Global dataset shows geography and life form predict modern plant extinction and rediscovery

Aelys M. Humphreys [1,2]*, Rafaël Govaerts [3]*, Sarah Z. Ficinski[1], Eimear Nic Lughadha[4] and Maria S. Vorontsova [1]

Most people can name a mammal or bird that has become extinct in recent centuries, but few can name a recently extinct plant. We present a comprehensive, global analysis of modern extinction in plants. Almost 600 species have become extinct, at a higher rate than background extinction, but almost as many have been erroneously declared extinct and then been rediscovered. Reports of extinction on islands, in the tropics and of shrubs, trees or species with narrow ranges are least likely to be refuted by rediscovery. Plant extinctions endanger other organisms, ecosystems and human well-being, and must be understood for effective conservation planning.

Extinction of biodiversity is a central part of our planet's past, present and future. Current understanding of ongoing extinction comes primarily from projections or assessments of extinction risk[1–4]. Direct data on modern extinction (having occurred in recent centuries) are scarce but relatively well documented for birds and mammals[4]. These data have been used to assess the severity of ongoing species extinction[5], but extrapolations from vertebrates underestimate ongoing losses in invertebrates[6,7]. A general understanding of modern extinction clearly requires analysis of a broad sample of biodiversity. To date, however, no global analysis has included plants (but see refs. [8,9]). This is problematic if we are to make accurate predictions of future losses of plants, as well as of other organisms, because extinctions are not expected in one group of organisms independently of others (for example, co-extinction of insects and their host plants[6,10]).

Here we analyse a previously unpublished database of seed plants that have become extinct since Linnaeus' *Species Plantarum*[11]. The database is a complete list of species reported as extinct, based on continuous literature review complementing data from the International Union for Conservation of Nature (IUCN) Red List of Threatened Species[4] (hereafter, Red List) over almost three decades (see Methods and Supplementary Dataset 1). We used this database

as a starting point for increasing understanding of modern extinction in plants, by testing whether what is known about modern extinction in animals and extinction risk in plants is true also for documented modern extinction of plants. We expected plant extinctions to be: (1) more numerous than currently recorded by the Red List, the global authority on species extinction risk, but incomplete for most groups[2,12]; (2) elevated in rate above background extinction rates, because this has been shown for animals[13,14]; (3) disproportionately high on isolated, oceanic islands because of the vulnerability of island biotas to anthropogenic change[8,15,16]; (4) mostly of woody species and mostly from the wet tropics, reflecting extinction risk in plants[2,4]; and (5) phylogenetically clustered (that is, concentrated in certain evolutionary groups), as shown for modern extinction of mammals[17]. Finally, we compared the geographic and phylogenetic distribution of extinct species with species that have been erroneously declared extinct and subsequently rediscovered.

We show that the number of known seed plant extinctions is more than four times that on the Red List, and that the status of 50 species listed as extinct needs updating due mainly to rediscovery or taxonomic change (Supplementary Dataset 2 and Supplementary Information). In addition, we list 491 extinct species not on the Red List. In total, we document 571 known cases of modern extinction in plants (Supplementary Dataset 1), originating from 1,319 species once published as extinct and representing 1,234 currently accepted species (571 extinct, 431 rediscovered and 232 synonymized species). Fewer than 50% of reported species extinctions are therefore still presumed accurate.

Extinction of seed plants is occurring at a faster rate than the normal turnover of species. We found that, on average, 2.3 species have become extinct each year for the past 2.5 centuries. However, most species have not been known for 250 years, recently described species may have higher extinction rates than those described earlier[14] and species may become extinct before being formally

Table 1 | Rates of modern extinction in seed plants compared to vertebrates

	Total number of seed plant species described	Number of seed plant species extinct	Average seed plant taxonomic age (years)[a]	Seed plant extinction rate (E/MSY)[a]	Amphibian extinction rate (E/MSY)[b]	Bird extinction rate (E/MSY)[b]	Mammal extinction rate (E/MSY)[b]
Before 1900	129,529	256	171 (195)	11.6 (10.1)	66	49 (73[c])	72
1900–2018	204,793	315	60 (84)	25.6 (18.3)	107	132	243

Extinction rate is expressed as E/MSY. [a]Estimates without and, in brackets, with correction for the lag time between collection and description as a new species (24 years on average[18]). [b]Estimates from ref. [14]. [c]Estimate from ref. [13].

