Language extinction triggers the loss of unique medicinal knowledge

Rodrigo Cámara-Leret[a,1] and Jordi Bascompte[a]

[a]Department of Evolutionary Biology and Environmental Studies, University of Zurich, CH-8057 Zurich, Switzerland

Over 30% of the 7,400 languages in the world will no longer be spoken by the end of the century. So far, however, our understanding of whether language extinction may result in the loss of linguistically unique knowledge remains limited. Here, we ask to what degree indigenous knowledge of medicinal plants is associated with individual languages and quantify how much indigenous knowledge may vanish as languages and plants go extinct. Focusing on three regions that have a high biocultural diversity, we show that over 75% of all 12,495 medicinal plant services are linguistically unique—i.e., only known to one language. Whereas most plant species associated with linguistically unique knowledge are not threatened, most languages that report linguistically unique knowledge are. Our finding of high uniqueness in indigenous knowledge and strong coupling with threatened languages suggests that language loss will be even more critical to the extinction of medicinal knowledge than biodiversity loss.

indigenous languages | ecosystem services | biocultural diversity

Indigenous people have accumulated a sophisticated knowledge about plants and their services—including knowledge that confers significant health benefits (1)—that is encoded in their languages (2). Indigenous knowledge, however, is increasingly threatened by language loss (3) and species extinctions (4, 5). On one hand, language disuse is strongly associated with decreases in indigenous knowledge about plants (6). On the other hand, global change will constrain the geographic ranges of many human-utilized endemic plants and crops (7, 8). Together, language extinction and reductions in useful plant species within the coming century may limit the full potential of nature's contributions to people and the discovery of unanticipated uses. So far, however, our understanding of the degree to which the loss of indigenous languages may result in the loss of linguistically unique knowledge and how this risk compares to that posed by ecological extinction has been limited (Fig. 1).

Unraveling the structure of indigenous knowledge about medicinal services has important implications for its resilience (9). Most indigenous cultures transmit knowledge orally (10). Therefore, if knowledge about medicines is shared widely among indigenous groups that speak different languages, knowledge resilience would be high. That is, even if some indigenous languages go extinct, their medicinal plant knowledge would still be safeguarded in other surviving languages with whom such knowledge is shared. To assess the extent of this, we analyzed three large ethnobotanical datasets for North America (11), northwest Amazonia (12), and New Guinea (13). Together, these data span 3,597 medicinal plant species and 12,495 plant services associated with 236 indigenous languages (*Materials and Methods*). We defined a "medicinal plant service" as the combination of a plant species and a medicinal subcategory (e.g., *Ficus insipida* + Digestive System).

Our results show that indigenous knowledge about medicinal plants exhibits a strong pattern of linguistic uniqueness in all regions, with 73%, 91%, and 84% of the medicinal services in North America, northwest Amazonia, and New Guinea being

cited by only one language, respectively (Fig. 2). The fraction of unique knowledge that is explained by cultural turnover (i.e., languages having different knowledge of the same species), as opposed to species turnover (i.e., each language using different plant species), is 56%, 18%, and 50% in North America, northwest Amazonia, and New Guinea, respectively. Our finding of a strong pattern of unique knowledge raises the question of whether unique knowledge is mostly found in languages that are threatened.

Our analysis indicates that threatened languages support 86% and 100% of all unique knowledge in North America and northwest Amazonia, respectively. This result highlights that the Americas are an indigenous knowledge hotspot (i.e., most medicinal knowledge is linked to threatened languages) and, thus, a key priority area for future documentation efforts. By contrast, threatened languages account for 31% of all unique knowledge in New Guinea. However, in the absence of an island-wide linguistic survey, the true status of New Guinea's languages is difficult to assess. A recent study in Papua New Guinea showed that only 58% of 6,190 students, compared to 91% of their parents, are fluent in indigenous languages (14). Crucially, the varied medicinal plant uses known to students fluent in indigenous languages are replaced by a few uses mostly concentrated in nonnative plant species in the students who do not speak indigenous languages. Such a dramatic decline in language skills in a single generation suggests that our language-threat data from Glottolog (15) likely underestimate the percentage of unique knowledge associated with threatened languages in New Guinea.

