

Bioquality hotspots in the tropical African flora

Cicely A. M. Marshall^{1*}, Jan J. Wieringa² & William D. Hawthorne¹

¹Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, United Kingdom.

²Naturalis Biodiversity Center, National Herbarium of the Netherlands, Darwinweg 2, 2333 CR Leiden, The Netherlands.

*Correspondence to: Cicely Marshall, email: cicely.marshall@plants.ox.ac.uk; Telephone: 01865 275 140 or 01865 244 460.

Summary

Identifying areas of high biodiversity is an established way to prioritize areas for conservation [1–3], but global approaches have been criticized for failing to render global biodiversity value at a suitable scale for local management [4–6]. We assembled 3.1 million species distribution records for 40,583 vascular plant species of tropical Africa from sources including plot data, herbarium databases, checklists, and GBIF, and cleaned the records for geographic accuracy and taxonomic consistency. We summarised the global ranges of tropical African plant species into four, weighted, categories of global rarity called Stars. We applied the Star weights to summaries of species distribution data at fine resolutions to map the bioquality (range restricted global endemism) of areas [7]. We generated confidence intervals around bioquality scores to account for the remaining uncertainty in the species inventory. We confirm the broad significance of the Horn of Africa, Guinean forests, coastal forests of east Africa, and Afromontane regions for plant biodiversity, but reveal also the variation in bioquality within these broad regions and others, particularly at local scales. Our framework offers practitioners a quantitative, scalable and replicable approach for measuring the irreplaceability of particular local areas for global biodiversity conservation, and comparing those areas within their global and regional context.

Results & Discussion

Distribution data for tropical African plants

Biodiversity hotspots were originally identified using the richness of species endemic to large, biogeographic realms which had been significantly degraded [1], largely because species distribution data were available only at this coarse resolution [8]. This situation has improved rapidly as online public repositories (e.g. GBIF), collection digitisation efforts (e.g. JStor's Global Plants Initiative), and data journals (e.g. Check List) were established, increasing the available number of geolocated species records.

We assembled 3.1 million global species distribution records for tropical African vascular plants, from plot data, herbarium databases, checklists, and the Global Biodiversity Information Facility (GBIF). We limited our GBIF search to records supported by herbarium specimens and those without reported geographic issues. Our tropical African species list was derived from the African Plants Database, and includes 40,583 accepted species or intraspecific names which were checked for synonymy and comprehensiveness against other resources. We refer to species for simplicity, but all analysis was conducted on the lowest named taxonomic unit at or below the species level.

Of the 3.1 million distribution records, 0.5 million specimens were collected without coordinates. We geolocated these records by comparing the text locality information provided in the collectors' notes to standardised gazetteer dictionary files, and assigned to the records either point-with-radius, or polygon, coordinates, depending on the detail available in the notes. Records assigned polygons were included in the following analyses if they fitted inside the sampling units in question. We used similar geolocation methods to detect records for which the supplied coordinates and supplied text locality information conflicted; such records were checked and corrected by hand or were omitted from analysis. Many older specimens, often including types, were collected without coordinates. Using these novel methods we were able to compile records for almost all vascular plant taxa present in tropical Africa, ensure taxonomic consistency and geographic accuracy, and respect the geographic resolution of the original collection in the analysis.

We estimated the completeness of our species distribution data by comparing our species sampling levels against published estimates of species richness [9] (Figure 1A). There are many areas for which species sampling is far from complete, particularly for central Africa [10]. We must continue our efforts to fill these data gaps, but we cannot afford to ignore the biogeographic signal present in existing data or the plants we seek to record will be gone.

Star rating: Species-level conservation assessment

We summarised the global range for all plant species in tropical Africa into 4 categories of global range, called Stars [7] (Figure 2). Globally rare species are the important elements of biodiversity to conserve locally, in order to conserve species richness globally. Black Star species have the narrowest global ranges (c. 2.7 degree squares occupancy on average), Green Star species are the globally commonest (c. 72 degree squares), and Gold and Blue Star species are intermediate. Star ratings are species-specific, mutually exclusive and globally applicable, so that each species or intraspecific taxon in the world can have only one Star. Global ranges were categorized, rather than using a continuous occupancy metric, to produce a memorable framework which retains the necessary subtlety to reveal robust biogeographic patterns. Given that the full degree square occupancy of all species globally is not yet known (Figure 1A), the categorical system also allows for interpretation of the appropriate Star rating for species which are inadequately represented in herbaria, for example due to geographic or ecological biases in collections. We reviewed each species' Star in light of the best available information from online floras and other botanic resources, unless it was already a Green Star species (globally widespread). Although this introduces a degree

of subjectivity to the system, the results better reflect the true breadth of knowledge regarding species' distributions than a strict reliance on digitised records would. Each Star category carries a weight which is inverse to the mean range (measured as degree square occupancy) for all the included species of that Star category, so that rarer species and Stars have a higher weight (see Supplemental Experimental Procedures).