[1]Comparative Plant and Fungal Biology, Royal Botanic Gardens, Kew, Richmond, UK. [2]Department of Ecology, Environment and Plant Sciences, Stockholm University, Stockholm, Sweden. [3]Biodiversity Informatics and Spatial Analysis, Royal Botanic Gardens, Kew, Richmond, UK. [4]Conservation Science, Royal Botanic Gardens, Kew, Richmond, UK. *e-mail: aelys.humphreys@su.se; r.govaerts@kew.org
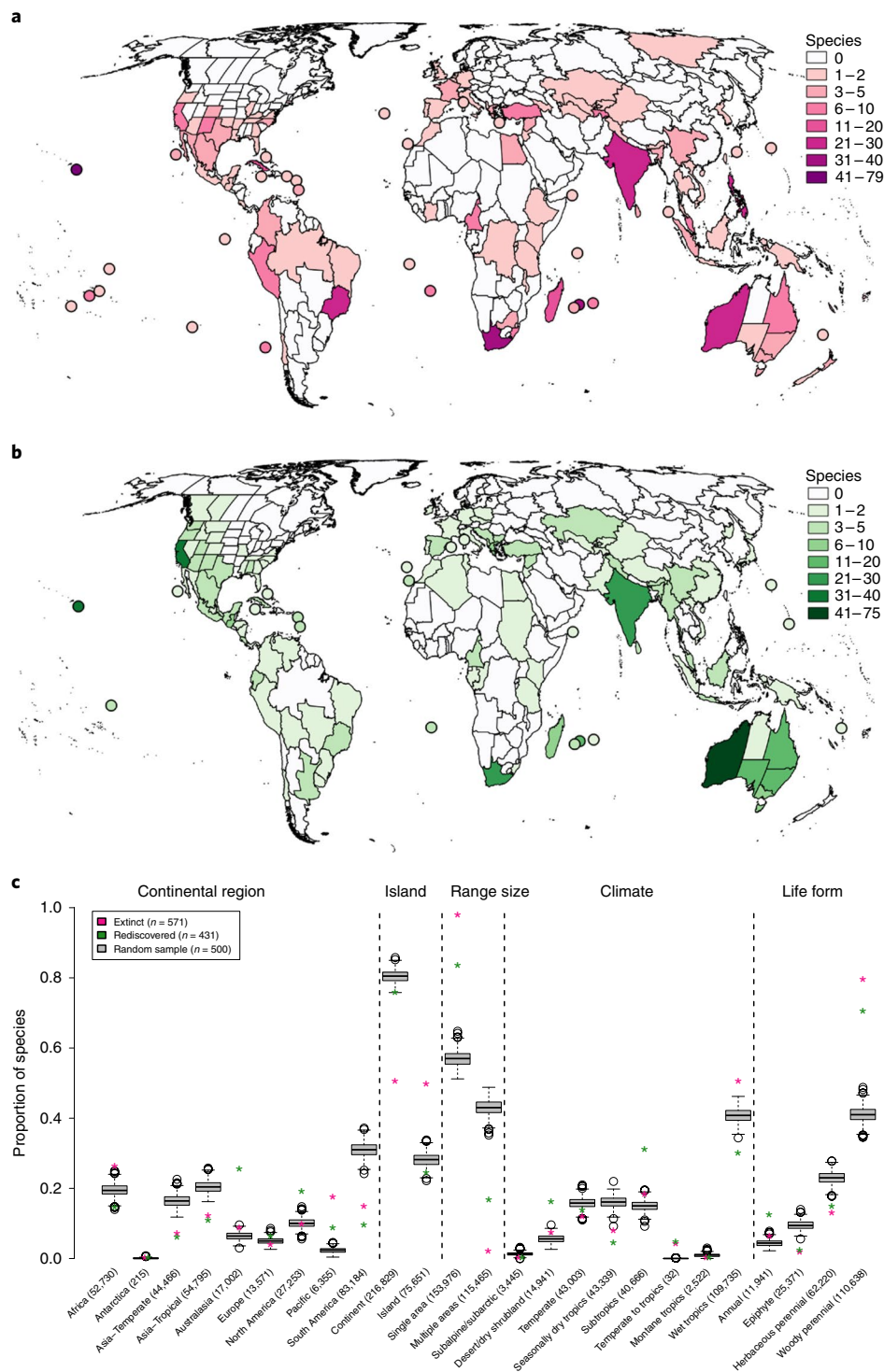
Fig. 1 | Distribution of extinct and rediscovered seed plant species among geographical regions, climate zones and life forms. a, The geographical pattern of modern extinction in seed plants. Hawaii stands out as having the most recorded extinctions (79), followed by the Cape Provinces of South Africa (37) and Mauritius (32), with Australia, Brazil, India and Madagascar also being among the top regions. Seventeen extinctions are of cultivated plants that have never been known in the wild (Supplementary Dataset 1). **b**, The geographical pattern of rediscovery of seed plants that have been erroneously declared extinct. Australia stands out for contributing the most rediscovered species overall (133). About half of the regions with the most recorded extinctions also have the most rediscovered species (Supplementary Information). **c**, Comparison of the proportions of extinct and rediscovered species on different continents, on islands versus continental regions, in a single geographical area versus multiple, in different climate zones and life forms, compared to the global distribution of seed plants. Boxplots show the range of proportions calculated across 500 random draws of 500 species from a database of 269,441 seed plant species (World Checklist of Selected Plant Families[23]). Upper and lower box bounds represent the 75th and 25th percentiles, respectively, and the horizontal line represents the median value and circles denote outliers. Numbers along the x axis represent the total number of seed plant species known in each category. Observed values are shown with asterisks (pink, extinct; green, rediscovered).