> ## Significance
>
> The United Nations proclamation of 2022–2032 as the International Decade of Indigenous Languages aims to raise global awareness about their endangerment and importance for sustainable development. Indigenous languages contain the knowledge that communities have about their surrounding plants and the services they provide. The use of plants in medicine is a particularly relevant example of such ecosystem services. Here, we find that most medicinal knowledge is linguistically unique—i.e., known by a single language—and more strongly associated with threatened languages than with threatened plants. Each indigenous language is therefore a unique reservoir of medicinal knowledge—a Rosetta stone for unraveling and conserving nature's contributions to people.

[1]To whom correspondence may be addressed. Email: rodrigo.camaraleret@ieu.uzh.ch.

Fig. 1. Medicinal plant knowledge and its association with indigenous languages. The figure illustrates a regional pharmacy with remedies (jars with plants) cited by languages (jar labels). In this paper, we assess to what degree the knowledge contained in this pharmacy would be eroded by the extinction of either indigenous languages or plants.

Once we have quantified the overall amount of unique knowledge, we next proceed by mapping how it is distributed across the linguistic phylogeny. This will serve to identify whether unique knowledge is uniformly distributed across all linguistic groups or whether a few linguistic groups deserve more protection than others. First, we built language phylogenies for all of the indigenous languages in our sample. Next, we calculated the degree of phylogenetic clustering of unique knowledge using Pagel's lambda (λ) (16); values of λ close to one indicate strong phylogenetic clustering, whereas values close to zero indicate data without phylogenetic dependence. We did not find clustering of unique knowledge along the language phylogenies in any of the

three regions (Fig. 3; *SI Appendix*, Table S1). This indicates that when planning for medicinal knowledge conservation, the entire linguist spectrum—rather than a few "hot" branches—needs to be considered.