Star rating can be compared with the IUCN Red Listing approach when criterion B2 (AOO) is invoked [11], but Star rating requires no explicit measure of population change, regional ratings are not necessary or allowed, and the grid size for AOO calculations is standardised to one degree square (or 100 x 100 km, whichever is larger), for all plant species. Globally, three times as many vascular plant species have a Star rating compared with a Red List Category (62,868 cf. 20,147; 100% cf. 8% tropical African plant species assessed). Star rating offers a biologically pure assessment of a species' range which is relatively fast to conduct, and is useful for scientific analyses of distribution patterns as well as conservation assessment. As a consequence of this study, all tropical African vascular plant species have a Star rating, so the system can now be used to support or prioritise conservation anywhere in tropical Africa, and could be extended to other taxa.

Bioquality hotspots in tropical Africa

We used the Star ratings and species distribution summary tables to produce a quantitative measure of plant biodiversity value for areas across tropical Africa. Indexes respecting species global ranges reflect a particular component of what specialists tend to recognise as the biodiversity value of a place. We refer to this attribute of plant biodiversity as *bioquality*, and the particular index used to measure bioquality is the Genetic Heat Index (GHI) [7,12]. GHI is calculated for a unique species list for an area, by averaging over the weights of the Stars for those species found in the area. An area with a high proportion of globally rarer species in its flora achieves a high GHI and a high bioquality hotspot score.

This is similar to calculating range-size rarity [13,14], except we measured ranges globally rather than within the study area, to produce scores which are comparable globally. Range size rarity uses the continuous degree square occupancy of species, whereas we have binned ranges into the four Star categories to produce results which are not artificially precise, given that the full degree square occupancy of all species is not yet known (Figure 1A). The biggest difference is that the GHI divides by the number of species present (to produce a weighted average), which means that the GHI does not measure richness or diversity. This has the possible disadvantage that areas with high absolute numbers of rare species achieve lower GHI scores if they also include many common species, but a number of significant advantages: Areas are not downgraded if their species inventory is not complete, making the measure robust to missing data. GHI scores decrease where vegetation is invaded by globalized species. Species richness increases with the size of area under consideration: ignoring richness means that GHI scores can be calculated and meaningfully compared for areas of any shape or size, including the very local.

To conserve species globally, it is not important to prioritise individual areas with high species richness. Rather, it is important to protect areas where a high proportion of the individuals belong to globally rare species, otherwise those species would be lost from the global species pool [15]. In fact, the number of species in an area, whether rare, threatened or simply present (richness), is now generally recognised as a poor metric for identifying conservation priorities, because richness alone reveals little more than the availability of data, the size and shape of the area under consideration [5], and the biome type.

When GHI is calculated from an essentially complete species list for an area, then confidence intervals are not necessary. Neither would they be necessary if species were sampled incompletely

but representatively with respect to the true balance of Stars in the full flora, because as a weighted average the GHI includes no measure of richness. However, we cannot tell whether the recorded flora is currently biased towards the globally rarest (or commonest) species. We therefore estimated bootstrapped confidence intervals for the GHI for each degree square, given the apparent GHI (Figure 1C) and current estimated species sampling completeness (Figure 1A), to produce a confidence interval within which the true GHI value of each area is expected fall, even if sampling is currently biased with respect to Star (Figure 1B and 1D). This is one way in which uncertainty can be quantified and reliable conclusions drawn, whilst the species inventory is incomplete.

Figure 1B reveals tropical Africa's biodiversity patterns in their most complete, repeatable, and intimate detail yet. On the whole, the results fit comfortably with previous studies of the distribution of Africa's plant biodiversity [13,16–20], by highlighting the generally rather low endemism in the Sahara, Sahel and Sudanian regions, and medium to high endemism for the Guineo-Congolian, Zambezian, Somalia-Masai, Karoo-Namib, Zanzibar-Inhambane and Afromontane regions. The Somalia-Masai (Horn of Africa) flora comes out as one of the hottest floras in tropical Africa; while the large number of endemic species has been recognised [21], Somalia's high bioquality has perhaps been underappreciated relative to Africa's wetter and montane forest regions [13], most likely due to undersampling and relatively lower species richness.