described. To correct for this, we calculated modern extinction rates separately for species described before and since the year 1900, including an average lag time of 24 years between the first collection of a plant and its description as a new species[18]. We found that recently described plants are becoming extinct at almost twice the rate of those described before 1900, but at rates almost an order of magnitude lower than for vertebrates (Table 1).

Our estimated rate of ongoing extinction (18–26 extinctions per million species years (E/MSY), Table 1) is up to 500 times the background extinction rate for plants (0.05–0.35 E/MSY[13]). However, previous studies have suggested current extinction rates closer to 1,000 times the normal turnover rate with others predicting that rates will soon be 10,000 times faster[5,13,14,16]. Similarly, our data suggest that 0.2% of standing plant diversity is extinct while mollusc losses have been estimated at 7% (ref. [7]) and birds and mammals at 5% (ref. [19]). Furthermore, 30,000 plant extinctions were projected by 2015 (ref. [20]) and 50,000 vascular plant species are estimated to be threatened with extinction[2]. Why are our findings lower than estimates for animals and forecasts for plants? The average extinction lag time is thought to be longer for plants than for animals, and thousands of living plant species are thought to be functionally extinct[20]. This is consistent with 89% of rediscovered species having high extinction risk (Supplementary Table 1), with several being known from only a few surviving individuals. Therefore, our estimated extinction rate, while elevated, is still likely to prove an underestimate of ongoing extinction of plant diversity.

The geographical pattern of modern extinction of plants is strikingly similar to that for animals[6,12,16,17]: all of the top extinction areas are high-diversity regions with a tropical or Mediterranean climate, including islands (Fig. 1a). The proportions of extinct species from islands (50%) and the Pacific (18%) are significantly higher than expected based on the global distribution of seed plants (Fig. 1c; $P < 0.01$ based on random draws, see Methods). This probably reflects the high proportion of unique species (endemics) in island biotas and their vulnerability to biological invasion[8]. Consistent with this, we found that extinct species have narrower ranges than seed plants as a whole (Fig. 1c; 98% of extinct species were known from a single region, compared to 57% of all seed plants). This confirms that biodiversity 'hotspots', with exceptional numbers of endemics and undergoing extensive habitat change[21], are key to understanding global patterns of recent and future extinctions[16].