So far, we have focused on how unique knowledge is distributed along the cultural dimension. Let us turn now to examine the other component of the indigenous knowledge network (9), namely, the plants. To understand the degree of threat faced by medicinal plants, we queried the International Union for Conservation of Nature (IUCN) Red List of Threatened Species (17). We found conservation assessments for 22%, 31%, and 32% of the medicinal species recorded in North
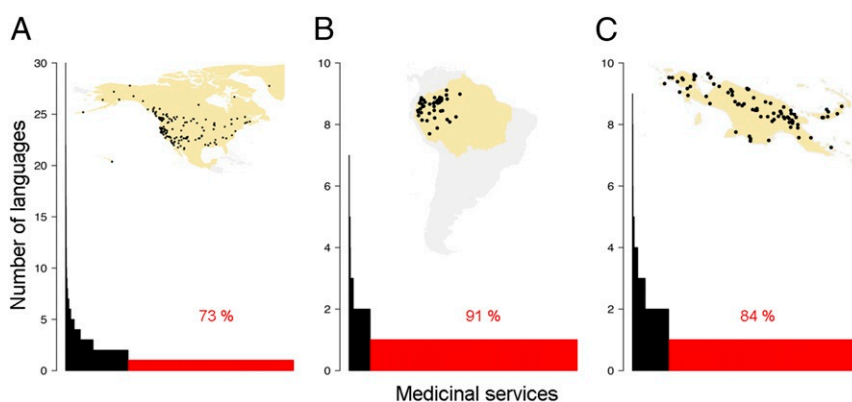
Fig. 2. Most medicinal knowledge is unique to a single language. Histograms depict the number of indigenous languages that cite a medicinal service. (*A*) North America. (*B*) Northwest Amazonia. (*C*) New Guinea. Red bars show medicinal plant services only known to one language. Dots within the maps indicate the distribution of languages.
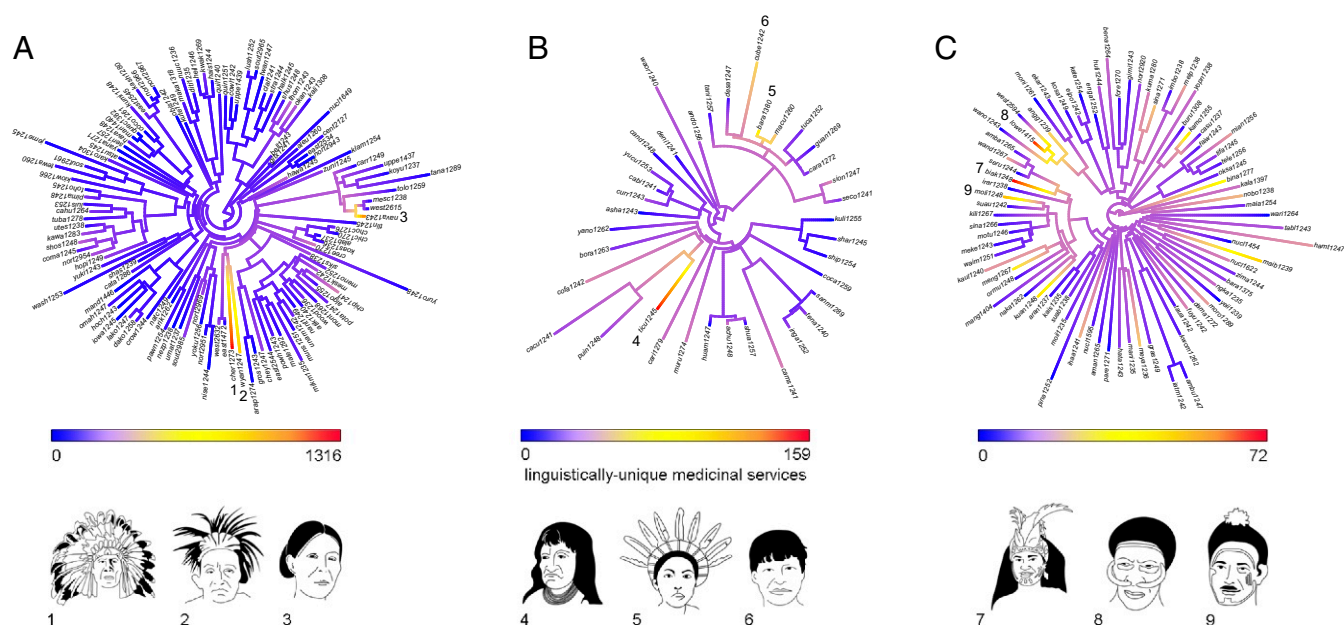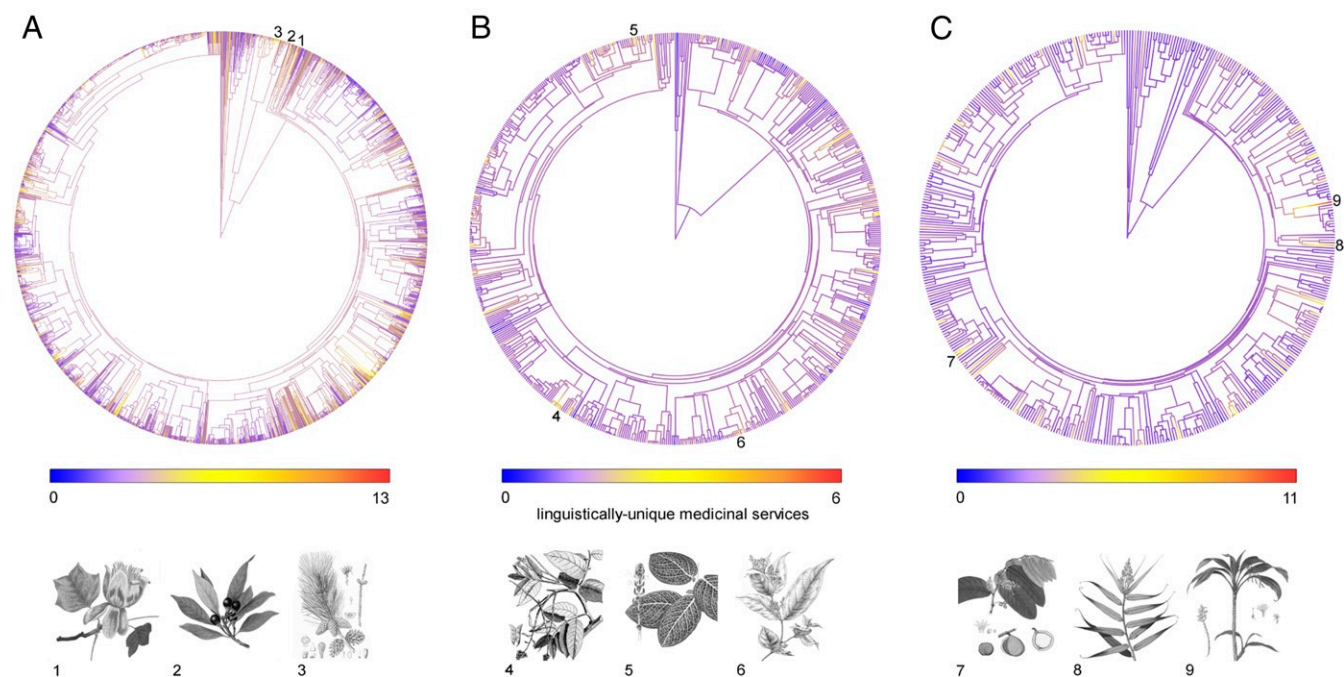
Fig. 3. Distribution of unique knowledge across languages. Trees represent language phylogenies of North America (*n* = 119 languages) (*A*); northwest Amazonia (*n* = 37 languages) (*B*); and New Guinea (*n* = 80 languages) (*C*). Illustrations represent indigenous groups whose languages have the highest number of unique medicinal services per region. These languages are indicated by their corresponding numbers in the linguistic trees: 1, Cherokee; 2, Huron–Wyandot; 3, Navajo; 4, Ticuna; 5, Barasana–Eduria; 6, Cubeo; 7, Biak; 8, Lower Grand Valley Dani; and 9, Molima. Language names at phylogeny tips are abbreviated following Glottolog codes. For the list of language names and Glottolog codes, see *SI Appendix,* Table S2.

