

Document downloaded from the institutional repository of the University of Alcalá: <http://dspace.uah.es/dspace/>

This is a postprint version of the following published document:

Molina-Venegas, R. et al., 2021. Maximum levels of global phylogenetic diversity efficiently capture plant services for humankind.. *Nature Ecology & Evolution*, 5, pp.583–588.

Available at <http://dx.doi.org/10.1038/s41559-021-01414-2>

© 2021 The Author(s), under exclusive licence to Springer Nature Limited

(Article begins on next page)



This work is licensed under a
Creative Commons Attribution-NonCommercial-NoDerivatives
4.0 International License.

1 **Maximum levels of global phylogenetic diversity efficiently capture plant services**
2 **for humankind**

3 Rafael Molina-Venegas^{1*}, Miguel Ángel Rodríguez¹, Manuel Pardo-de-Santayana^{2,3}, Cristina Ronquillo¹
4 & David J. Mabberley⁴

- 5 1. *GLoCEE - Global Change Ecology and Evolution Group, Department of Life Sciences,*
6 *University of Alcalá, 28805, Spain.*
7 2. *Department of Biology (Botany), Autonomous University of Madrid, 28049, Spain.*
8 3. *Research Centre in Biodiversity and Global Change (CIBC-UAM)*
9 4. *Wadham College, University of Oxford, OX1 3PN, United Kingdom; Department of Biology,*
10 *Macquarie University, and National Herbarium of New South Wales, Sydney, 2000, Australia.*
11

12 *Correspondence to: rafmolven@gmail.com
13

14 The divergent nature of evolution suggests that securing the human benefits that are directly
15 provided by biodiversity may require counting on disparate lineages of the Tree of Life. However,
16 quantitative evidence supporting this claim is still tenuous. Here, we draw on a global review of
17 plant-use records demonstrating that maximum levels of phylogenetic diversity capture
18 significantly greater numbers of plant-use records than random selection of taxa. Our study
19 establishes an empirical foundation that links evolutionary history to human well-being, and it
20 will serve as a discussion baseline to promote better-grounded accounts of the services that are
21 directly provided by biodiversity.
22
23

24 Evolution is the process that led to all living organisms and hence the foundation for the
25 human benefits that are directly provided by biodiversity¹, including not only basic resources
26 but also psychological and self-fulfilment needs². Since evolution is a divergent process, some
27 researchers have claimed that securing the services that are directly provided by biodiversity
28 may require counting on disparate lineages of the Tree of Life¹, because they might provide
29 complementary benefits³. Although this theoretical background is deeply rooted in the academic
30 literature⁴⁻⁶, empirical evidence connecting evolutionary history to human well-being is still
31 surprisingly tenuous⁷ and not without controversy^{8,9}. While some authors hold that maximizing
32 phylogenetic diversity should lead to recognition of high levels of useful feature diversity^{1,8},
33 others have suggested that the phylogenetic approach can be misleading⁹. This controversy
34 likely reflects that the connection between evolutionary history and human well-being remains
35 largely theoretical¹⁰ (but see Forest et al.¹¹ for an empirical local assessment), which is only an
36 initial move towards its consolidation as a scientific paradigm.

37 Here, we provide quantitative evidence that maximum levels of global plant
38 phylogenetic diversity (PD_{max}) capture more human benefits (i.e. plant-use records sorted into
39 28 standard categories of use¹²) and at higher diversity levels (i.e. records more evenly
40 distributed between the categories) than does random selection of taxa, supporting the long-
41 standing notion that maximizing phylogenetic diversity is a valuable means to retrieve high
42 levels of useful feature diversity⁴⁻⁶. Our genus-level analysis is based on the most
43 comprehensive time-calibrated vascular plant phylogeny available^{13,14}, including all accepted
44 vascular plant genera worldwide (a total of 13489) as well as 9478 genus-level plant-use records
45 (presence/absence) obtained from a systematic review of botanical literature and authoritative
46 websites¹⁵.

47 The PD_{max} strategy overcame random selection of taxa at any sample size (Fig. 1a), with
48 relative gains varying between 4% and 46% (Fig. 1b). This result suggests that in the absence of
49 any other source of information beyond evolutionary history, prospecting disparate lineages of
50 the phylogeny could help to make the most of the natural services that are the result of
51 evolution. With regard to individual plant-use categories, PD_{max} retrieved a higher number of

52 records relative to random selection in 92% of the comparisons (Fig. 2 and Supplementary Fig.
53 1). Moreover, given that relative record gains with PD_{max} were overall higher for the less
54 common categories (Supplementary Fig. 2), PD_{max} also retrieved significantly more equitable
55 distributions of records among categories at most sample sizes (Fig. 1c). This indicates that
56 PD_{max} recovers more plant-uses in general than random selection, and that it does so optimizing
57 the capture of some of the rarest uses, thus resulting in a more balanced palette of human
58 benefits. Both PD_{max} and random selection strategies retrieved the maximum possible richness
59 of plant-use categories (n = 28) across most sample sizes, yet random selection failed in
60 retrieving maximum richness of categories at 10% and 20% sample sizes in a few cases.

61 Our genus-level approach is superior to the species level in that the latter would suffer
62 from unacceptable omission errors –ethnobotanical knowledge will most likely remain vastly
63 under-documented for long below the genus level¹⁶⁻¹⁸– and extreme lack of phylogenetic
64 information¹³, yet it may introduce some uncertainty because the operational unit of plant-use is
65 often the species. As such, retrieving a useful genus that comprises just a few species could be
66 considered more valuable than a highly diversified one with the same use, because the
67 uncertainty regarding the species that are actually useful within each genus would be less in the
68 former case. Nonetheless, a reanalysis of the data after downweighting our genus-level plant-
69 use observations in direct proportion to species richness per genus revealed an even stronger
70 pattern (Supplementary Fig. 3). Moreover, the relationship between PD and plant benefits held
71 in separate continental regions of the world (TDWG level-1 standards, Supplementary Figs. 4-
72 5), which suggests that our results are consistent across floras that have evolved in distinct
73 biogeographic regions and over different timescales.