Smaller scale bioquality hotspots are visible around Mount Cameroon, Mount Mulanje and Mount Chimanimani. In Guineo-Congolia, bioquality peaks in the high rainfall forests of Cameroon and Gabon towards the coasts, is higher for western Upper Guinea than in the east, and bioquality is somewhat lower but comparable for Congolia (though data are sparser). Bioquality peaks in the Zambezian region in south eastern Democratic Republic of Congo, and in southwest central Angola. For the Karoo-Namib, the coastline of southern Angola is particularly hot; the flora of the eastern coast of Africa (Zanzibar-Inhambane regional mosaic) is particularly hot in south east Tanzania.

Bioquality at local scales

Our bioquality metric (GHI) is based on a weighted average of globally rare plants, and as proportions scale meaningfully with richness and area, the scale (grain) and shape of sampling units for an analysis can be matched to its application. The data for such fine-scale bioquality analyses can be derived for a project area by on-the-ground sampling. In particular, Rapid Botanic Survey is a botanical survey technique specifically designed to collect this information with the minimum possible effort (see Supplemental Experimental Procedures), although a meaningful GHI score can be calculated from any reasonably taxonomically complete survey data e.g. relevés or all-species transects [22].

Figure 3 reveals the local variation in bioquality found by local sampling within one of these degree squares, around Yepeka (Nimba mountains, northern Liberia), and across different vegetation types and altitudes. Such local-scale information is particularly useful for land management planning. Bioquality around the Nimba mountains is lower for the more populated, lowland area around the central road corridor, and peaks in the closed canopy slope forests at higher elevations, with some variation apparent even within this forest type. It is clear that this 'hotspot' at the one degree square scale is a patchwork of hot and cold spots at a finer scale. It is useful to be able to measure how hot an area is at this rather local scale, because it is at this scale where decisions impacting biodiversity are often taken.

The background map shows minGHI at 0.5 x 0.5 degree square resolution, and reveals bioquality patterns in greater detail than the one degree map of Figure 1B, although fewer data points can be resolved to this higher resolution grid.

Bioquality as a conservation framework

Bioquality is measured using the global range of plant species. Vascular plants are often used as an indicator taxon for biodiversity measurements because they are relatively well known taxonomically and geographically, and define the terrestrial habitats in which other taxa live. If high bioquality is used to define priorities for conservation, or to inform local land management, it makes sense to consider many other aspects of an area [23], including species other than plants [24], social factors [25], economic cost/benefit analyses [26], ecosystem-wide benefits [27], phylogenetic diversity and evolutionary processes [28], and rates or risk of habitat loss [1]. We keep such measures out of our plant bioquality analysis, and promote viewing them as independent GIS layers, because mixing criteria in a single metric makes results harder to interpret and to make globally consistent. We accept that the proportion of globally rare plant species in a flora is by no means the only important factor when designing a land management plan, but it is a critical one.

As a consequence of this study, all mainland tropical African plant taxa have a Star rating and GHIs can now be calculated easily anywhere in tropical Africa where the species composition is at least partly known. This should prove useful in the context of Environmental Impact Assessment, or Protected Area planning, because a local scale hotspot map and database can: Describe a baseline; inform the positioning of infrastructure or protected areas; identify appropriate offset areas; allow precise monitoring of impacts and changes through time (with resurvey); and help devise management plans for the globally rarest species. We accept as a premise of the system that the data are never complete, and that taxonomic boundaries also shift, so the system is built to be robust in light of new information.

As much as 79% of Earth's land surface has now been prioritized for conservation under one system or another [8], and we do not wish to define yet another set of broad areas of conservation importance. Instead, our framework offers conservationists and land managers a quantitative and replicable approach for measuring the irreplaceability of particular local areas for global biodiversity conservation, and comparing those areas within their global and regional context.

Author contributions

Conceptualisation: C.M. and W.H.; Methodology and Software: C.M and W.H.; Resources and Data Curation: C.M. J.W, W.H.; Writing – Original Draft: C.M.; Writing – Review & Editing: C.M, J.W., W.H.