America, northwest Amazonia, and New Guinea, respectively. Of the total medicinal flora with IUCN assessments, 4%, 1%, and 4% were classified as threatened in North America, northwest Amazonia, and New Guinea, respectively (*Materials and Methods*). To ascertain whether the observed patterns may change as more species are formally assessed, we also obtained conservation predictions from a machine-learning study (18) (*Materials and Methods*), which contains assessments for 57%, 25%, and 49% of the medicinal species recorded in North America, northwest Amazonia, and New Guinea, respectively. According to

Fig. 4. Distribution of unique knowledge across medicinal floras. Trees represent medicinal plant phylogenies of North America (*n* = 2,475 species) (*A*); northwest Amazonia (*n* = 645 species) (*B*); and New Guinea (*n* = 477 species) (*C*). Illustrations and their corresponding numbers show the plant species with more unique medicinal services per region. 1, *Liriodendron tulipifera*; 2, *Persea borbonia*; 3, *Pinus glabra*; 4, *Tachigali paniculata*; 5, *Fittonia albivenis*; 6, *Tetrapterys styloptera*; 7, *Inocarpus fagifer*; 8, *Flagellaria indica*; and 9, *Cordyline fruticosa*. All illustrations from www.plantillustrations.org belong to the public domain.

ECOLOGY

SUSTAINABILITY SCIENCE

that study, the probability of a medicinal species belonging to a threatened category ranged from 0.0002 to 0.8341 in North America (mean ± SD, 0.156 ± 0.158), 0.149 to 0.822 in northwest Amazonia (mean 0.483 ± 0.119), and 0.063 to 0.679 in New Guinea (mean 0.357 ± 0.141), respectively. In summary, both the IUCN conservation assessments and machine-learning predictions suggest that most medicinal plant species in our sample are not threatened. Finally, we found that about 1% of all unique knowledge in each region was associated with both threatened languages and threatened plants (*SI Appendix*, Table S3). However, there is considerable uncertainty about the potential loss of unique knowledge from the extinction of plants because 64% and 69% of the unique knowledge in North America and northwest Amazonia that is associated with threatened languages belongs to plants that lack plant-conservation assessments. IUCN conservation assessments are urgently needed for these plant species.

To assess whether unique knowledge is strongly clustered biologically, we built phylogenies of the medicinal floras of each region and calculated Pagel's lambda (Fig. 4). We only found significant clustering of unique knowledge in North America, although values were low (*SI Appendix*, Table S1). This relatively weak phylogenetic signal across the three regions suggests that when planning for biocultural conservation, the entire medicinal flora—rather than a few clades—must be considered.

Here, we have shown that in North America, northwest Amazonia, and New Guinea, indigenous knowledge of medicinal plant services exhibits a low redundancy across languages that is typical of systems with high information content (19, 20). This low redundancy in medicinal knowledge among languages does not support the notion of high cross-cultural consensus—i.e., that cultures resemble each other in their knowledge—but instead highlights the unique biocultural heritage each culture holds. The invention and diversification of languages involve two opposing forces. On the one hand, sharing facilitates the exchange of information and the spread of valuable ideas that may enhance the fitness within populations. On the other hand, the diversification of languages is the result of innovations, and eventually linguistic barriers may limit information spread. In areas of high linguistic or biological diversity and/or geographic barriers, the balance between sharing and innovating may tip toward the latter. This may result in the amplification of differences among cultures, as we have shown here for the case of medicinal knowledge.