74 The striking success of the PD_{max} strategy lies in the phylogenetic structure of the
75 categories. As such, we found a strong positive relationship between the PD that is encapsulated
76 by each plant-use category and the relative gain in records per category under the PD_{max} strategy
77 (Extended Data Fig. 1), meaning that greater gains are predicted for phylogenetically dispersed
78 categories. In fact, the only category that was significantly underrepresented with PD_{max} relative

79 to random selection concerns rubber plants (Fig. 2 and Supplementary Fig. 1), which are
80 strongly clumped in the phylogeny (Supplementary Table 2). Our results complement previous
81 findings reported in local studies that high levels of PD can increase multifunctionality via
82 complementarity of beneficial attributes among phylogenetically distant taxa³. For example,
83 regarding the production of natural poisons against harmful or nuisance invertebrates, we found
84 that maximum levels of global PD capture more plant taxa generating them than random
85 selection (Fig. 2), which in turn may imply an increased potential to control the detrimental
86 effects of disparate invertebrate lineages. While the latter hypothesis cannot be tested with our
87 data, observations that most of the antagonistic plant-invertebrate interactions that ultimately
88 shaped this benefit are phylogenetically conserved^{19,20} (i.e. invertebrate species often attack a
89 narrow range of closely-related host plants) and geographically restricted²¹ support this idea. It
90 follows that, in the shadow of global change, counting on a variety of invertebrate poisons and
91 deterrents from distinct plant lineages may help to counter phylogenetically diverse pests
92 coming from disparate parts of the world^{22,23}.

93 It is important to note that an unobserved link between a human need and a taxon does
94 not necessarily imply that the link will not be found in the future. The ecological apparency
95 hypothesis states that among equally valuable taxa with regard to a certain use, the most
96 apparent or salient ones are preferred simply because they are readily available²⁴. Furthermore,
97 cultural factors could also explain the preferential use of certain taxa at the expense of others
98 that might equally fulfill the need²⁵. By analogy to the ecological prediction that higher
99 competition between closely related taxa of similar phenotypes can lead to greater phylogenetic
100 diversity²⁶, human preference patterns in the use of available plant resources might have
101 increased phylogenetic overdispersion in local ethnofloras. Therefore, ecological and cultural
102 factors, together with the fact that both plant lineages and the human cultures that prospect them
103 are geographically restricted to a greater or lesser extent, may have contributed to the striking
104 success of the PD_{max} strategy over random selection in capturing the human benefits that are
105 associated with plant biodiversity.

107 Services (IPBES) has recently approved undertaking the assessment of the use of ‘wild’ species,
108 including the identification of opportunities to establish measures that ensure and promote
109 sustainable practices²⁷. The ultimate goal of this conservation initiative is securing the “option
110 values” of biodiversity, this is, the present and future benefits that are associated with the
111 continued existence of a wide variety of taxa in nature, and phylogenetic diversity is
112 increasingly recognized as a valuable indicator of such maintenance of options²⁸. Concurring
113 with the IPBES philosophy that the world is in need of a broadly appreciation of option values
114 as a key contribution of nature to people^{29,30}, our study establishes a solid empirical foundation
115 that links evolutionary history to human well-being, and it will serve as a discussion baseline to
116 promote better-grounded accounts of the services that are directly provided by biodiversity^{31,32}.

117 118 **Methods**

119 **Plant-use dataset.** We compiled a genus-level dataset of plant-use records for all vascular plant
120 taxa described to date using the information gathered in the fourth edition of *Mabberley’s plant-*
121 *book*¹⁵. *Mabberley’s plant-book* is the most comprehensive and authoritative encyclopaedic
122 review of global plant classification (genera) and their uses published hitherto. From 1974 to
123 2017 all the information included in *Mabberley’s plant-book* was gathered, sorted, evaluated
124 and synthesized by David Mabberley, who systematically reviewed over 1000 botanical sources
125 including modern Floras, handbooks, periodicals, monographs and websites (all references can
126 be found in Mabberley¹⁵. We conducted a double-check manual screening of all plant-uses
127 described in *Mabberley’s plant-book* and sorted them into 28 standard categories of use
128 following the guidelines in the Economic Botany Data Collection Standard¹² (hereafter
129 “Collection Standard”). When two or more applications of the same category were described for
130 a given taxon, we considered them as a single plant-use record. For example, if the wood of a
131 taxon is used to build poles, furniture and toys (i.e. three different applications), we simply
132 recorded that the taxon provides timber. This procedure resulted in a binary classification of
133 9478 plant-use records across the 28 categories, including benefits related to human and animal