Most extinct plants were woody perennials and/or from the wet tropics or subtropics. This reflects the diversity of seed plants globally, but the proportions for extinct plants are much higher than expected (Fig. 1c, $P < 0.01$; for example, 80% of extinct species are woody perennials compared to 40% of seed plants overall). In contrast, we found lower than expected proportions of epiphytes and herbaceous perennials, suggesting that herbaceous plants may be less prone to extinction due to larger population sizes and more persistent soil seed banks. Alternatively, these results indicate that our knowledge of modern extinction in plants is biased by historic and ongoing focus on trees[22], reflecting cultural, ecological and economic interests, a phenomenon analogous to zoologists' focus on birds and fish rather than on insects.

There is no phylogenetic pattern to plant extinction—that is, extinctions were randomly distributed among evolutionary groups. Extinct species are distributed among a quarter of seed plant families[23], with most extinctions reported from species-rich and globally widespread families (Fig. 2a, Supplementary Fig. 1 and Supplementary Information). The lack of phylogenetic signal to plant extinctions at the family level suggests that unrecorded modern extinctions and future extinctions are unlikely to be predicted from plants' evolutionary relationships. This is at odds with findings for modern extinction in mammals[17] and extinction risk in vertebrates[24] but consistent with most studies on extinction risk in plants[25], reinforcing that extinction processes for animals and plants are different.
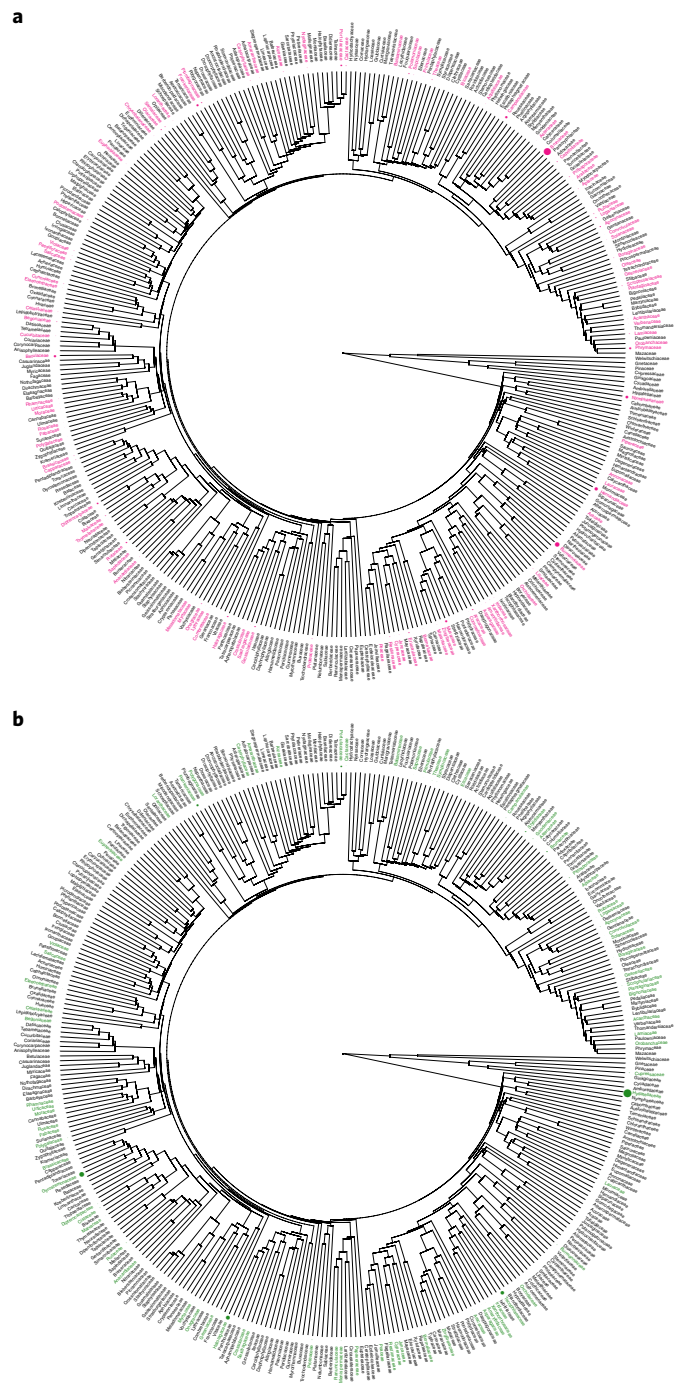