Acknowledgments

Cyrille Chatelain provided the African Plants Database Tropical Africa species list. Data capture, editing and broadcast of the African Plants Database is the product of a collaboration between the South African National Biodiversity Institute, the Conservatoire et Jardin botaniques de la Ville de Genève, Tela Botanica and the Missouri Botanical Garden. BISAP (Biogeographical Information System on African Plant Diversity) has been established at the Nees Institute, University of Bonn in the context of BIOTA Africa project together with several external partners, especially Jon Lovett (University of Leeds) and Peter Linder (Zurich). For more details about the dataset and a full acknowledgement of all contributors please refer to Linder et al. 2005 [29] and Küper et al. 2006 [30]. We thank M. Swaine for data and discussions; S. Harris, A. Hector, L. Turnbull, L. Hill, Z. Goodwin and R. Scotland for comments and discussions; D. Filer and A. Liddell for the TOPO website and the Star Server. Survey work in Nimba County was funded by ArcelorMittal Liberia and Euronimba Liberia Ltd in the context of Environmental and Social Impact Assessments. C.M. acknowledges support from the Clarendon Fund and Merton College; W.H. acknowledges support from the James Martin 21st Century School and BP Biofuels.

References

- [1] Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858.
- [2] Olson, D.M., and Dinerstein, E. (2002). The Global 200: Priority Ecoregions for Global Conservation. *Ann. Mo. Bot. Gard.* 89, 199.
- [3] Stattersfield, A. J., Crosby, M. J., Long, A. J. and Wege, D. C. (1988). Endemic Bird Areas of the World: Priorities for Biodiversity Conservation. BirdLife International (Cambridge, UK).
- [4] Mace, G.M., Balmford, A., Boitani, L., Cowlishaw, G., Dobson, A.P., Faith, D.P., Gaston, K.J., Humphries, C.J., Vane-Wright, R.I., Williams, P.H., et al. (2000). It's time to work together and stop duplicating conservation efforts *Nature* 405, 393–393.
- [5] Brummitt, N., and Lughadha, E.N. (2003). Biodiversity: Where's Hot and Where's Not. *Conserv. Biol.* 17, 1442–1448.
- [6] da Fonseca, G.A.B., Balmford, A., Bibby, C., Boitani, L., Corsi, F., Brooks, T., Gascon, C., Olivieri, S., Mittermeier, R.A., Burgess, N., et al. (2000). ... following Africa's lead in setting priorities. *Nature* 405, 393–394.
- [7] Hawthorne, W.D. (1996). Holes and the sums of parts in Ghanaian forest: regeneration, scale and sustainable use. *Proc. R. Soc. Edinb. Sect. B Biol. Sci.* 104, 75–176.
- [8] Brooks, T.M., Mittermeier, R.A., Fonseca, G.A.B. da, Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D., and Rodrigues, A.S.L. (2006). Global Biodiversity Conservation Priorities. *Science* 313, 58–61.
- [9] Barthlott, W., Mutke, J., Rafiqpoor, D., Kier, G., and Kreft, H. (2005). Global Centers of Vascular Plant Diversity. *Nova Acta Leopoldina NF 92* 342, 61–83.
- [10] Stropp, J., Ladle, R.J., M. Malhado, A.C., Hortal, J., Gaffuri, J., H. Temperley, W., Olav Skøien, J., and Mayaux, P. (2016). Mapping ignorance: 300 years of collecting flowering plants in Africa. *Glob. Ecol. Biogeogr.* 25, 1085–1096.
- [11] IUCN. The IUCN Red List of Threatened Species Version 2015-4, 2016. <http://www.iucnredlist.org/#>.
- [12] Baksh-Comeau, Y.S., Maharaj, S.S., Adams, C.D., Harris, S.A., Filer, D.L., and Hawthorne, W.D. (2016). An annotated checklist of the vascular plants of Trinidad and Tobago with analysis of vegetation types and botanical “hotspots.” *Phytotaxa* 250, 1–431.
- [13] Küper, W., Sommer, J.H., Lovett, J.C., Mutke, J., Linder, H.P., Beentje, H.J., Rompaey, R.S.A.R.V., Chatelain, C., Sosef, M., and Barthlott, W. (2004). Africa's Hotspots of Biodiversity Redefined. *Ann. Mo. Bot. Gard.* 91, 525–535.
- [14] Williams, P., Gibbons, D., Margules, C., Rebelo, A., Humphries, C., and Pressey, R. (1996). A Comparison of Richness Hotspots, Rarity Hotspots, and Complementary Areas for Conserving Diversity of British Birds. *Conserv. Biol.* 10, 155–174.
- [15] Mittermeier, R.A., Turner, W.R., Larsen, F.W., Brooks, T.M., and Gascon, C. (2011). Global Biodiversity Conservation: The Critical Role of Hotspots. In *Biodiversity Hotspots*, F.E. Zachos, and J.C. Habel, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 3–22.
- [16] Linder, H.P. (2001). Plant diversity and endemism in sub-Saharan tropical Africa. *J. Biogeogr.* 28, 169–182.
- [17] Wieringa, J.J., and Poorter, L. (2004). Biodiversity hotspots in West Africa; patterns and causes. In *Biodiversity of West African Forests: An Ecological Atlas of Woody Plant Species*, L. Poorter, F. Bongers, F. N'. Kouamé, and W.D. Hawthorne, ed. (CABI), pp. 61–72.
- [18] Lovett, J.C., Rudd, S., Taplin, J., and Frimodt-Møller, C. (2000). Patterns of plant diversity in Africa south of the Sahara and their implications for conservation management. *Biodivers. Conserv.* 9, 37–46.
- [19] White, F. (1993). The AETFAT Chorological Classification of Africa: History, Methods and Applications. *Bull. Jard. Bot. Natl. Belg. Bull. Van Natl. Plantentuin Van Belg.* 62, 225.
- [20] Lovett, J.C. (1998). Importance of the Eastern Arc Mountains for Vascular Plants. *J. East Afr. Nat. Hist.* 87, 59–74.