134 nutrition (human food, human-food additives, vertebrate food, invertebrate food), materials
135 (wood, stems, fibres, leaves, seeds/fruits, tannins/dyestuffs, gums/resins, lipids, waxes, scents,
136 latex/rubber), fuels (fuelwood, charcoal, biofuels), medicine (both human and veterinary),
137 poisons (vertebrate poison, invertebrate poison), social (antifertility agents, smoking
138 materials/drugs, symbolic/magic/inspiration) and environmental uses (ornamental,
139 bioindicators/bioremediators, soil improvers, hedging/shelter). A detailed description of the
140 categories is provided in Supplementary Table 1. Despite the use of leaves and seeds/fruits as
141 materials are considered as “miscellaneous” in the Collection Standard, we took them up front
142 as independent categories because we found many records in *Mabberley’s plant-book* that fit
143 into these categories (typically leaves for thatching and seeds/fruits for handicrafts). The
144 environmental categories “erosion control”, “revegetators”, “soil improvers” and “agroforestry”
145 described in the Collection Standard were considered as one single category (i.e. soil improvers)
146 because they were very difficult to tease apart in many cases (e.g. some plants are used in
147 agroforestry because they prevent soil erosion, and revegetators often improve soil quality). The
148 same rationale applies to the Collection Standard categories “shade/shelter” and
149 “boundaries/barriers/supports”, which were merged into one single category (i.e. hedges and
150 shelters). The Collection Standard also recognized different sub-categories of medicine, human
151 food and poisons¹², but we did not distinguish between them here because such information is
152 often unknown and does not make much sense in the context of our global assessment. For
153 example, while we are interested in recording the value of a taxon as human food,
154 distinguishing between the parts of the plant that are actually eaten (sub-categories for human
155 food in the Collection Standard) is rather irrelevant for the purposes of the study. A few records
156 could not be assigned to any of the categories described in the Collection Standard (e.g. spores
157 and inflorescences used as materials), which recommends gathering such cases into
158 “miscellaneous” categories¹². However, we simply disregarded them because such a mixture of
159 poorly represented categories would not make sense in the context of our study. Finally, the
160 category “cork and cork substitutes” described in the Collection Standard was disregarded
161 because we found very few records in *Mabberley’s plant-book* (likely because cork and cork

162 substitutes are provided only by a few species and primarily from *Quercus*). We considered
163 both fully realized (> 99% of the cases) and mooted uses (as long as they were properly
164 documented in the literature), and doubtful entries were disregarded in any case. The resultant
165 plant-use binary matrix (i.e. presence/absence of uses per genus) was used in all the analyses
166 described below. Additionally, we derived a downweighted plant-use matrix by dividing the
167 entries in the binary matrix (plant-use observations at the genus-level) by the total number of
168 accepted species per genus (following *Plants of the World Online*³³). This second matrix was
169 used in a second round of analyses to take into account the uncertainty in the relationship
170 between plant-use records in the genus-level dataset and the species that are actually useful, as
171 the latter information is often unknown.

172 Of all the taxa included in the dataset, 33% showed at least one category of use, with a
173 maximum number of plant-use records per taxa of 17 (Supplementary Fig. 6). The most
174 common category was “ornamental” (26%), followed by “medicine” (16%), “human food”
175 (13%) and “timber” (8%), while the rest of categories occurred at a frequency lower than 5%
176 (Supplementary Fig. 7). The phi correlation coefficient among the categories varied between -
177 0.008 and 0.332, suggesting overall weak relationships among them.

178

179 **Phylogenetic data.** We generated a genus-level time-calibrated molecular phylogeny using the
180 mega-tree GBOTB.extended¹⁴, which is a combination of the GBOTB tree for seed plants of
181 Smith & Brown¹³ and the pteridophytes clade in Zanne et al.³⁴ phylogeny with updates and
182 corrections (i.e. taxonomic standardization to *The Plant List*³⁵ nomenclatural and spelling
183 criteria). This combined phylogeny represents the most comprehensive and sophisticated
184 molecular phylogeny for vascular plants published hitherto. For each accepted genus in
185 *Mabberley’s plant-book*, we picked one representative species at random from the largest
186 monophyletic cluster of the genus in GBOTB.extended (if available). In the very few cases
187 where more than one largest monophyletic cluster was found, we first selected one of the
188 clusters at random and then picked one representative species. The GBOTB.extended phylogeny
189 was then pruned to retain only the representative species of the genera. After resolving a few

190 discrepancies and synonymy issues between *Mabberley's plant-book*¹⁵ and *The Plant List*³⁵
191 (using the nomenclatural criteria in *Plants of the World Online*³³ as a complementary reference
192 to solve disputes), we found that 71% of the genera accepted in *Mabberley's plant-book*
193 included at least one representative species in the phylogeny. This purely molecular
194 phylogenetic topology (hereafter “molecular tree”) revealed that all the taxonomic families of
195 the genera included in the tree formed monophyletic clades except for Nymphaeaceae,
196 Olacaceae, and Tectariaceae, which were paraphyletic, and the polyphyletic Diplaziopsidaceae
197 (see Supplementary Table 3 for a list of genera with taxonomic families). To take into account
198 uncertainty in the phylogenetic relationships of the taxa that were missed in the molecular tree
199 (hereafter “phylogenetically uncertain taxa” or PUT³⁶), we derived a distribution of
200 phylogenetic hypotheses from the latter using a systematic randomization procedure that was
201 taxonomically and phylogenetically informed. The workflow implies defining for each PUT its
202 “most derived consensus clade” (MDCC) (i.e. the clade in the molecular tree that most certainly
203 contains the PUT) based on expert knowledge³⁶ (e.g. taxonomy, morphology, geographic
204 distribution, etc). Once the MDCCs of the PUTs are defined, a distribution of phylogenetic
205 hypotheses can be generated by replicating the random insertion of the PUTs within their
206 respective MDCCs a high number of times (e.g. 100 times per posterior tree³⁶). The resultant
207 phylogenetic hypotheses can be then used to replicate the analyses and average the results over
208 the entire distribution of trees^{9,14,36}. Smith & Brown¹³ provided just one maximum likelihood
209 tree rather than a posterior distribution, and therefore we derived 100 alternative phylogenetic
210 hypotheses from the maximum likelihood tree as follows.