Only about 6% of higher plants have been screened for biological activity (21). Therefore, assessing to what degree linguistically unique medicinal services are truly effective in the Western sense is beyond the scope of this paper. In many instances, these plants have been proven medicinally effective (12, 22–27), albeit there are also exceptions (28, 29). Regardless of that, here, we treat this knowledge as what it is: part of the cultural heritage of indigenous people.

The United Nations declared 2022–2032 as the International Decade of Indigenous Languages to raise awareness about their importance for sustainable development and their endangerment across the world. Our study suggests that each indigenous language brings unique insights that may be complementary to other societies that seek potentially useful medicinal remedies. Therefore, the predicted extinction of up to 30% of indigenous languages by the end of the 21st century (3) would substantially compromise humanity's capacity for medicinal discovery.

Materials and Methods

Plant Services. We obtained a list of medicinal plant species and services associated with individual indigenous groups from three regions: 1) North America: from the Native American Ethnobotany database (11)—the largest repository of indigenous knowledge for the region; 2) northwest Amazonia:

from Richard E. Schultes's book on the medicinal plants of northwest Amazonia, which integrates nearly half a century of his field research (12); and 3) New Guinea: from an ethnobotanical review of 488 references and 854 herbarium specimens (13).

We classified uses from the three data sources into medicinal subcategories following the classification in the *Economic Botany Data Collection Standard* (31), with modifications explained by Cámara-Leret et al. (32). Medicinal subcategories included Blood and cardio-vascular system; Cultural diseases and disorders; Dental health; Digestive system; Endocrine system; General ailments with unspecified symptoms; Infections and infestations; Metabolic system and nutrition; Muscular-skeletal system; Nervous system and mental health; Poisoning; Pregnancy, birth and puerperium; Reproductive system and reproductive health; Respiratory system; Sensory system; Skin and subcutaneous tissue; Urinary system; Veterinary; Not specified; and Other medicinal uses. We defined "unique knowledge" as a medicinal service cited exclusively by one indigenous language.

The amount of unique knowledge may change as more indigenous groups are studied and as more in-depth studies are made on indigenous groups already reported in the literature. We hypothesize that it will increase, for three reasons. First, research in South America indicates that the amount of unique knowledge is positively correlated with the total number of medicinal uses registered in a community (33). This is confirmed in New Guinea by a study that showed that linguistic uniqueness occurs even among the best-studied indigenous groups who occupy montane habitats with a similar flora (34). Therefore, any undersampling in our study regions—especially in New Guinea—would, in fact, underestimate unique knowledge because generalist knowledge tends to be the first that is documented in the field. Second, our classification of medicinal plant services is conservative because it omits "plant parts" (e.g., bark, leaf, fruit, and seed). Because different plant parts may be used for different purposes [different plant parts may have different chemical compounds (35)], our classification underestimates the detection of medicinal knowledge that is restricted to one language. Third, our classification of medicinal plant services is based on 20 medicinal subcategories that are broad in scope. For example, "Infections and infestations" encompasses reports related to, e.g., malaria, leishmaniasis, and measles. This means that if one indigenous language cites plant A as a remedy for malaria and another indigenous language cites plant A for measles, both reports would be classified as "Plant A: Infections and Infestations" and considered shared, rather than unique, knowledge.

Language Phylogenies and Threat. Medicinal services in the literature were associated with 119 indigenous languages in North America, 37 languages in northwest Amazonia, and 80 languages in New Guinea. For each region, we built language trees through phylogenetic inference using machine-learning techniques on the word lists of the Automated Similarity Judgment Program (ASJP) database (36) and used the Glottolog classification as a constraint tree (30). The ASJP list consists of the 40 most stable items, as determined by Holman et al. (37), from the 100-item list of Swadesh (38). To assess the degree of threat faced by languages in our sample, we queried the Glottolog (15), which derives an Agglomerated Endangerment Status (AES) from the databases of The Catalog of Endangered Languages, United Nations Educational, Scientific, and Cultural Organization Atlas of the World's Languages in Danger, and Ethnologue. For a list of the languages analyzed, see *SI Appendix*, Table S2.