Fig. 2 | Phylogenetic distribution of seed plant families with extinct and rediscovered species. a, Extinction and **b**, rediscovery of species. There is no phylogenetic signal to what families contain species that **a**, have been declared extinct in recent centuries or **b**, have subsequently been rediscovered. This finding is robust to whether extinction is analysed as a proportion of species per family or the presence/absence of at least one extinct/rediscovered species. Families with at least one extinct/rediscovered species are shown in pink (extinct) or green (rediscovered), and proportions of extinct/rediscovered species are shown as filled circles. Note that the proportion of extinct/rediscovered species per family is vanishingly small in most cases (≤3.7% for extinct and ≤8.3% for rediscovered species; Supplementary Fig. 1), and circles have been scaled for visibility.

A final contribution of this study is the quantification of rediscovery of species once reported as extinct. This not only improves

the accuracy of our understanding of extinction but allows for potential remedial work as well. Almost as many species as have become extinct have been rediscovered (Supplementary Dataset 1). Overall, we report a rediscovery rate of 35%, or 16 species per year in the past three decades. This is consistent with 39% rediscovery over two decades for Australian plants[26] and 36% for mammals[27], but exceeds the 9% rate for molluscs[12]. The striking consistency between plants and mammals might represent a genuine error rate associated with extinction declarations. The lower rate for molluscs probably reflects a focus on island taxa, and the rediscovery rate for plants from oceanic islands is similarly low (9%; Fig. 1c).

Geographical and phylogenetic patterns of rediscovery mirror those for extinction (Figs. 1 and 2 and Supplementary Information), and there is overlap between families with many extinctions and those with many rediscoveries (Supplementary Fig. 1). Search effort is a correlate of rediscovery in mammals[27], and our demonstration of rediscovery in the very taxa and regions with reported extinctions reduces the risk that the presumed extinctions are erroneous. Despite this overall similarity, important differences between extinction and rediscovery suggest that rediscovery is most likely on continents (as opposed to islands), in the (dry) subtropics and of plants that are annual and/or have distributions that span more than one geographic area (Fig. 1c and Supplementary Information). Broad distributions have also been associated with rediscovery of mammals[27]. Higher than expected proportions of annuals among rediscovered plants might be because annuals are not always easily detectable and are thus more likely to be erroneously reported as extinct. In contrast, our data suggest that reports of extinction for trees, shrubs and species with narrow distributions from islands or the wet tropics are more likely to be accurate. Thus, our results reveal mechanisms by which plants may erroneously be declared extinct and biases in our knowledge of plant extinctions, which is skewed to well-studied areas of long-standing botanical interest and to woody rather than herbaceous plants.

In conclusion, our study greatly advances understanding of ongoing extinction in plants and suggests that geography and life form best predict ongoing extinction. The study of extinction inevitably comes with caveats. Extinctions of poorly known taxa may go unreported resulting in underestimation of rates; conversely, even for better-known taxa, low detectability may result in rate overestimation, revealed only by rediscovery. Our study indicates that these caveats can be alleviated through increased study in poorly known, biodiverse areas, and by furthering understanding of the environmental, socioeconomic and temporal relationships between extinction and rediscovery. We urge botanists to compile data on search effort, species density, abundance and detectability and to engage local people in the search for their missing biodiversity. Such efforts will improve our understanding of genuine extinctions and help target future conservation action.

Methods

The database on seed plant extinction was built using all available sources of information: global, regional and national Red Lists, taxonomic revisions, floras, research papers, field trips and herbarium visits (Supplementary Dataset 1). The database was originally compiled from global and regional Red Lists in the 1980s. Since then, all incoming literature at the Royal Botanic Gardens, Kew has been screened weekly for any new reports of plants that are globally extinct or extinct in the wild (as defined by the Red List[4]), or reports of rediscovery or taxonomic change that would render a previous extinction declaration invalid. New data have continuously been added to the database. Rediscovery was defined as the discovery of a single living population. As a consequence, the continuously maintained dataset analysed here is different to the Red List data, which are not continuously updated and not compiled following comprehensive literature searches to ascertain all reported extinctions.