211 First, we retrieved for each genus in the dataset the taxonomic rank immediately above
212 in the taxonomic hierarchy (typically subtribe, tribe or subfamily in ascending order, hereafter
213 “taxonomic ranks”) from the NCBI Taxonomy database, the standard nomenclature and
214 classification repository for the International Nucleotide Sequence Database Collaboration³⁷.
215 For some families, this information was not available in the NCBI repository, in which case we
216 retrieved the taxonomic ranks from *Mabberley's plant-book*¹⁵. In the cases where taxonomic
217 ranks were neither available in the latter source, we simply assigned the family rank to the

218 genera. The mapping of taxonomic ranks in the molecular tree reveals whether or not they
219 represent natural lineages (i.e. monophyletic or paraphyletic³⁸), and we took advantage of such
220 information to define the MDCCs for our PUTs. If the taxonomic rank of a PUT mapped as
221 purely monophyletic or purely paraphyletic in the molecular tree, the subset of phylogenetic
222 branches connecting all the genera in the tree that shared the same taxonomic rank as the PUT
223 (hereafter “sharing taxa”) defined the MDCC (see Supplementary Figs. 8a-9a). In few cases, the
224 taxonomic ranks did not map as purely monophyletic or paraphyletic due to (1) the presence of
225 “outliers” that mapped away from the main cluster of sharing taxa or (2) the presence of
226 “intruders” from a different taxonomic rank within the main cluster. Such outliers and intruders
227 might represent incorrect taxonomic assignments or even artefacts derived from the
228 phylogenetic inference rather than evidence of unnatural (i.e. polyphyletic) groups. Thus, we
229 calculated two different indices for each potential monophyletic or paraphyletic cluster of
230 sharing taxa (because of the presence of outliers, intruders, or both) in the phylogeny. The
231 outlier ratio (OR) for a given set of sharing taxa is the ratio between the number of outliers
232 observed for the set (relative to the largest cluster) and the number of sharing taxa in the set, and
233 the intruder ratio (IR) is the ratio between the number of intruders observed within the largest
234 cluster of sharing taxa and the size of the cluster (see Supplementary Figs. 8-9). If (and only if)
235 both ratios were ≤ 0.05 , the subset of phylogenetic branches connecting all the sharing taxa in
236 the largest cluster (i.e. including intruders if any but not outliers) defined the MDCC of the
237 PUT. Otherwise, the MDCC was defined as the smallest phylogenetic clade that included all the
238 sharing taxa in the tree (i.e. including outliers and/or intruders, see Supplementary Figs. 8-9). In
239 those cases where one single genus represented the only sharing taxon of a PUT in the
240 molecular tree, the terminal node (i.e. the phylogenetic tip) defined the MDCC of the PUT only
241 if the node represented a singleton taxonomic family or subfamily. Otherwise (e.g. singleton
242 tribes or subtribes), the parent node of the singleton sharing taxon defined the MDCC instead
243 (see Supplementary Fig. 10). Once all the PUTs were assigned to a MDCC (see Supplementary
244 Table 4), they were added to a randomly selected branch of their corresponding MDCC, the
245 probability of being added along any branch of the clade being directly proportional to the

246 length of the branch. We used a uniform distribution to determine the exact position to insert the
247 PUTs along the selected branches³⁹. This procedure was replicated 100 times to obtain a
248 distribution of phylogenetic hypotheses.

249

250 **Finding the subsets of genera that maximize phylogenetic diversity.** We used the
251 phylogenetic diversity (PD) index as a metric of the evolutionary history encompassed by a set
252 of taxa⁴ because PD is the most commonly used metric in exercises that aim at maximizing
253 phylogenetic diversity^{4,8,9,40}. The greedy algorithm⁴¹ was used to find heuristically the subset of
254 genera in the phylogeny that maximized the PD metric (PD_{max}) for a sample size $S = 10, 20, 30,$
255 $40, 50, 60, 70, 80,$ and 90% of the total pool ($n = 13489$), respectively. Because there are
256 multiple subsets of size S that maximize PD in a phylogeny, we produced 10 PD_{max} subsets of
257 genera per alternative phylogenetic hypothesis ($n = 100$) and sample size S . Thus, we obtained
258 1000 different PD_{max} subsets for each sample size S^9 .

259

260 **Assessing the performance of the PD_{max} strategy.** With regard to human benefits provided by
261 plant biodiversity, the PD_{max} strategy could be considered more efficient than random selection
262 of taxa if the former captures (i) a greater richness of plant-use categories, (ii) a greater number
263 of plant-use records (in total and per category), and (iii) a greater equitability in the distribution
264 of the records among the categories (Pielou's evenness index⁴²). Thus, for each sample size S ,
265 we computed these variables using 1000 PD_{max} subsets and averaged the results to obtain one
266 observed value per sample size and variable⁹. We used standardized effect sizes (SES) to
267 compare observed values against null distributions generated by randomly picking subsets of S
268 taxa 1000 times:

$$269 \quad \text{SES} = \frac{M_{\text{obs}} - M_{\text{null}}}{SD_{\text{null}}} \quad (1)$$

270 where SES is the standardized effect size score for a given variable and sample size,
271 M_{obs} is the observed averaged value of the variable when taxa selection is phylogenetically
272 informed (i.e. using PD_{max} subsets), M_{null} is the mean of the null distribution (averaged value of

273 the variable when taxa are picked at random), and SD_{null} is the standard deviation of the null
274 distribution.

275

276 **Phylogenetic diversity of plant-use categories.** We computed the amount of evolutionary
277 history (PD) that is encapsulated in each plant-use category in our dataset⁴. PD is not
278 statistically independent of taxa richness, which differed greatly between the categories
279 (Supplementary Table 2). Therefore, in order to make PD values comparable between them, we
280 computed SES scores using equation 1. Null distributions of PD were generated for each
281 category by shuffling taxa labels across the phylogenetic tips 1000 times⁴³, and SES scores were
282 averaged across the 100 phylogenetic hypotheses used in the study. All analyses were
283 conducted in R⁴⁴ using the packages *picante*⁴⁵, *phytools*³⁹ and the greedyPD function developed
284 by Mazel et al.⁹.