Vascular Plant Phylogenies and Threat. We verified plant-species taxonomy using recently published checklists to the vascular plants of the Americas (39) and New Guinea (40). Using the list of medicinal plant species in each region, we queried the mega-tree GBOTB.-extended of Smith and Brown (41) with the *phylo.maker* function of the R package V.PhyloMaker (42). The phylogenies used in all subsequent analyses comprised 2,475 species in North America, 645 species in northwest Amazonia, and 477 species in New Guinea. To assess the threat faced by medicinal plant species, we queried the conservation assessments published by the IUCN Red List of Threatened Species (17), which classifies species as Data Deficient, Least Concern, Near Threatened, Vulnerable, Endangered, Critically Endangered, Extinct in the Wild, and Extinct. Following IUCN, species assessed to be Near Threatened, Vulnerable, and Endangered were considered threatened. Because most plant species lack IUCN conservation assessments, we also obtained endangerment probabilities from a recent study that used machine-learning to predict the conservation status of 30,497 plant species (18).

1. T. W. McDade et al., Ethnobotanical knowledge is associated with indices of child health in the Bolivian Amazon. Proc. Natl. Acad. Sci. U.S.A. **104**, 6134–6139 (2007).
2. B. Berlin, Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies (Princeton University Press, Princeton, NJ, 2014).
3. G. F. Simons, M. P. Lewis, "The world's languages in crisis: A 20-year update" in Responses to Language Endangerment. E. Mihas, B. Perley, G. Rei-Doval, K. Wheatley, Eds. (Studies in Language Companion Series, John Benjamins, Amsterdam, Netherlands, 2013), vol. 142, pp. 3–19.
4. W. J. Sutherland, Parallel extinction risk and global distribution of languages and species. Nature **423**, 276–279 (2003).
5. S. Aswani, A. Lemahieu, W. H. H. Sauer, Global trends of local ecological knowledge and future implications. PloS One **13**, e0195440 (2018).
6. A. Saynes-Vásquez, J. Caballero, J. A. Meave, F. Chiang, Cultural change and loss of ethnoecological knowledge among the Isthmus Zapotecs of Mexico. J. Ethnobiol. Ethnomed. **9**, 40 (2013).
7. R. Cámara-Leret et al., Climate change threatens New Guinea's biocultural heritage. Sci. Adv. **5**, eaaz1455 (2019).
8. S. Pironon et al., Potential adaptive strategies for 29 sub-Saharan crops under future climate change. Nat. Clim. Change **9**, 758–763 (2019).
9. R. Cámara-Leret, M. A. Fortuna, J. Bascompte, Indigenous knowledge networks in the face of global change. Proc. Natl. Acad. Sci. U.S.A. **116**, 9913–9918 (2019).
10. D. Nettle et al., Vanishing Voices: The Extinction of the World's Languages (Oxford University Press on Demand, Oxford, UK, 2000).
11. D. Moerman, Native American ethnobotany: A database of foods, drugs, dyes and fibers of Native American peoples, derived from plants (2020). naeb.brit.org/. Accessed 24 May 2020.
12. R. E. Schultes, R. F. Raffauf, The Healing Forest: Medicinal and Toxic Plants of the Northwest Amazonia (Dioscorides Press, Portland, OR, 1990).
13. R. Cámara-Leret, Z. Dennehy, Indigenous knowledge of New Guinea's useful plants: A review. Econ. Bot. **75**, 405–415 (2019).
14. A. Kik et al., The world's hotspot of linguistic and biocultural diversity under threat. bioRxiv [Preprint] (2021). https://doi.org/10.1101/2021.04.12.439439 (Accessed 26 April 2021).
15. H. Hammarström, R. Forkel, M. Haspelmath, S. Bank, Glottolog 4.3 (2021). https://glottolog.org/. Accessed 1 February 2020.
16. Mark. Pagel, Inferring the historical patterns of biological evolution. Nature **401**, 877–884 (1999).
17. International Union for Conservation of Nature. The IUCN red list of threatened species (Version 2020-2, International Union for Conservation of Nature, Cambridge, UK, 2020).
18. T. A. Pelletier, B. C. Carstens, D. C. Tank, J. Sullivan, A. Espíndola, Predicting plant conservation priorities on a global scale. Proc. Natl. Acad. Sci. U.S.A. **115**, 13027–13032 (2018).
19. C. E. Shannon, A mathematical theory of communications. Bell Syst. Tech. J **27**, 379–423 (1948).
20. R. Margalef, Information theory in biology. Gen. Syst. **3**, 36–71 (1958).
21. R. Verpoorte, Pharmacognosy in the new millennium: Leadfinding and biotechnology. J. Pharm. Pharmacol. **52**, 253–262 (2000).
22. P. A. Cox, L. R. Sperry, M. Tuominen, L. Bohlin, Pharmacological activity of the Samoan ethnopharmacopoeia. Econ. Bot. **43**, 487–497 (1989).
23. M. J. Balick, "Ethnobotany and the identification of therapeutic agents from the rainforest" in Bioactive Compounds from Plants, D. J. Chadwick, J. Marsh, Eds. (Ciba Foundation Symposium, Wiley and Sons, Chichester, UK, 1990), vol. 154, pp. 22–39.
24. W. H. Lewis, M. P. Elvin-Lewis, Medicinal plants as sources of new therapeutics. Ann. Mo. Bot. Gard., **82**, 16–24 (1995).
25. L. Svetaz, et al., Value of the ethnomedical information for the discovery of plants with antifungal properties. A survey among seven Latin American countries. J. Ethnopharmacol. **127**, 137–158 (2010).
26. T. A. K. Prescott, M. Briggs, R. Kiapranis, M. S. J. Simmonds, Medicinal plants of Papua New Guinea's Miu speaking population and a focus on their use of plant–slaked lime mixtures. J. Ethnopharmacol. **174**, 217–223 (2015).
27. T. A. K. Prescott et al., Tropical ulcer plant treatments used by Papua New Guinea's Apsokok nomads. J. Ethnopharmacol. **205**, 240–245 (2017).
28. B. Telban, The role of medical ethnobotany in ethnomedicine: A New Guinea example. J. Ethnobiol. **8**, 149–169 (1988).
29. W. L. Applequist et al., Antimalarial use of Malagasy plants is poorly correlated with performance in antimalarial bioassays. Econ. Bot. **71**, 75–82 (2017).
30. G. Jäger, Global-scale phylogenetic linguistic inference from lexical resources. Sci. Data **5**, 1–16 (2018).
31. F. E. M. Cook et al., Economic Botany Data Collection Standard (Royal Botanic Gardens, Kew, UK, 1995).
32. R. Cámara-Leret, N. Paniagua-Zambrana, H. Balslev, M. J. Macía, Ethnobotanical knowledge is vastly under-documented in northwestern South America. PloS One **9**, e85794 (2014).
33. R. Cámara-Leret, N. Paniagua-Zambrana, J.-C. Svenning, H. Balslev, M. J. Macía, Geospatial patterns in traditional knowledge serve in assessing intellectual property rights and benefit-sharing in northwest South America. J. Ethnopharmacol. **158**, 58–65 (2014).
34. R. Cámara-Leret, Z. Dennehy, Information gaps in indigenous and local knowledge for science-policy assessments. Nat. Sustain. **2**, 736–741 (2019).
35. R. F. Raffauf, Some notes on the distribution of alkaloids in the plant kingdom. Econ. Bot. **24**, 34–38 (1970).
36. S. Wichmann, E. W. Holman, C. H. Brown, Eds., The ASJP database (Version 17, 2016) https://zenodo.org/record/3835942. Accessed 14 April 2020.
37. E. W. Holman et al., Explorations in automated language classification. Folia Ling. **42**, 331–354 (2008).
38. M. Swadesh, Towards greater accuracy in lexicostatistic dating. Int. J. Am. Ling. **21**, 121–137 (1955).
39. C. U. Ulloa et al., An integrated assessment of the vascular plant species of the Americas. Science **358**, 1614–1617 (2017).
40. R. Cámara-Leret et al., New Guinea has the world's richest island flora. Nature **584**, 579–583 (2020).
41. S. A. Smith, J. W. Brown, Constructing a broadly inclusive seed plant phylogeny. Am. J. Bot. **105**, 302–314 (2018).
42. Y. Jin, H. Qian, V.Phylomaker: An R package that can generate very large phylogenies for vascular plants. Ecography **42**, 1353–1359 (2019).

ECOLOGY

SUSTAINABILITY SCIENCE