- [21] Thulin, M. (1993). Flora of Somalia, Volume 1 (Royal Botanic Gardens, Kew).
- [22] Moore, J.J. (1962). The Braun-Blanquet System: A Reassessment. *J. Ecol.* 50, 761–769.
- [23] Hunter, M.L., and Hutchinson, A. (1994). The Virtues and Shortcomings of Parochialism: Conserving Species That Are Locally Rare, but Globally Common. *Conserv. Biol.* 8, 1163–1165.
- [24] Grenyer, R., Orme, C.D.L., Jackson, S.F., Thomas, G.H., Davies, R.G., Davies, T.J., Jones, K.E., Olson, V.A., Ridgely, R.S., Rasmussen, P.C., et al. (2006). Global distribution and conservation of rare and threatened vertebrates. *Nature* 444, 93–96.
- [25] Cincotta, R.P., Wisniewski, J., and Engelman, R. (2000). Human population in the biodiversity hotspots. *Nature* 404, 990–992.
- [26] O'Connor, C., Marvier, M., and Kareiva, P. (2003). Biological vs. social, economic and political priority-setting in conservation. *Ecol. Lett.* 6, 706–711.
- [27] Loreau, M. (2001). Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges. *Science* 294, 804–808.
- [28] Vane-Wright, R.I., Humphries, C.J., and Williams, P.H. (1991). What to protect?—Systematics and the agony of choice. *Biol. Conserv.* 55, 235–254.
- [29] Linder, H.P., Lovett, J.C., Mutke, J.M., Barthlott, W., JÜRGENS, N., Rebelo, T., and Küper, W. (2005). A numerical re-evaluation of the sub-Saharan phytochoria of mainland Africa. *Biol. Skr.* 55, 229–252.
- [30] Küper, W., Sommer, J.H., Lovett, J.C., and Barthlott, W. (2006). Deficiency in African plant distribution data – missing pieces of the puzzle. *Bot. J. Linn. Soc.* 150, 355–368.

Figure legends

Figure 1. Bioquality hotspots in the tropical African flora. A: Ratio of species richness in our database to total species richness estimated using Barthlott et al. 2005 [9]. B: Bioquality mapped at one degree square resolution using minGHI, a reliable minimum estimate of GHI; minGHI is a conservative GHI estimate expected to be closer to the true GHI if collections are currently biased towards globally rare species. C: GHI values for bioquality (assumes no species sampling bias with respect to Star). D: maxGHI, maximum likely GHI assuming species sampling is currently biased towards the globally commonest species (probably the least likely scenario). Confidence intervals (minGHI to maxGHI) are larger where species sampling is poorer (compare panels A, B, C, D). ‘True’ GHI values, assuming perfect collection, would fall between minGHI and maxGHI estimates for each cell. See also Tables S1 and S2 Excel files.

Figure 2. Example distribution patterns for a species of each Star. Black Star species occupy on average 2.7 degree squares globally. Gold Star species occupy 8, Blue Star species occupy 24, and Green Star species occupy 72 degree squares globally (or 100 x 100 km, whichever is the larger). Mapped distribution for *Allophylus africanus* includes distribution data for named formas and varieties. See also Table S1 Excel file.

Figure 3. Bioquality at local scales. GHI calculated from 310 Rapid Botanic Survey (RBS) samples across northern Nimba County, Liberia. The hotter GHI (>200) scores equivalent to the minGHI estimate for the degree square as a whole were recorded in forest in this region, although not all the forest had such a high GHI. Background map shows minGHI for 0.5 x 0.5 degree squares. See also Tables S1 and S2 Excel files.