285

286 **Continental-scale analyses.** In order to assess whether the relationship between PD and plant
287 benefits holds across floras that have evolved in distinct biogeographic regions, we also
288 conducted all the analyses described above at the continental scale. To do so, we compiled a
289 checklist of the native genera of each TDWG level-1 region (Biodiversity Information
290 Standards⁴⁶), namely, Africa (n = 4487), Australasia (n = 2067), Europe + Asia-Temperate (n =
291 4117), North America (n = 3307), Asia-Tropical (n = 4071) and South America (n = 4783),
292 using distributional information available in *Plants of the World Online*³³ and also *Mabberley's*
293 *plant-book*¹⁵ in the few cases where this information could not be retrieved from the former
294 source. The TDWG regions “Pacific” (minor Pacific islands) and “Antarctic” were disregarded
295 because they showed comparatively lower diversities, and “Europe” and “Asia-Temperate”
296 were merged into one single unit because the taxonomic turnover between the two regions (i.e.
297 β_{sim} distance⁴⁷) was very low (Supplementary Table 5), meaning that most of the genus-level
298 flora of “Europe” (the less diverse of the two) is shared with that of “Asia-Temperate”. Thus,
299 we finally analyzed six continental datasets separately. We note that widespread genera might
300 not always include useful species across their entire distribution range, which would lead to

301 overestimating the ethnofloras of the regions. Thus, in order to account for this uncertainty, we
302 also conducted the continental-scale analyses using only the genera that were endemic to each
303 region (Africa = 2294; Australasia = 776; Europe + Asia-Temperate = 1887; North America =
304 824; Asia-Tropical = 809; South America = 2387).

305

306 **Data availability**

307 The data that support the findings of this study are available at

308 <https://doi.org/10.6084/m9.figshare.13625546.v1>

309

310

311 **Code availability**

312 All the code used in this research is available as functions that were either implemented in

313 published R packages or provided as Supplementary Material in a previous Open Access study.

314

315 **References**

- 316 1. D. P. Faith, S. Magallón, A. P. Hendry, E. Conti, T. Yahara, M. J. Donoghue,
317 Evosystem services: an evolutionary perspective on the links between biodiversity and
318 human well-being. *Curr. Opin. Env. Sust.* **2**, 66–74 (2010).
- 319 2. R. Cámara-Leret, S. Faurby, M. J. Macía, H. Balslev, B. Göldel, J.-C. Svenning, W.
320 D. Kissling, N. Rønsted, C. H. Saslis-Lagoudakis, Fundamental species traits explain
321 provisioning services of tropical American palms. *Nat Plants.* **3**, 16220 (2017).
- 322 3. C. Oka, M. Aiba, T. Nakashizuka, Phylogenetic clustering in beneficial attributes of
323 tree species directly linked to provisioning, regulating and cultural ecosystem services.
324 *Ecol. Indic.* **96**, 477–495 (2019).
- 325 4. D. P. Faith, Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1–
326 10 (1992).
- 327 5. R. I. Vane-Wright, C. J. Humphries, P. H. Williams, What to protect?—Systematics
328 and the agony of choice. *Biol. Conserv.* **55**, 235–254 (1991).
- 329 6. R. H. Crozier, Genetic diversity and the agony of choice. *Biol. Conserv.* **61**, 11–15
330 (1992).

- 331 7. C. M. Tucker, T. Aze, M. W. Cadotte, J. L. Cantalapiedra, C. Chisholm, S. Díaz, R.
332 Grenyer, D. Huang, F. Mazel, W. D. Pearse, Assessing the utility of conserving
333 evolutionary history. *Biol. Rev.* (2019).
- 334 8. N. R. Owen, R. Gumbs, C. L. Gray, D. P. Faith, Global conservation of phylogenetic
335 diversity captures more than just functional diversity. *Nat. Commun.* **10**, 859 (2019).
- 336 9. F. Mazel, M. W. Pennell, M. W. Cadotte, S. Díaz, G. V. Dalla Riva, R. Grenyer, F.
337 Leprieur, A. O. Mooers, D. Mouillot, C. M. Tucker, W. D. Pearse, Prioritizing
338 phylogenetic diversity captures functional diversity unreliably. *Nat. Commun.* **9**
339 (2018), doi:10.1038/s41467-018-05126-3.
- 340 10. F. Mazel, M. W. Pennell, M. W. Cadotte, S. Díaz, G. V. D. Riva, R. Grenyer, F.
341 Leprieur, A. O. Mooers, D. Mouillot, C. M. Tucker, W. D. Pearse, Reply to: “Global
342 conservation of phylogenetic diversity captures more than just functional diversity.”
343 *Nat. Commun.* **10**, 858 (2019).
- 344 11. F. Forest, R. Grenyer, M. Rouget, T. J. Davies, R. M. Cowling, D. P. Faith, A.
345 Balmford, J. C. Manning, Ş. Procheş, M. van der Bank, others, Preserving the
346 evolutionary potential of floras in biodiversity hotspots. *Nature.* **445**, 757–760 (2007).
- 347 12. F. E. M. Cook, Economic Botany Data Collection Standard. International Working
348 Group on Taxonomic Databases for Plant Sciences (TDWG). Royal Botanic Gardens,
349 Kew, UK (1995).
- 350 13. S. A. Smith, J. W. Brown, Constructing a broadly inclusive seed plant phylogeny. *Am.*
351 *J. Bot.* **105**, 302–314.
- 352 14. Y. Jin, H. Qian, V. PhyloMaker: an R package that can generate very large
353 phylogenies for vascular plants. *Ecography.* **42**, 1353–1359 (2019).
- 354 15. D. J. Mabberley, Mabberley’s plant-book: a portable dictionary of plants, their
355 classification and uses. Fourth edition. Cambridge University Press (2017).
- 356 16. P. A. Cox, Will tribal knowledge survive the millennium? *Science.* **287**, 44–45 (2000).
- 357 17. R. Cámara-Leret, N. Paniagua-Zambrana, H. Balslev, M. J. Macía, Ethnobotanical
358 knowledge is vastly under-documented in northwestern South America. *PLOS ONE.*
359 **9**, e85794 (2014).
- 360 18. R. Cámara-Leret, Z. Dennehy, Information gaps in indigenous and local knowledge
361 for science-policy assessments. *Nat. Sustain.* **2**, 736–741 (2019).
- 362 19. V. Novotny, Y. Basset, S. E. Miller, G. D. Weiblen, B. Bremer, L. Cizek, P. Drozd,
363 Low host specificity of herbivorous insects in a tropical forest. *Nature.* **416**, 841–844
364 (2002).
- 365 20. G. S. Gilbert, R. Magarey, K. Suiter, C. O. Webb, Evolutionary tools for
366 phytosanitary risk analysis: phylogenetic signal as a predictor of host range of plant
367 pests and pathogens. *Evol. Appl.* **5**, 869–878 (2012).

