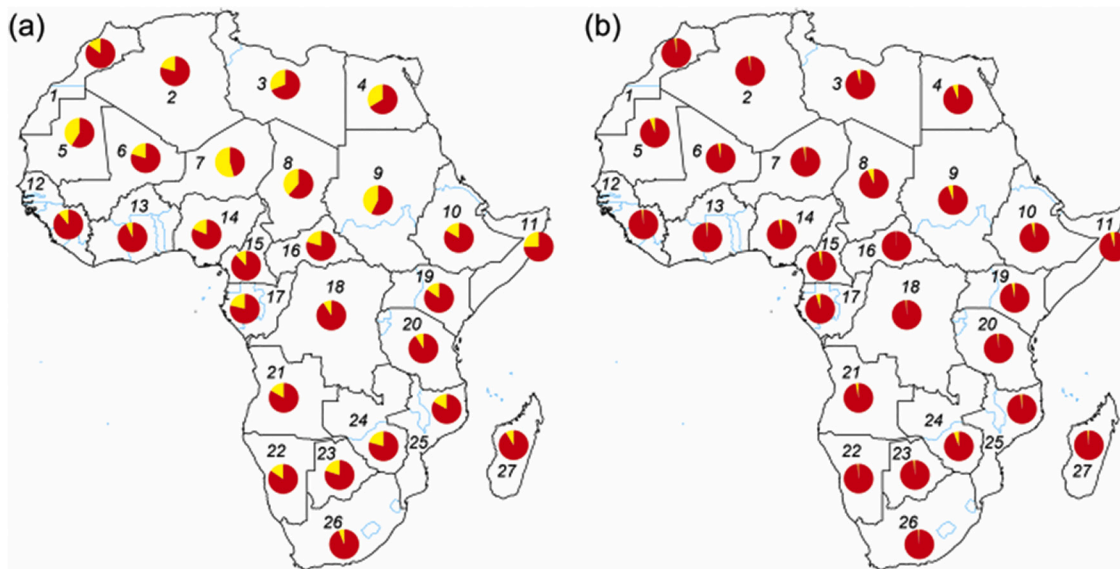
Fig. 5. Completeness (red) of botanical-region-level species lists of vascular plants (a) derived from geo-referenced occurrence records in GBIF and RAINBIO and (b) derived from all the eight online botanical databases used in this study (see Methods). See Fig. 3a for species richness of native vascular plants in each of the 27 botanical regions in Africa. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Information about the samples used to build the species-area relationships for six regions (as shown in Fig. S2) and the results of simple linear models with $\log_{10}$-transformed species richness being regressed on $\log_{10}$-transformed area (km^2).

| Region	No. of samples	Smallest sample ($\log_{10} km^2$)	Largest sample ($\log_{10} km^2$)	Model Constant (c-value)	Slope (z-value)	R^2
A	19	3.190	6.047	1.944	0.338	0.922
B	146	2.615	6.087	1.578	0.343	0.744
C	48	3.454	5.915	1.376	0.418	0.818
D	13	1.544	6.000	2.154	0.298	0.904
E	112	1.519	6.103	1.935	0.365	0.816
F	44	3.440	6.078	1.530	0.371	0.970

Note for region: Region A = Cameroon + Gabon + Equatorial Guinea + Republic of the Congo; Region B = Ethiopia + Eritrea + Djibouti; Region C = Kenya + Uganda; Region D = Tanzania + Rwanda + Burundi; Region E = South Africa + Lesotho + Swaziland; Region F = Benin + Ghana + Guinea + Guinea-Bissau + Côte d'Ivoire + Liberia + Sierra Leone + Togo.