The rate of species extinction was expressed as E/MSY, taking into account the average time elapsed since species were described, separately before and after the year 1900 (ref. [14]), including the average 24-year lag time between first collection

and description as a new species[18]. Dates (date of the earliest name belonging to each species concept) were obtained from the World Checklist for Selected Plant Families[23] (WCSP) for all currently accepted seed plant species ($n = 334{,}322$).

Phylogenetic signal in extinction and rediscovery was tested for at the family level using a published, dated phylogeny[28]. Two of the families in our dataset (Pandanaceae and Zamiaceae) were not present in the tree and were therefore excluded from analyses. We used Pagel's λ (ref. [29]) and Blomberg's K (ref. [30]) to calculate the phylogenetic signal of the proportion of extinct or rediscovered species per family, normalized by logit transformation. We also treated the data as a binary variable, scored as the presence or absence of extinct or rediscovered species for each family, and assessed the phylogenetic signal using the D-statistic[31] (Supplementary Information). Significance of K and D were determined using 1,000 randomizations each.

To test whether the distribution of extinct and rediscovered species among geographical regions, climate zones and life forms differs from that for seed plants as a whole, we used data on these variables from the WCSP[23]. The checklist is complete at the family level, and includes geographical data for 269,441 species and life form data for 210,170 species. The geographic distribution data follow a custom, standardized system of country and first-level province boundaries developed for plants and approved as a Taxonomic Databases Working Group (TDWG) standard[32]. Geographical data were manually scored as 'island' versus 'continent', and range sizes were defined as the number of country or first-level province (TDWG level 3) regions in which each species occurs. To provide sets of seed plants for comparison with our dataset, we generated 500 random draws of 500 species and used these to generate expected proportions of species among continents, climate zones, life forms and on islands. Observed data were compared to the distribution of proportions in these random draws.

Data availability

All new data on plant extinction and rediscovery analysed during this study are available as supplementary files linked to this published article. The data used for comparison with all seed plants are from the World Checklist of Selected Plant Families and are, or will soon become, publicly available from http://wcsp.science.kew.org.

References

1. Pimm, S. L. & Raven, P. *Nature* **403**, 843–845 (2000).
2. Brummitt, N. A. et al. *PLoS ONE* **10**, e0135152 (2015).
3. Pelletier, T. A., Carstens, B. C., Tank, D. C., Sullivan, J. & Espindola, A. *Proc. Natl Acad. Sci. USA* **115**, 13027–13032 (2018).
4. *The IUCN Red List of Threatened Species* Version 3.1 (IUCN, accessed June 2016).
5. Ceballos, G. et al. *Sci. Adv.* **1**, e1400253 (2015).
6. Dunn, R. R. *Conserv. Biol.* **19**, 1030–1036 (2005).
7. Regnier, C. et al. *Proc. Natl Acad. Sci. USA* **112**, 7761–7766 (2015).
8. Bellard, C., Rysman, J. F., Leroy, B., Claud, C. & Mace, G. M. *Nat. Ecol. Evol.* **1**, 1862 (2017).
9. Gray, A. The ecology of plant extinction: rates, traits and island comparisons. *Oryx*, 1–5 (2018).
10. Fonseca, C. R. *Conserv. Biol.* **23**, 1507–1515 (2009).
11. Linnaeus, C. *Species Plantarum* (Salvius, 1753).
12. Regnier, C., Fontaine, B. & Bouchet, P. *Conserv. Biol.* **23**, 1214–1221 (2009).
13. De Vos, J. M., Joppa, L. N., Gittleman, J. L., Stephens, P. R. & Pimm, S. L. *Conserv. Biol.* **29**, 452–462 (2015).
14. Pimm, S. L. et al. *Science* **344**, 987 (2014).
15. Bellard, C., Cassey, P. & Blackburn, T. M. *Biol. Lett.* **12**, 20150623 (2016).
16. Pimm, S. L., Russell, G. J., Gittleman, J. L. & Brooks, T. M. *Science* **269**, 347–350 (1995).
17. Turvey, S. T. & Fritz, S. A. *Philos. Trans. R. Soc. B* **366**, 2564–2576 (2011).
18. Bebber, D. P. et al. *Proc. Natl Acad. Sci. USA* **107**, 22169–22171 (2010).
19. Mooers, A. O., Goring, S. J., Turvey, S. T. & Kuhn, T. S. in *Holocene Extinctions* (ed. Turvey, S.T.) 263–277 (Oxford Univ. Press, 2009).
20. Cronk, Q. Plant extinctions take time. *Science* **353**, 446–447 (2016).
21. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. *Nature* **403**, 853–858 (2000).
22. Beech, E., Rivers, M., Oldfield, S. & Smith, P. P. *J. Sustain. For.* **36**, 454–489 (2017).
23. *World Checklist of Selected Plant Families* (WCSP, 2018); http://wcsp.science.kew.org/
24. Welch, J. N. & Beaulieu, J. M. *Diversity* **10**, 63 (2018).
25. Nic Lughadha, E. M. et al. in *State of the World's Plants 2017* (ed. Willis, K. J.) 72–77 (Royal Botanic Gardens, 2017).