- 368 21. J. Calatayud, J. L. Hórreo, J. Madrigal-González, A. Migeon, M. Á. Rodríguez, S.
369 Magalhães, J. Hortal, Geography and major host evolutionary transitions shape the
370 resource use of plant parasites. *PNAS*. **113**, 9840–9845 (2016).
- 371 22. G. T. Pecl, M. B. Araújo, J. D. Bell, J. Blanchard, T. C. Bonebrake, I.-C. Chen, T. D.
372 Clark, R. K. Colwell, F. Danielsen, B. Evengård, L. Falconi, S. Ferrier, S. Frusher, R.
373 A. Garcia, R. B. Griffis, A. J. Hobday, C. Janion-Scheepers, M. A. Jarzyna, S.
374 Jennings, J. Lenoir, H. I. Linnetved, V. Y. Martin, P. C. McCormack, J. McDonald, N.
375 J. Mitchell, T. Mustonen, J. M. Pandolfi, N. Pettorelli, E. Popova, S. A. Robinson, B.
376 R. Scheffers, J. D. Shaw, C. J. B. Sorte, J. M. Strugnell, J. M. Sunday, M.-N. Tuanmu,
377 A. Vergés, C. Villanueva, T. Wernberg, E. Wapstra, S. E. Williams, Biodiversity
378 redistribution under climate change: Impacts on ecosystems and human well-being.
379 *Science*. **355** (2017).
- 380 23. P. Lehmann, T. Ammunét, M. Barton, A. Battisti, S. D. Eigenbrode, J. U. Jepsen, G.
381 Kalinkat, S. Neuvonen, P. Niemelä, J. S. Terblanche, B. Økland, C. Björkman,
382 Complex responses of global insect pests to climate warming. *Front. Ecol. Environ*.
383 **18**, 141–150 (2020).
- 384 24. R. F. P. de Lucena, P. M. de Medeiros, E. de L. Araújo, A. G. C. Alves, U. P. de
385 Albuquerque, The ecological apparency hypothesis and the importance of useful
386 plants in rural communities from Northeastern Brazil: An assessment based on use
387 value. *J. Environ. Manag.* **96**, 106–115 (2012).
- 388 25. G. Menendez-Baceta, L. Aceituno-Mata, V. Reyes-García, J. Tardío, M. Salpeteur, M.
389 Pardo-de-Santayana, The importance of cultural factors in the distribution of
390 medicinal plant knowledge: A case study in four Basque regions. *J. Ethnopharmacol*.
391 **161**, 116–127 (2015).
- 392 26. C. O. Webb, D. D. Ackerly, M. A. McPeck, M. J. Donoghue, Phylogenies and
393 Community Ecology. *Annu. Rev. Ecol. Syst.* **33**, 475–505 (2002).
- 394 27. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
395 Global. Information on scoping for the thematic assessment of sustainable use of wild
396 species. IPBES (2018). Available at: [https://ipbes.net/sustainable-use-wild-species-](https://ipbes.net/sustainable-use-wild-species-assessment)
397 [assessment](https://ipbes.net/sustainable-use-wild-species-assessment)
- 398 28. IPBES, The IPBES regional assessment report on biodiversity and ecosystem services
399 for Asia and the Pacific. Karki, M., Senaratna Sellamuttu, S., Okayasu, S., and Suzuki,
400 W. (eds). Secretariat of the Intergovernmental Science-Policy Platform on
401 Biodiversity and Ecosystem Services, Bonn, Germany (2018).
- 402 29. M. Pardo-de-Santayana, M. Macía, The benefits of traditional knowledge. *Nature*.
403 **518**, 487–488 (2015).

- 404 30. S. Díaz, U. Pascual, M. Stenseke, B. Martín-López, R. T. Watson, Z. Molnár, R. Hill,
405 K. M. A. Chan, I. A. Baste, K. A. Brauman, S. Polasky, A. Church, M. Lonsdale, A.
406 Larigauderie, P. W. Leadley, A. P. E. van Oudenhoven, F. van der Plaats, M. Schröter,
407 S. Lavorel, Y. Aumeeruddy-Thomas, E. Bukvareva, K. Davies, S. Demissew, G.
408 Erpul, P. Failler, C. A. Guerra, C. L. Hewitt, H. Keune, S. Lindley, Y. Shirayama,
409 Assessing nature's contributions to people. *Science*. **359**, 270–272 (2018).
- 410 31. A. Antonelli, et al. (205 authors), State of the world's plants and fungi 2020. Royal
411 Botanic Gardens, Kew (2020).
- 412 32. T. Ulian, M. Diazgranados, S. Pironon, S. Padulosi, U. Liu, L. Davies, M.-J. R.
413 Howes, J. S. Borrell, I. Ondo, O. A. Pérez-Escobar, S. Sharrock, P. Ryan, D. Hunter,
414 M. A. Lee, C. Barstow, Ł. Łuczaj, A. Pieroni, R. Cámara-Leret, A. Noorani, C. Mba,
415 R. N. Womdim, H. Muminjanov, A. Antonelli, H. W. Pritchard, E. Mattana,
416 Unlocking plant resources to support food security and promote sustainable
417 agriculture. *Plants, People, Planet*. **2**, 421–445 (2020).
- 418 33. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published
419 on the Internet; <http://www.plantsoftheworldonline.org/>
- 420 34. A. E. Zanne, D. C. Tank, W. K. Cornwell, J. M. Eastman, S. A. Smith, R. G. FitzJohn,
421 D. J. McGlinn, B. C. O'Meara, A. T. Moles, P. B. Reich, D. L. Royer, D. E. Soltis, P.
422 F. Stevens, M. Westoby, I. J. Wright, L. Aarssen, R. I. Bertin, A. Calaminus, R.
423 Govaerts, F. Hemmings, M. R. Leishman, J. Oleksyn, P. S. Soltis, N. G. Swenson, L.
424 Warman, J. M. Beaulieu, Three keys to the radiation of angiosperms into freezing
425 environments. *Nature*. **506**, 89–92 (2014).
- 426 35. The Plant List. Version 1.1. Published on the Internet; <http://www.theplantlist.org/>
- 427 36. T. F. Rangel, R. K. Colwell, G. R. Graves, K. Fučíková, C. Rahbek, J. A. F. Diniz-
428 Filho, Phylogenetic uncertainty revisited: Implications for ecological analyses.
429 *Evolution*. **69**, 1301–1312 (2015).
- 430 37. S. Federhen, The NCBI Taxonomy database. *Nucleic Acids Res.* **40**, D136–D143
431 (2012).
- 432 38. E. Hörandl, T. F. Stuessy, Paraphyletic groups as natural units of biological
433 classification. *Taxon*. **59**, 1641–1653 (2010).
- 434 39. L. J. Revell, phytools: an R package for phylogenetic comparative biology (and other
435 things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
- 436 40. A. S. L. Rodrigues, K. J. Gaston, Maximising phylogenetic diversity in the selection
437 of networks of conservation areas. *Biol. Conserv.* **105**, 103–111 (2002).
- 438 41. M. Bordewich, A. G. Rodrigo, C. Semple, Selecting taxa to save or sequence:
439 desirable criteria and a greedy solution. *Syst. Biol.* **57**, 825–834 (2008).

- 440 42. E. C. Pielou, The measurement of diversity in different types of biological collections.
441 *J. Theor. Biol.* **13**, 131–144 (1966).
- 442 43. S. W. Kembel, Disentangling niche and neutral influences on community assembly:
443 assessing the performance of community phylogenetic structure tests. *Ecol. Lett.* **12**,
444 949–960 (2009).
- 445 44. R Core Team. R: A language and environment for statistical computing. (R foundation
446 for statistical computing Vienna, Austria, 2020).
- 447 45. S. W. Kembel, P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D.
448 Ackerly, S. P. Blomberg, C. O. Webb, Picante: R tools for integrating phylogenies and
449 ecology. *Bioinformatics.* **26**, 1463–1464 (2010).
- 450 46. Brummitt, R. K. (2001). World geographical scheme for recording plant distributions
451 (PDF) (2nd ed.). International working group on taxonomic databases for plant
452 sciences (TDWG).
- 453 47. A. Baselga, Partitioning the turnover and nestedness components of beta diversity.
454 *Glob. Ecol. Biogeogr.* **19**, 134–143 (2010).

455

456

457 **Acknowledgments**

458 We thank the Scientific Computation Center of Andalusia (CICA) for the computing services
459 they provided and Herlander Lima for his assistance in downloading plant distributional
460 information from the web. This work was supported by the Regional Government of the
461 Community of Madrid and the University of Alcalá through the project “Plant evolutionary
462 history and human well-being in a changing world; assessing theoretical foundations using
463 empirical evidence and new phylogenetic tools”, which was granted to RM-V (CM/JIN/2019-
464 005). RM-V was supported by the TALENTO program of the Regional Government of the
465 Community of Madrid (2018-T2/AMB-10332). MAR was supported by the Ministry of Science
466 and Innovation of Spain (grant CGL2017-86926-P).

467

468 **Author contributions**

469 RM-V conceived the ideas, led the assemblage of the plant-use dataset with the help of MPS
470 and DJM, conducted the analyses and led the writing. CR led the assemblage of the continental

471 datasets. MAR helped to design the structure of the draft. All the authors read, edited and
472 commented on the manuscript.

473

474 **Competing interests**

475 The authors declare no competing interests.

476

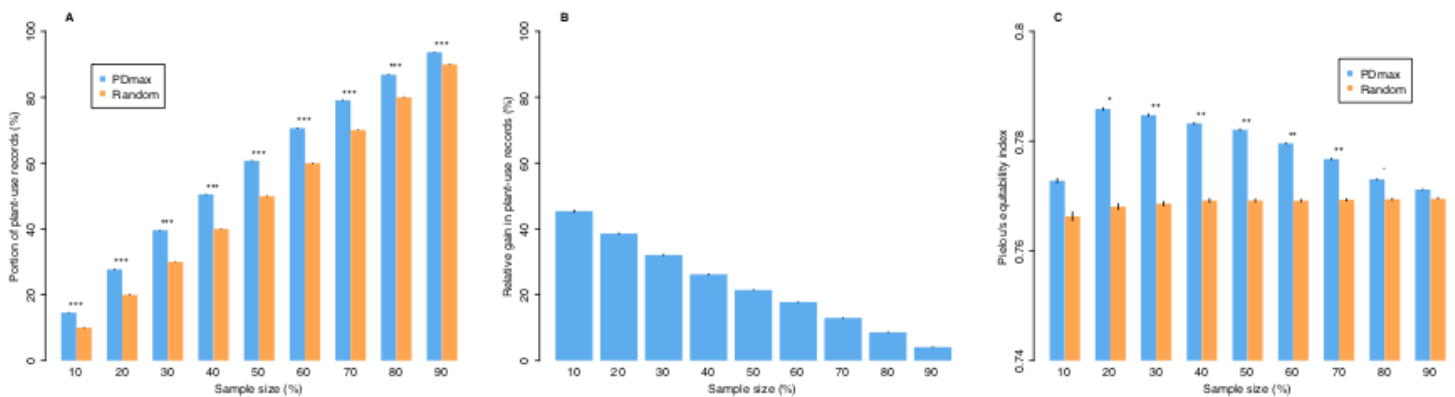
477 **Additional information**

478 Extended data is available for this paper at XXX

479 Supplementary information is available for this paper at XXX

480

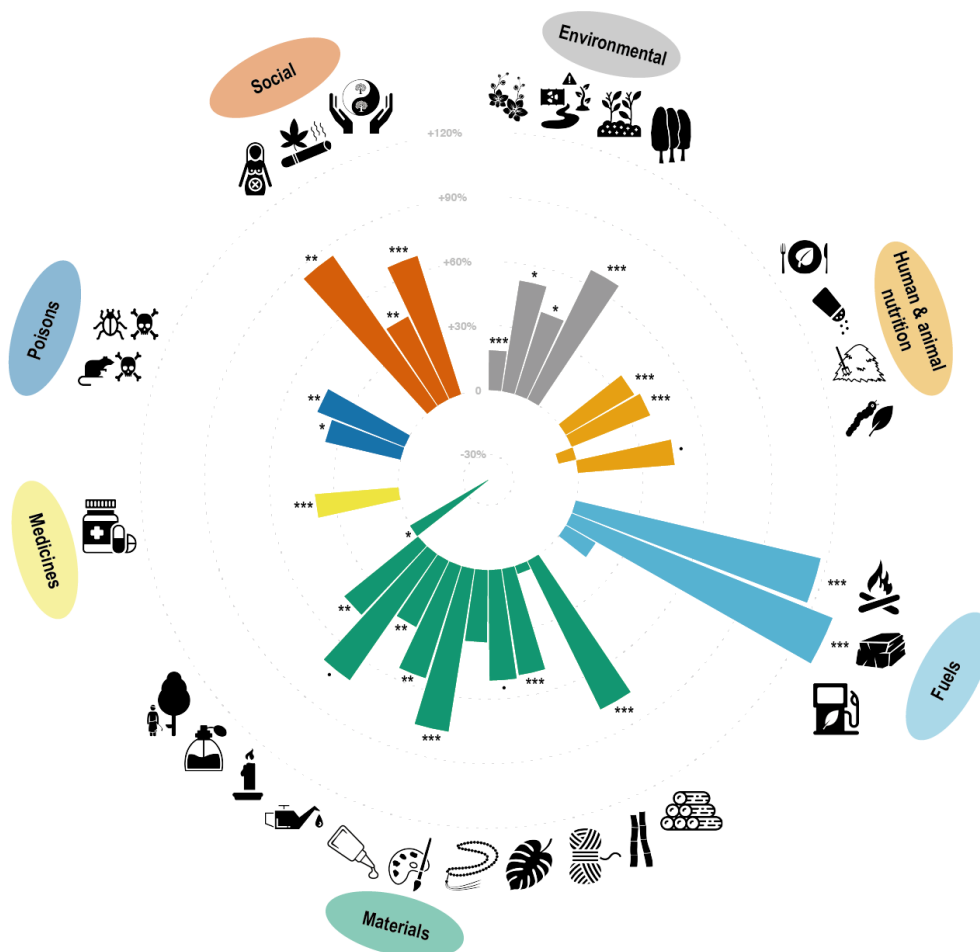
481 **Fig. 1. Relative gain in plant-use records and equitability in their distribution among**
 482 **categories.** a, Portion of the total pool of plant-use records ($n = 9478$ counted for all use categories
 483 combined) retrieved with the PD_{max} and random selection strategies across sample sizes. b, Gain
 484 in plant-use records obtained with PD_{max} relative to random selection across sample sizes. c,
 485 Equitability (Pielou's evenness index) in the distribution of plant-use records among the 28
 486 categories with PD_{max} and random sampling strategies across sample sizes. The symbols in a and
 487 c indicate statistical significance (based on SES scores) for a nominal alpha of 10% “.”, 5% “*”,
 488 1% “**” and 0.1% “***”, respectively (two-tailed tests), and the vertical thin bars at the center
 489 of the percentage bars represent confidence intervals at 95%.
 490
 491



493

494 **Fig. 2. Relative gains in plant-use records per category.** The bars represent the relative
 495 gains obtained with PD_{max} relative to random selection at $S = 20\%$ of the total pool of taxa, the
 496 sample size at which the maximum equitability in the distribution of records among use categories
 497 was observed (see Fig. 1c). The symbols on the bars indicate statistical significance (based on
 498 SES scores) for a nominal alpha of 10% “.”, 5% “*”, 1% “***” and 0.1% “****”, respectively (two-
 499 tailed tests). The colours represent different groups of categories following the Economic Botany
 500 Data Collection Standard (see Supplementary Table 1). Note that in order to optimize the
 501 visibility of the figure, values below the -30% threshold do not scale linearly (only one category
 502 with relative gain at -61%). From twelve o’clock and clockwise: (1) ornamental, (2) bioindicators
 503 and bioremediators, (3) soil improvers, (4) hedges and shelters, (5) human food, (6) human food
 504 additives, (7) vertebrate food, (8) invertebrate food, (9) fuelwood, (10) charcoal, (11) biofuels,
 505 (12) timber, (13) stems, (14) fibres, (15) leaves, (16) seeds and fruits, (17) tannins and dyestuffs,
 506 (18) resins and gums, (19) lipids, (20) waxes, (21) scents, (22) rubber, (23) medicines, (24)
 507 vertebrate poisons, (25) invertebrate poisons, (26) antifertility agents, (27) smoking materials and
 508 drugs, (28) symbolism, magic and inspiration.

509
 510



511