4. Discussion

4.1. The flora of Africa

Our study is the first attempt to synthesize available inventory-based botanical information about the flora of Africa. Based on our synthesis, continental Africa and the islands considered in this study collectively harbour 65,414 native species of vascular plants. Our study area does not include some of those small islands in the Atlantic Ocean and the West Indian Ocean that have been considered as part of Africa. We conducted a supplementary analysis to assess how many additional species of vascular plants in those islands have not been included in the botanical data used in our analyses. Specifically, we downloaded distributional data from POWO for 15 islands or archipelagos surrounding Africa but outside our study area (i.e. Aldabra, Ascension, Azores, Cabinda, Chagos Archipelago, Comoros, Gulf of Guinea Is., Mauritius, Mozambique Channel Is., Réunion, Rodrigues, Selvagens, Seychelles, Socotra, and St. Helena); we found that these islands collectively have ~4800 species of native vascular plants in POWO but 98% of them are also present in the islands or mainland of our main
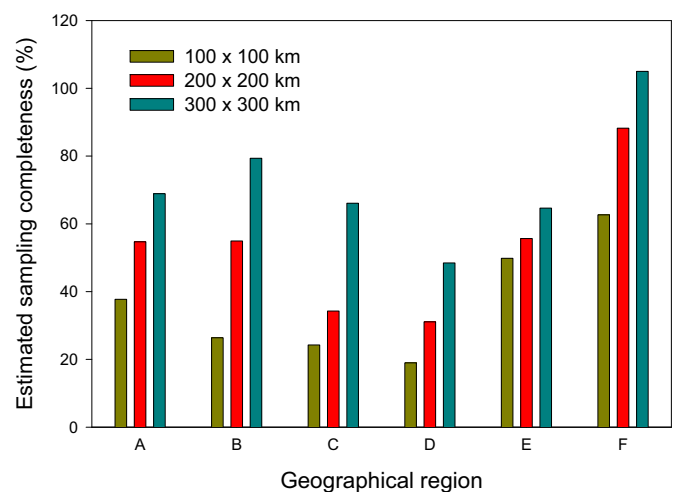
Fig. 6. Completeness (%) of species lists of native plants derived from two online sources (GBIF and RAINBIO) at three spatial scales in six geographical regions across Africa (as shown in Fig. S2).

study. Thus, including these small islands in a study on the African flora will only slightly increase its species richness.

The data analyzed in this study included 4021 species of vascular plants that were described as new species based on specimens collected from Africa during the past 30 years (from 1990 to 2019), i.e. 134 species per year, according to the TROPICOS database of the Missouri Botanical Garden. There appears to be a declining trend in the number of new species found in Africa (Fig. S4). A forecast using simple linear regression based on the number of new species described between 1990 and 2019 would yield about 68,500 native species of vascular plants in Africa by 2050.

Africa is 1.7 times as large as South America (30.4 versus 17.8 million km^2) but vascular plant species richness of Africa is only 80% of that of South America (65,414 versus 82,052 species, this study and Ulloa et al., 2017, respectively). When tropical regions of these two continents were compared, the anomaly in species richness between the two continents was enhanced: tropical Africa has 20.0 million km^2 and 31,323 species of vascular plants while tropical South America (i.e. the

South America north of Southern Cone; Ulloa et al., 2017) has 13.7 million km^2 and 79,134 species of vascular plants (Ulloa et al., 2017; http://www.tropicos.org/Project/VPA). Tropical Southeast Asia, which has only about 5 million km^2 of land and harbours more than 50,000 species of vascular plants (Couvreur, 2015; Whitmore, 1998), is also much richer than tropical Africa with respect to plant species. The anomaly in plant species richness between tropical Africa and other tropical regions appears to hold at local (forest stand) scales. For example, Parmentier et al. (2007) showed that species diversity of trees with diameter at breast height (dbh) ≥ 10 cm is higher for tropical rain forests in the Amazon basin than those in Central Africa. Africa has been labeled as the "odd man out" (Richards, 1973), due to its lower plant species diversity in tropical rain forests, compared with its counterparts in the Neotropics and Southeast Asia. Several hypotheses have been proposed to explain this diversity anomaly (see Couvreur, 2015 for a review), which include the following: Compared with the Neotropics and Southeast Asia, Africa has a higher number of "dry" months (Richards, 1973), has a smaller species pool adapted to high precipitation and temperature (Parmentier et al., 2007), has a smaller total area of tropical rain forests (195 million ha in continental Africa, 652 million ha in the Neotropics, 302 million ha in Southeast Asia, Mayaux et al., 2013), had smaller areas with stable climates during the Pliocene or Pleistocene (Couvreur, 2015), had higher extinction rates (Raven and Axelrod, 1974) and lower speciation rates (Baker and Couvreur, 2013), was less affected by tectonic movement (Couvreur, 2015), and had greater human impacts and disturbances on species diversity (Richards, 1973). Testing these hypotheses can shed light on understanding the origin and maintenance of the species richness of the African flora and the anomaly in species richness between Africa and South America, which were connected to each other 120 million years ago (Manish and Pandit, 2018).

4.2. Country-level floras

Floras and checklists of vascular plants at the country level in Africa have been well documented in the literature (Table S1). For some countries, their floras have been updated multiple times (e.g. *Flora of Egypt*; Boulos, 1995, 1999, 2009). To minimize difference in sampling area, we combined small-sized administrative countries to generate larger botanical regions. This would increase the completeness of resulting species lists because species lists tend to be more complete when species lists at a smaller spatial scale are combined to form species lists at a larger spatial scale, as shown in previous studies (e.g. Qian et al., 2018) as well as the present study (Fig. 6). Species richness of native vascular plants for each of the countries in Africa has been reported in previous studies (e.g. Pysek et al., 2017). Because data on species richness reported in those studies for native plants were taken from the original literature (primarily individual floras and checklists) and because botanical nomenclature in different literature sources was not standardized, species richness reported in previous studies may not be comparable between countries. Furthermore, because some country-level floras were published decades ago, data reported in those studies may be outdated, which may lead to underestimating species richness for some countries. In contrast, our study used the newest versions of country-level floras and species checklists published in the literature (Table S1), used the most recent versions of the online sources (last accesses in 2020 in all cases), included all new species described in the recent past decades (up to 2019; Fig. S4), and standardized plant names in different data sources. Thus, species richness reported in this study is up-to-date and comparable among different countries. The average number of native species of vascular plants in the 52 counties in Africa reported in Pysek et al. (2017) is 3984 but the average number of native vascular plant species for these 52 countries is 4286 based on the data used in our study (Table S2), indicating that species lists used in our study are more complete.

GBIF includes 83.0% of all country-level occurrences for Africa.

Although the data extracted from GBIF may result in more complete species lists, compared with those derived from other online sources, it is a great challenge to use GBIF data to generate species lists partly because many occurrence records with GBIF are erroneous in terms of geographic locations and identifications (Stropp et al., 2016) and partly because native and non-native distributions are not differentiated. If non-native species are not filtered out when using GBIF to generate a species list of native plants for a country, thousands of non-native species may be erroneously included in the resulting species list as native species. As our study shows (Fig. 3), over 7000 species of vascular plants in GBIF for botanical region #26 (including South Africa, Lesotho and Swaziland) are non-native species to the botanical region. Few, if any, previous studies have taken the issue of non-native species into consideration when using data in GBIF. Including a large number of non-native species in a study that aims to address questions about native species would likely substantially bias the results of the study. The Global Naturalized Alien Flora database (GloNAF; van Kleunen et al., 2019) may be helpful to filter out naturalized non-native species at the country level, but a great number of non-native species in each country are not naturalized species and thus are not included in country-level naturalized alien species lists in GloNAF. If the nativity status (i.e. native versus non-native) of each species in each focal region can be determined, using APD, POWO and GBIF collectively, in conjunction with other online botanical databases, to generate country-level or larger species lists, the completeness of the resulting species lists will be typically >90% (Fig. 5).

4.3. Species lists derived from geo-referenced occurrence data for areas at the spatial scales ranging from 10,000 to 90,000 km^2

Of the online botanical data sources examined in the present study, GBIF and RAINBIO are the two which hold downloadable geo-referenced occurrence records that can be used to generate species lists for geographic areas at any size, including commonly used grid cells at spatial scales ranging from 10,000 to 90,000 km^2 (Linder, 2001; McKnight et al., 2007). GBIF holds much more geo-referenced occurrence records than does RAINBIO. At the continental scale, adding data from RAINBIO to GBIF increased species occurrences only by 5% at the spatial scale of 100 km × 100 km (i.e. 763,929 versus 803,730 occurrences). When 1533 grid cells of 100 km × 100 km located in latitudes between 15° N and 15° S, where tropical rain forests in Africa are located, which are the focus of the RAINBIO database, were considered, adding data from RAINBIO to GBIF increased species occurrences by 9% at the spatial scale of 100 km × 100 km (i.e. 392,043 versus 429,822 occurrences). Thus, using geo-referenced occurrence records with both GBIF and RAINBIO to generate species lists for Africa would increase the degree of the completeness of the resulting species lists, compared to those derived from either database alone.

Evaluating the completeness of a species list derived from geo-referenced occurrence records for an area requires estimating species richness of the area if an inventory-based complete species list for the area is not available. Commonly used approaches to estimate species richness include species accumulation curve (SAC; Tittensor et al., 2010) and species-area relationship (SAR; Rosenzweig, 1995). The SAC approach is based on the relationship between the number of occurrence records and the number of species in the occurrence records in a sampling area, and uses the final slope of the species accumulation curve (usually the last 10%) to estimate species richness and sampling completeness of the area (Stropp et al., 2016; Yang et al., 2013). A flat slope indicates saturation in the sampling and thus a high degree of inventory completeness. The SAC approach assumes that the probability of recording a species as occurring in a given sampling area correlates positively with the number of occurrence records collected for that area (Lobo et al., 2018), and that occurrence records are sampled randomly and that species occurrences are neither spatially nor temporally auto-correlated (Colwell and Coddington, 1994; Gotelli and Colwell, 2001).

However, these assumptions do not always hold (Lobo et al., 2018). The shape of an SAC is sensitive to the number of occurrence records used to build the SAC, as shown in Yang et al. (2013). SAC-based approaches may substantially bias estimates of completeness of species lists derived from occurrence records. For this reason, we did not use any SAC-based approach in the present study. Instead, we used a SAR-based approach to estimate species richness, which is a commonly used approach to estimate species richness (Rosenzweig, 1995).

Our SAR models were built based on a relatively large number of sampling areas for each of the six selected regions (Fig. S2). Because the models explained the vast majority (86% on average) of the variation in species richness in sampling areas and because the three spatial scales of grid cells examined in this study (i.e. 10,000, 40,000 and 90,000 km^2) are all within the range of minimum and maximum sizes of sampling areas for all the six regions, it is reasonable to believe that the models can accurately predict species richness at the three spatial scales. At a given spatial scale (e.g. 10,000 km^2) within a region, species richness may vary between grid cells within the region, even though the size of each region analyzed in our study is not very large. However, for each region, because we compared the estimated species richness with the average of species richness observed in all grid cells within the region, our estimate of the completeness of species lists derived from geo-referenced occurrence records should be reasonably accurate. We have reported the c- and z-values of each SAR model in Table 2. These models can be used to assess the completeness of species lists derived from GBIF and any other data sources at any spatial scales within the range of sampling sizes of each model (Table 2). New data are regularly added to the GBIF database. Our models can also be used to assess the completeness of species lists of vascular plants derived from different versions of the GBIF database not only at the three spatial scales that we examined for Africa but also any spatial scale between 10,000 and 90,000 km^2.

Stropp et al. (2016) assessed species lists of angiosperms (flowering plants) derived from GBIF at the spatial scale of 25 km × 25 km for Africa. The GBIF data used in their study were retrieved on October 10, 2012 (Stropp et al., 2016). After data cleaning, 934,676 occurrence records were left for their study. They found that less than 1% of the 25 km × 25 km grid cells in Africa have a completeness of ≥50%. The GBIF data used in our study were downloaded in March 2020, with 2.49 million occurrence records having values of latitude and longitude correctly located within their respective countries in Africa. Thus, the number of occurrence records derived from GBIF in our study is more than twice as large as that of their study. In addition, the GBIF data used in our study were supplemented by the data from RAINBIO. Thus, our study used a much larger data set to assess the completeness of species lists for all major groups of vascular plants (i.e. pteridophytes, gymnosperms, and angiosperms) derived from geo-referenced occurrence records. We found that the completeness of species lists derived from geo-referenced occurrence records was 36.6, 53.1 and 72.1%, respectively, at the spatial scales of 100 km × 100 km, 200 km × 200 km, 300 km × 300 km. The degree of completeness is much higher in our study than that of Stropp et al. (2016), partly because of more occurrence records and larger spatial scales used in our study.

Our study showed that at the spatial scale of 100 km × 100 km, the completeness of species lists for vascular plants derived from currently available geo-referenced occurrence data is only about one third of estimated species richness. With this low degree of completeness, the current GBIF data alone, or after being supplemented with other data sources such as RAINBIO, is not appropriate for use in addressing most, if not all, biogeographic and macroecological questions for the flora of Africa. Similarly, because species lists derived from GBIF and RAINBIO include only about half of estimated species richness at the spatial scale of 200 km × 200 km, it may also not be appropriate to use such derived species lists in biogeographic and macroecological studies. Our study showed that about three quarters of estimated species richness were included in species listed derived from GBIF and RAINBIO at the spatial

scale of 300 km × 300 km, some biogeographic and macroecological questions may be addressed at this spatial scale using data with GBIF and RAINBIO. Considering that the number of geo-referenced occurrence records with GBIF for Africa increased greatly from 2012 to 2020 (compare Stropp et al.'s study with our study), we predict that many new occurrence data will be included in GBIF for Africa in the next decade. The users of the GBIF database may use the SAR models reported in the present study as benchmarks to determine the completeness of species lists derived from GBIF from time to time.

We did not assemble a species-area relationship for any region in Northern Africa (i.e. Sahara) to assess the completeness of species lists derived from GBIF, because data to build a reliable species-area relationship is lacking across Northern Africa. Considering that nearly half of Northern Africa does not have any geo-referenced occurrence record in GBIF (Fig. S3), the degree of the completeness of species lists derived from GBIF is expected to be lower, compared with other parts of Africa. Thus, species lists of vascular plants for grid cells in Northern Africa derived from GBIF may not be used in biogeographic and macroecological studies requiring relatively complete species lists.

Funding

Data availability statement

All data used in this study have been published and are accessible to readers from the sources cited in this article.

Declaration of Competing Interest

There is no competing interest with this study.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2021.101382.

References

Angiosperm Phylogeny Group, 2016. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 181, 1–20.

Arrhenius, O., 1921. Species and area. J. Ecol. 9, 95–99.

Asase, A., Peterson, A.T., 2016. Completeness of digital accessible knowledge of the plants of Ghana. Biodivers. Inform. 11, 1–11.

Asase, A., Sainge, M.N., Radji, R.A., Ugbogu, O.A., Peterson, A.T., 2020. A new model for efficient, need-driven progress in generating primary biodiversity information resources. Appl. Plant Sci. 8, e11318.

Baker, W.J., Couvreur, T.L.P., 2013. Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. II. Diversification history and origin of regional assemblages. J. Biogeogr. 40, 286–289.

Boulos, L., 1995. Flora of Egypt: Checklist. Al Hadara Pub, Cairo, Egypt.

Boulos, L., 1999. Flora of Egypt. Al Hadara Pub, Cairo, Egypt.

Boulos, L., 2009. Flora of Egypt: Checklist. Revised Annotated Edition. Al-Hadara Publishing, Cairo.

Cayuela, L., Cerda, Í.G.-d.l., Albuquerque, F.S., Golicher, D.J., 2012. TAXONSTAND: an R package for species names standardisation in vegetation databases. Methods Ecol. Evol. 3, 1078–1083.

Colwell, R.K., Coddington, J.A., 1994. Estimating terrestrial biodiversity through extrapolation. Philos. Trans. R. Soc. B: Biol. Sci. 345, 101–118.

Couvreur, T.L.P., 2015. Odd man out: why are there fewer plant species in African rain forests? Plant Syst. Evol. 301, 1299–1313.

Davis, S.D., Heywood, V.H., Hamilton, A.C., 1994. Centres of Plant Diversity. A Guide and Strategy for their Conservation, Vol. 1. IUCN Publications Unit, Cambridge, U.K.

Flora of Tropical East Africa Editors, 1952–2012. Flora of Tropical East Africa. Royal Botanic Gardens, Kew, London.

Fridley, J.D., Qian, H., White, P.S., Palmer, M.W., 2006. Plant species invasions along the latitudinal gradient in the United States: comment. Ecology 87, 3209–3213.

Fritz, S.A., Rahbek, C., 2012. Global patterns of amphibian phylogenetic diversity. J. Biogeogr. 39, 1373–1382.

Germishuizen, G., 2006. A Checklist of South African Plants. SABONET/SANBI, Pretoria.

Goldblatt, P., 1978. An analysis of the flora of southern Africa: its characteristics, relationships, and origins. Ann. Mo. Bot. Gard. 65, 369–436.

Good, R., 1974. The Geography of the Flowering Plants, 3rd edn. White Plains, New York, Longman.

Gotelli, N.J., Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4, 379–391.

Govaerts, R., 2001. How many species of seed plants are there? Taxon 50, 1085–1090.

Hawkins, B.A., Rodríguez, M.Á., Weller, S.G., 2011. Global angiosperm family richness revisited: linking ecology and evolution to climate. J. Biogeogr. 38, 1253–1266.

Jin, Y., Qian, H., 2019. V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. Ecography 42, 1353–1359.

Klopper, R.R., Chatelain, C., Banninger, V., Habashi, C., Steyn, H.M., De Wet, C., Arnold, T.H., Gautier, L., Smith, G.F., Spichiger, R., 2006. Checklist of the Flowering Plants of Sub-Saharan Africa: An Index of Accepted Names and Synonyms. SABONET Report No. 42.

Klopper, R., Gautier, L., Chatelain, C., Smith, G., Spichiger, R., 2007. Floristics of the angiosperm flora of sub-Saharan Africa: an analysis of the African plant checklist and database. Taxon 56, 201–208.

Koffi, K.J., Kouassi, A.F., Adou Yao, C.Y., Bakayoko, A., Ipou, I.J., Bogaert, J., 2015. The present state of botanical investigations in Côte d'Ivoire. Biodiv. Inform. 10, 56–64.

Linder, H.P., 2001. Plant diversity and endemism in sub-Saharan tropical Africa. J. Biogeogr. 28, 169–182.

Linder, H.P., 2014. The evolution of African plant diversity. Front. Ecol. Evol. 2, 38. https://doi.org/10.3389/fevo.2014.00038.

Linnaeus, C., 1759. Flora Capensis. Dissertation by Wännman C.H., Uppsala. Subsequently published in Linnaeus, C. (1760). Amoen. Acad. 5, 353–370.

Lobo, J.M., Hortal, J., Yela, J.L., Millán, A., Sánchez-Fernández, D., García-Roselló, E., González-Dacosta, J., Heine, J., González-Vilas, L., Guisande, C., 2018. KnowBR: an application to map the geographical variation of survey effort and identify well-surveyed areas from biodiversity databases. Ecol. Indic. 91, 41–248.

Manish, K., Pandit, M.K., 2018. Geophysical upheavals and evolutionary diversification of plant species in the Himalaya. PeerJ 6, e5919.

Mayaux, P., Pekel, J.-F., Desclée, B., Donnay, F., Lupi, A., Achard, F., Clerici, M., Bodart, C., Brink, A., Nasi, R., Belward, A., 2013. State and evolution of the African rainforests between 1990 and 2010. Philos. Trans. R. Soc. B: Biol. Sci. 368 https://doi.org/10.1098/rstb.2012.0300.

McKnight, M.W., White, P.S., McDonald, R.I., Lamoreux, J.F., Sechrest, W., Ridgely, R.S., Stuart, S.N., 2007. Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. PLoS Biol. 5, 2424–2432.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.

Parmentier, I., Malhi, Y., Senterre, B., Whittaker, R.J., Alonso, A., Balinga, M.P.B., Bakayoko, A., Bongers, F., Chatelain, C., Comiskey, J.A., Cortay, R., Kamdem, M.-N. D., Doucet, J.-L., Gautier, L., Hawthorne, W.D., Issembe, Y.A., Kouamé, F.N., Kouka, L.A., Leal, M.E., Lejoly, J., Lewis, S.L., Nusbaumer, L., Parren, M.P.E., Peh, K. S.-H., Phillips, O.L., Poorter, L., Sheil, D., Sonké, B., Sosef, M.S.M., Sunderland, T.C. H., Stropp, J., ter Steege, H., Swaine, M.D., Tchouto, M.G.P., Gemerden, B.S.V., Van Valkenburg, J.L.C.H., Wöll, H., 2007. The odd man out? Might climate explain the lower tree α-diversity of African rain forests relative to Amazonian rain forests? J. Ecol. 95, 1058–1071.

Pteridophyte Phylogeny Group, 2016. A community-derived classification for extant lycophytes and ferns. J. Syst. Evol. 54, 563–603.

Pysek, P., Pergl, J., Essl, F., Lenzner, B., Dawson, W., Kreft, H., Weigelt, P., Winter, M., Kartesz, J., Nishino, M., Antonova, L.A., Barcelona, J.F., Cabezas, F.J., Cárdenas, D., Cárdenas-Toro, J., Castano, N., Chacón, E., Chatelain, C., Dullinger, S., Ebel, A.L., Figueiredo, E., Fuentes, N., Genovesi, P., Groom, Q.J., Henderson, L., Inderjit Kupriyanov, A., Masciadri, S., Maurel, N., Meerman, J., Morozova, O., Moser, D., Nickrent, D., Nowak, P.M., Pagad, S., Patzelt, A., Pelser, P.B., Seebens, H., Shu, W., Thomas, J., Velayos, M., Weber, E., Wieringa, J.J., Baptiste, M.P., van Kleunen, M., 2017. Naturalized alien flora of the world: species diversity, taxonomic and phylogenetic patterns, geographic distribution and global hotspots of plant invasion. Preslia 89, 203–274.

Qian, H., 1998. Large-scale biogeographic patterns of vascular plant richness in North America: an analysis at the generic level. J. Biogeogr. 25, 829–836.

Qian, H., 2009. Beta diversity in relation to dispersal ability for vascular plants in North America. Glob. Ecol. Biogeogr. 18, 327–332.

Qian, H., Fridley, J.D., Palmer, M.W., 2007. The latitudinal gradient of species-area relationships for vascular plants of North America. Am. Nat. 170, 690–701.

Qian, H., Deng, T., Beck, J., Sun, H., Xiao, C., Jin, Y., Ma, K., 2018. Incomplete species lists derived from global and regional specimen-record databases affect macroecological analyses: a case study on the vascular plants of China. J. Biogeogr. 45, 2718–2729.

Qian, H., Deng, T., Jin, Y., Mao, L., Zhao, D., Ricklefs, R.E., 2019. Phylogenetic dispersion and diversity in regional assemblages of seed plants in China. Proc. Natl. Acad. Sci. U. S. A. 116, 23192–23201.

Raven, P.H., Axelrod, D.I., 1974. Angiosperm biogeography and past continental movement. Ann. Mo. Bot. Gard. 61, 539–673.

Richards, P.W., 1973. Africa, the 'Odd man out'. In: Meggers, B.J., Ayensu, E.S., Duckworth, W.D. (Eds.), Tropical Forest Ecosystems of Africa and South America: A Comparative Review. Smithsonian Institution Press, Washintong DC.

Ricklefs, R.E., Lovette, I.J., 1999. The roles of island area per se and habitat diversity in the species-area relationships of four lesser Antillean faunal groups. J. Anim. Ecol. 68, 1142–1160.

Rosenzweig, M.L., 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge.

Sandel, B., Weigelt, P., Kreft, H., Keppel, G., van der Sande, M.T., Levin, S., Smith, S., Craven, D., Knight, T.M., 2020. Current climate, isolation and history drive global patterns of tree phylogenetic endemism. Glob. Ecol. Biogeogr. 29, 4–15.

Stropp, J., Ladle, R.J., Malhado, A.C.M., Hortal, J., Gaffuri, J., Temperley, W.H., Skøien, J.O., Mayaux, P., 2016. Mapping ignorance: 300 years of collecting flowering plants in Africa. Glob. Ecol. Biogeogr. 25, 1085–1096.

Takhtajan, A.L., 1986. Floristic Regions of the World. University of California Press, Berkeley.

Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., Berghe, E.V., Worm, B., 2010. Global patterns and predictors of marine biodiversity across taxa. Nature 466, 1098–1101.

Tutin, T.G., Heywood, V.H., Burges, N.A., Moore, D.M., Valentine, D.H., Walters, S.M., Webb, D.A., 1964-1980. Flora Europaea, vols. 1-5. Cambridge University Press, Cambridge.

Ulloa, C.U., Acevedo-Rodríguez, P., Beck, S., Belgrano, M.J., Bernal, R., Berry, P.E., Brako, L., Celis, M., Davidse, G., Forzza, R.C., Gradstein, S.R., Hokche, O., León, B., León-Yánez, S., Magill, R.E., Neill, D.A., Nee, M., Raven, P.H., Stimmel, H., Strong, M.T., Villaseñor, J.L., Zarucchi, J.L., Zuloaga, F.O., Jørgensen, P.M., 2017. An integrated assessment of the vascular plant species of the Americas. Science 358, 1614–1617.

van Kleunen, M., Dawson, W., Essl, F., Pergl, J., Winter, M., Weber, E., Kreft, H., Weigelt, P., Kartesz, J., Nishino, M., Antonova, L.A., Barcelona, J.F., Cabezas, F.J., Cárdenas, D., Cárdenas-Toro, J., Castaño, N., Chacón, E., Chatelain, C., Ebel, A.L., Figueiredo, E., Fuentes, N., Groom, Q.J., Henderson, L., Inderjit, Kupriyanov, A., Masciadri, S., Meerman, J., Morozova, O., Moser, D., Nickrent, D.L., Patzelt, A., Pelser, P.B., Baptiste, M.P., Poopath, M., Schulze, M., Seebens, H., Shu, W.-S., Thomas, J., Velayos, M., Wieringa, J.J., Pyšek, P., 2015. Global exchange and accumulation of non-native plants. Nature 525, 100–103.

van Kleunen, M., Pyšek, P., Dawson, W., Essl, F., Kreft, H., Pergl, J., Weigelt, P., Stein, A., Dullinger, S., König, C., Lenzner, B., Maurel, N., Moser, D., Seebens, H., Kartesz, J., Nishino, M., Aleksanyan, A., Ansong, M., Antonova, L.A., Barcelona, J.F., Breckle, S. W., Brundu, G., Cabezas, F.J., Cárdenas, D., Cárdenas-Toro, J., Castaño, N., Chacón, E., Chatelain, C., Conn, B., de Sá Dechoum, M., Dufour-Dror, J.-M., Ebel, A. L., Figueiredo, E., Fragman-Sapir, O., Fuentes, N., Groom, Q.J., Henderson, L., Inderjit, Jogan, N., Krestov, P., Kupriyanov, A., Masciadri, S., Meerman, J., Morozova, O., Nickrent, D., Nowak, A., Patzelt, A., Pelser, P.B., Shu, W.-S., Thomas, J., Uludag, A., Velayos, M., Verkhosina, A., Villaseñor, J.L., Weber, E., Wieringa, J.J., Yazlik, A., Zeddam, A., Zykova, E., Winter, M., 2019. The Global Naturalized Alien Flora (GloNAF) database. Ecology 100, e02542.

Vetaas, O.R., Grytnes, J.-A., 2002. Distribution of vascular plant species richness and endemic richness along the Himalayan elevation gradient in Nepal. Glob. Ecol. Biogeogr. 11, 291–301.

Victor, J.E., Smith, G.F., Van Wyk, A.E., 2016. History and drivers of plant taxonomy in South Africa. Phytotaxa 269, 193–208.

White, F., 1983. The Vegetation of Africa. Unesco, Paris.

Whitmore, T.C., 1998. An Introduction to Tropical Rain Forests. Clarendon Press, Oxford.

Wiersema, J.H., León, B., 2013. World Economic Plants: A Standard Reference, 2nd edn. CRC Press.

Wilkinson, L., Hill, M., Welna, J.P., Birkenbeuel, G.K., 1992. SYSTAT for Windows: Statistics. SYSTAT Inc., Evanston.

Yang, W., Ma, K., Kreft, H., 2013. Geographical sampling bias in a large distributional database and its effects on species richness-environment models. J. Biogeogr. 40, 1415–1426.

Zhang, J., Qian, H., Girardello, M., Pellissier, V., Nielsen, S.E., Svenning, J.-C., 2018. Trophic interactions among vertebrate guilds and plants shape global patterns in species diversity. Proc. R. Soc. Lond. B 285, 20180949.

Zhou, Y., Liu, B., Mbuni, Y., Yan, X., Mwachala, G., Hu, G., Wang, Q., 2017. Vascular flora of Kenya, based on the Flora of tropical East Africa. PhytoKeys 16, 113–126.

Zhou, Y.-D., Boru, B.H., Wang, S.-W., Wang, Q.-F., 2021. Species richness and phylogenetic diversity of different growth forms of angiosperms across a biodiversity hotspot in the horn of Africa. J. Syst. Evol. 59, 141–150.