26. Keith, D. A. & Burgman, M. A. *Biol. Conserv.* **117**, 41–48 (2004).
27. Fisher, D. O. & Blomberg, S. P. *Proc. R. Soc. B* **278**, 1090–1097 (2011).
28. Magallón, S., Gomez-Acevedo, S., Sanchez-Reyes, L. L. & Hernandez-Hernandez, T. *New Phytol.* **207**, 437–453 (2015).
29. Pagel, M. *Nature* **401**, 877–884 (1999).
30. Blomberg, S. P, Garland, T. & Ives, A. R. *Evolution* **57**, 717–745 (2003).
31. Fritz, S. A. & Purvis, A. *Conserv. Biol.* **24**, 1042–1051 (2010).
32. Brummitt, R. K. *World Geographical Scheme for Recording Plant Distributions. For International Working Group on Taxonomic Databases For Plant Sciences (TDWG)* 153 (Hunt Institute for Botanical Documentation, Carnegie Mellon University, 2001).

Author contributions

A.M.H., R.G. and M.S.V. designed the study, based on data collected by R.G.; A.M.H. analysed the data with contributions from S.Z.F., E.N.L. and M.S.V.; and A.M.H. wrote the paper with contributions from all authors. All authors approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-019-0906-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to A.M.H. or R.G.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s):	Aelys M. Humphreys: aelys.humphreys@su.se, Rafaël Govaerts: r.govaerts@ kew. org
Last updated by author(s):	Apr 11, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software used
Data analysis	All analyses were performed in R version 3.4.0 Packages caper, motmot and phytools were used, with details provided in the reference list

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All the data associated with this paper are made available as supplementary files (Supplementary Data 1 and 2). We also draw on data from the World Checklist for Selected Plant Families (Govaerts, R. 2018, Royal Botanic Gardens, Kew), most of which is publicly available. The remainder will soon be released and can in the meantime be obtained by request (r.govaerts@kew.org).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences   ☐ Behavioural & social sciences   ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a long-term bibliographic/literature study, with a small proportion of data points verified through visits to herbaria.
Research sample	All information available during a recurrent bibliographic study since 1989. The data represent all information available to us.
Sampling strategy	All data on plant extinctions available were analysed. No sample size calculations were performed.
Data collection	Weekly bibliographic scanning to monitor changes to published records of plant extinction, carried out by Rafaël Govaerts from 1989, ongoing.
Timing and spatial scale	Ongoing bibliographic scanning since 1989; global in scope for all seed plants.
Data exclusions	We did not consider records below species rank and we did not consider species thought to be of hybrid origin.
Reproducibility	No experiments were conducted.
Randomization	We analysed all data available to us; therefore no randomization was performed.
Blinding	No blinding was carried out.

Did the study involve field work?   ☐ Yes   ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging