New Guinea has the world's richest island flora

https://doi.org/10.1038/s41586-020-2549-5

Received: 2 October 2019

Accepted: 29 June 2020

Published online: 05 August 2020

Check for updates

A list of authors and their affiliations appears at the end of the paper.

New Guinea is the world's largest tropical island and has fascinated naturalists for centuries[1,2]. Home to some of the best-preserved ecosystems on the planet[3] and to intact ecological gradients—from mangroves to tropical alpine grasslands—that are unmatched in the Asia-Pacific region[4,5], it is a globally recognized centre of biological and cultural diversity[6,7]. So far, however, there has been no attempt to critically catalogue the entire vascular plant diversity of New Guinea. Here we present the first, to our knowledge, expert-verified checklist of the vascular plants of mainland New Guinea and surrounding islands. Our publicly available checklist includes 13,634 species (68% endemic), 1,742 genera and 264 families—suggesting that New Guinea is the most floristically diverse island in the world. Expert knowledge is essential for building checklists in the digital era: reliance on online taxonomic resources alone would have inflated species counts by 22%. Species discovery shows no sign of levelling off, and we discuss steps to accelerate botanical research in the 'Last Unknown'[8].

Great uncertainty remains as to the number of New Guinea plant species known to science, with conflicting estimates ranging from 9,000 to 25,000 species[9,10]. To narrow this range, here we catalogue the entire known vascular flora (angiosperms, gymnosperms, ferns and lycophytes) of mainland New Guinea and its surrounding islands (hereafter 'New Guinea'; Fig. 1a, Extended Data Fig. 1). We do so through a large-scale collaborative effort in which 99 plant experts verified the identity of 23,381 taxonomic names derived from 704,724 specimens (see Methods). Overall, we find that New Guinea supports 13,634 described species, 1,742 genera and 264 families of vascular plants (Supplementary Tables 1, 2). This suggests that New Guinea is the world's most floristically diverse island, with a known vascular plant flora 19% larger than the 11,488 species recorded in Madagascar[11] and 22% larger than the 11,165 species recorded in Borneo (http://www.plantsoftheworldonline.org, accessed 27 April 2019). New Guinea contains almost three times the 4,598 spermatophyte species of Java[12] and 1.4 times the 9,432 vascular plant species of the Philippines[13]—the only Malesian island regions for which Floras have been published. The vascular plant flora of New Guinea is divided between two political entities (Fig. 1a): Papua New Guinea, with 10,973 species, has 44% more species than Indonesian New Guinea (Papua Barat and Papua provinces), which has 7,616. Papua New Guinea also has more genera (1,654 versus 1,511) and families (260 versus 248). These differences partly arise from the lower collecting density in Indonesian New Guinea[1,2] (Fig. 1a). Nevertheless, the order of country rankings in plant diversity is unlikely to change with further collections because Papua New Guinea has a larger area, and surface area is the strongest predictor of island plant diversity[14]. Our species total for Papua New Guinea differs markedly from the 29,756 species that were presented in an unverified list of the Global Biodiversity Information Facility[15] and our total number of genera for New Guinea is 28% lower than the 2,437 unverified genera reported in a previous macroecological study[16]. Together, these differences underscore the need for expert validation in the digital era, which we discuss below.

Floristic patterns

Five species-rich families make up 35% of the flora of New Guinea: Orchidaceae (2,856 species), Rubiaceae (784), Ericaceae (438), Poaceae (376) and Myrtaceae (352) (Fig. 1b, Extended Data Table 1). Orchidaceae account for 20% and 17% of the flora of Papua New Guinea and Indonesian New Guinea, respectively. The floristic importance of orchids is comparable to that in other megadiverse countries such as Ecuador (23% of total flora) and Colombia (15%)[17]. The five largest genera of vascular plants in New Guinea are *Bulbophyllum* (658 species; Orchidaceae), *Dendrobium* (614 species; Orchidaceae), *Syzygium* (207 species; Myrtaceae), *Ficus* (179 species; Moraceae) and *Rhododendron* (171 species; Ericaceae) (Fig. 2, Extended Data Table 2). Of the 1,742 genera found, 13 have more than 100 species and make up 21% of all species, whereas 692 genera are represented by a single species in New Guinea.

Endemism

Plant endemism in New Guinea is remarkably high: it is the only Malesian island group with more endemic than non-endemic species (9,301 endemic species; 68% of the total). This preponderance of endemic species was noted in earlier studies, although these were based on smaller floristic samples[9,18]. The uniqueness of New Guinea within Malesia may be explained by its greater land surface area and habitat diversity[5]; its location, marking the junction between Malesia, Australia and the Pacific; and its highly complex tectonic history[19]. Geographically, 53% of the endemic species have been found only in Papua New Guinea and 24% occur only in Indonesian New Guinea. Of the total species from Papua New Guinea, 64% are endemic, and 58% of the total species from Indonesian New Guinea are endemic. Such high richness of endemic species means that both countries have a unique responsibility for the survival of this irreplaceable biodiversity. Given the general trend of plant endemism to increase with elevation[20], the conservation of ecosystems along altitudinal gradients is particularly critical.

Angiosperms have higher species endemism (71%) than ferns and lycophytes (46%) or gymnosperms (41%). Endemism within families is highly uneven, with just eight angiosperm families comprising 50% of all endemics: Orchidaceae (2,464 endemic species), Rubiaceae (669), Ericaceae (431), Arecaceae (257), Myrtaceae (255), Gesneriaceae (218), Apocynaceae (196) and Lauraceae (195) (Fig. 1c). The families with the highest proportions of endemism are Ericaceae (98% of species

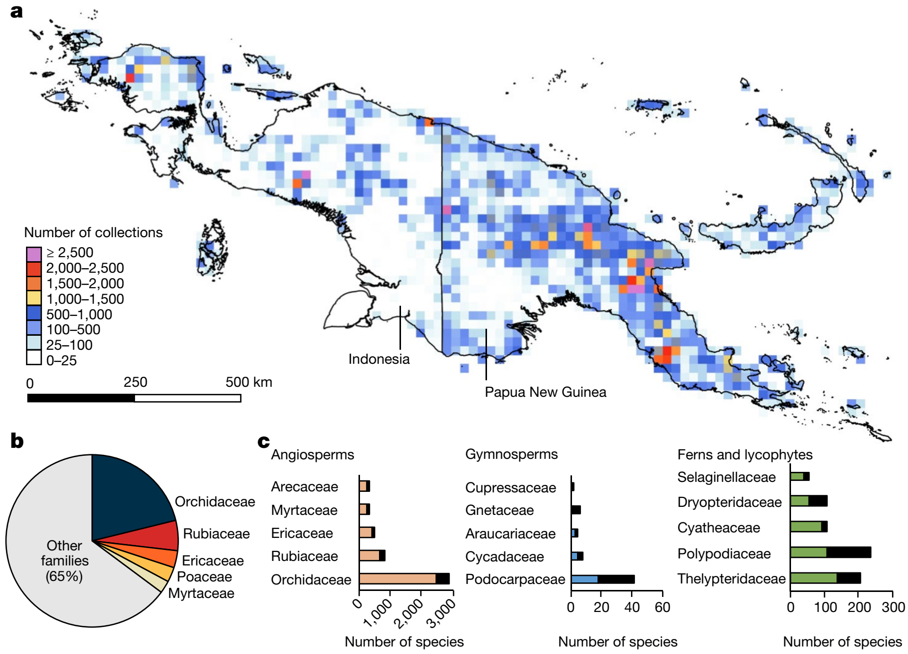


Fig. 1 | Floristic patterns in New Guinea. a, Map of the study area of mainland New Guinea and surrounding islands, showing the number of digitized collections per grid cell of 25 × 25 km. **b**, The five plant families that comprise 35% of the flora. **c**, Families with highest species endemism in angiosperms (orange), gymnosperms (blue) and ferns and lycophytes (green), arranged by increasing number of endemic species. Black bars depict the number of non-endemic species.

endemic), Gesneriaceae (96%) and Zingiberaceae (95%). All New Guinea species of *Vaccinium* (Ericaceae) are endemic and over 95% of species of *Begonia* (Begoniaceae), *Cyrtandra* (Gesneriaceae), *Glomera* (Orchidaceae), *Psychotria* (Rubiaceae), *Rhododendron* (Ericaceae), *Saurauia* (Actinidiaceae) and *Taeniophyllum* (Orchidaceae) are endemic. There are 61 endemic genera in New Guinea and these contain 164 species (ranging from 1–17 species per genus) or 2% of the endemic species (Extended Data Table 3). However, molecular research is urgently needed to test the monophyly of endemic genera, as phylogenetic data are absent for 59% of these (for example, GenBank: https://www.ncbi.nlm.nih.gov/Taxonomy/TaxIdentifier/tax_identifier.cgi).

Life forms

There are 3,962 species of trees in New Guinea, and these account for 29% of the flora (Fig. 3). The most-diverse 'tree families' (that is, those in which more than 50% of species are trees) are Myrtaceae (329 tree species), Lauraceae (240), Euphorbiaceae (204), Phyllanthaceae (167) and Moraceae (161). For comparison, Amazonia has 2.6 times more tree species, but in an area 6.4 times larger[21]. Taxonomic monographs have been completed for Moraceae for Flora Malesiana (an international project initiated in 1950 that aims to name, describe and inventory the vascular plants of the Malay Archipelago[22]), and partly for Euphorbiaceae and Phyllanthaceae, but monographs are urgently needed for the large families of trees Lauraceae and Myrtaceae (Supplementary Tables 3, 4). Species with 'non-tree' life forms (herbs, epiphytes, shrubs, climbers, palms and tree ferns) account for 71% of the vascular plant diversity of New Guinea (9,672 species; see Methods). The endemism of non-tree species resembles that for trees (68%) and the majority of the species diversity in New Guinea's endemic genera consists of non-trees (63% of species). Non-tree species diversity is greatest in Orchidaceae, Rubiaceae, Poaceae, Ericaceae and Arecaceae, and non-tree species of these families constitute about one third of the New Guinea flora. Except for Ericaceae, Flora Malesiana accounts are lacking for these species-rich non-tree families (Supplementary Tables 3, 4).

Expert knowledge in the digital era

We sought to ascertain what the total number of vascular plant species reported for New Guinea would be if we resolved names using online tools rather than expert knowledge. To assess this, we first submitted the list of 23,381 unique names to the Taxonomic Name Resolution Service (TNRS), an online name standardization platform[23] that is regularly used in macroecological studies[24]. We found that TNRS accepted 17,518 vascular plant species, or 75% of the names in the original list, whereas our 99 experts accepted 53%. There were significant differences in the number of species reported by TNRS and by experts; the numbers ranged from 0–275 species per family (mean, 16 ± 35; Wilcoxon signed-rank test, $V = 1{,}712$, $P < 0.001$). We reviewed all accepted TNRS names to assess whether these were native to New Guinea, because even accepted names can have geographic errors (non-native taxa). We found that 14% of taxonomically valid TNRS species were false presences. The families with the greatest incidence of false presences were Orchidaceae (244 species; 10% of total false presences), Poaceae (7%), Fabaceae (5%) and Myrtaceae (3%). To assess the quality of our checklist, we also performed an independent comparison with a New Guinea list in Plants of the World Online (POWO; http://www.plantsoftheworldonline.org, accessed 21 December 2019). POWO is a dynamic taxonomic portal based on mined literature that aims to become the most comprehensive single information resource covering all vascular plants by 2020. We found that POWO accepted 13,073 species for New Guinea, of which 1,714 species were synonyms and/or non-native taxa according to experts—making the final species count in POWO 17% lower than ours. Still, the POWO list had 259 native and accepted species (that is, not synonymized) that experts missed and which were subsequently added to the checklist. Overall, the independent comparisons with TNRS and POWO confirm the high quality of our checklist and highlight the need for expert knowledge in the digital era. Although New Guinea lags behind other tropical regions in taxonomic effort, uncertainty in taxonomic names and geographic occurrences is common even in better-studied regions. For example, an improved knowledge of the size of the Amazonian tree flora[21] was only achieved after a series of steps that reduced uncertainty[25,26]—underscoring the importance of dynamic lists and international collaboration networks. Because the importance of online taxonomic tools will continue to grow in the digital era, collaboration among taxonomists, ecologists and maintainers of online synonymy portals will be essential to enhance the quality of online tools such as TNRS.

Completing the New Guinea Flora

Our checklist with resolved plant names, geographic distributions and life forms (Supplementary Tables 1, 2) represents the first, to our

Fig. 2 | Representatives of species-rich genera with more than 80 species in New Guinea. **a**, *Bulbophyllum*; **b**, *Dendrobium*; **c**, *Crepidium*; **d**, *Taeniophyllum*; **e**, *Oberonia*; **f**, *Phreatia*; **g**, *Glomera*; **h**, *Syzygium*; **i**, *Rhododendron*; **j**, *Cyrtandra*; **k**, *Timonius*; **l**, *Freycinetia*; **m**, *Saurauia*; **n**, *Begonia*; **o**, *Medinilla*; **p**, *Ficus*; **q**, *Myristica*; **r**, *Psychotria*; **s**, *Vaccinium*. Photograph credits: A.S. (**a**–**f**), W.J.B. (**g**, **s**), Y.W.L. (**h**), T.U. (**i**–**l**, **o**, **q**), M.S.A. (**m**, **n**) and Z.E. (**p**, **r**).

knowledge, large-scale international attempt to catalogue the entire native flora of New Guinea beyond local lists[27]. Since the publication of the Flora of Java 50 years ago[12] and that of the Philippines in 2011[13], ours is the only other published vascular plant checklist of a large Malesian island or island group. An expert-vetted checklist for New Guinea will be invaluable for conservation planning, as accepted plant names and geographic distributions are the basis of policy-relevant International Union for Conservation of Nature (IUCN) Red List assessments, and are also used for modelling the effects of changes in climate and land use on species ranges. In addition, an authoritative checklist of plant names will improve the accuracy of biogeographic studies (for example, bioregionalization, molecular phylogenies) and trait-based approaches. DNA sequence data are lacking for most taxa in New Guinea, and our checklist will enable more-precise targeting of taxa for sequencing in species-rich groups with poor generic delimitation and high endemism (for example, Lauraceae). Finally, our checklist will aid in the discovery and characterization of more species by taxonomists. By cataloguing 13,634 plant species in the world's most biodiverse island in one year, our rapid collaborative assessment—facilitated by centuries of botanical collections and digital verifiable records—can also serve as a model for accelerating research in other hyperdiverse areas (for example, Borneo and Sumatra). Three conditions will help to increase the speed at which verified species checklists are produced in other hyperdiverse regions: (i) specimens and literature are accessible, physically and digitally in online portals; (ii) family experts exist and their institutions support them; and (iii) coordinator(s) have clear goals, time-delimited guidelines and promote international collaboration.

Species discovery shows no sign of levelling off, especially for non-tree life forms (Fig. 4) and we propose six steps to accelerate the cataloguing of the New Guinea flora. First, training the next generation of resident plant taxonomists is urgently needed. The plants of New Guinea have been studied mostly by people who are not residents[2], and 40% of the experts in our consortium are either retired or within ten years of turning 65 (International Plant Names Index, https://www.ipni.org). Unless the number of resident taxonomic leaders increases, the future of taxonomy in New Guinea will continue to depend on foreign experts. Thus, in-country and international training programmes (for example, postgraduate studies, parataxonomy courses[15,28]) will continue to be essential both for documenting the flora of the region and to increase exchange with Malesian plant taxonomy experts. To build capacity at all levels—from Indigenous citizen scientists to postgraduate students—universities and botanical gardens should align their training and research plans, and partner with embedded institutions such as non-governmental organizations (NGOs). Second, international-scale efforts to digitize and unify historical collections—as proposed by the Distributed System of Scientific Collections initiative (https://www.dissco.eu), for example—are critically needed to underpin research and to repatriate type specimens in digital format. So far, Indonesia's largest herbarium (Herbarium Bogoriense) has digitized around 20,000 type specimens (http://ibis.biologi.lipi.go.id/) but not the general collection; the Royal Botanic Gardens, Kew, the Royal Botanic Gardens and Domain Trust in Sydney and Singapore Botanic Gardens have digitized less than 30% of their New Guinea collections, and the Australian National Herbarium and the Papua New Guinea Forest Research Institute just 50%; the Naturalis Biodiversity Center in Leiden has photographed all specimens and most label information is available online; and only Queensland Herbarium is almost fully digitized. It is insufficient to digitize herbaria, however, if there are high rates of specimen misidentification[29]. Thus, our third recommendation is that critical taxonomic research—especially in species-rich genera (Extended Data Table 2)—needs long-term institutional and financial support if substantial advances are to be made. Otherwise, erroneous taxonomic determinations will persist, causing species numbers to be over- or underestimated. For example, the early twentieth century

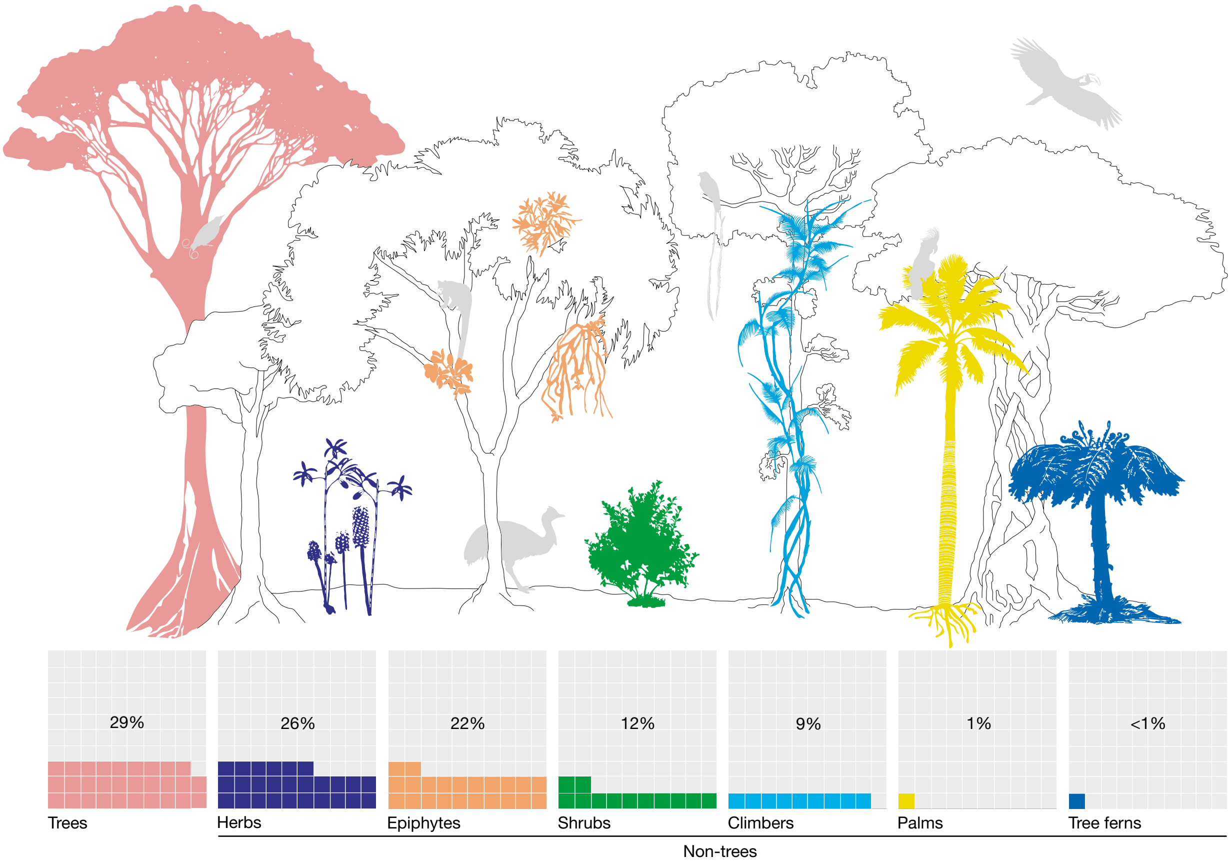


Fig. 3 | Breakdown of the New Guinea flora by life form. Fraction of species that are trees (pink), herbs (dark blue), epiphytes (orange), shrubs (green), climbers (light blue), non-climbing palms (yellow) and tree ferns (mid blue).

boom in botanical discoveries in New Guinea (Fig. 4) was largely due to Rudolf Schlechter, who described more than 1,000 new species and had long-term support. Often, scientists trained abroad who return home encounter heavy teaching loads, large administrative obligations and low salaries[30]. This may explain why only two complete Flora Malesiana accounts, and few genera in multi-authored accounts, have been written by an Indonesian person, and none by an individual from Papua New Guinea. Currently, there are very limited career opportunities for plant taxonomists in Indonesian New Guinea and Papua New Guinea. Boosting the role of resident botanists in understanding the New Guinea flora will thus require governmental measures that create jobs, improve professional conditions for taxonomists and reward scientific productivity and merit. A fourth step will be to increase the number and quality of user-friendly plant field guides[31]. This will be crucial to raise awareness of the region's plants and enhance collecting, identification and cataloguing efforts. As a fifth step, countries should support more international collaborations, because reciprocal exchanges to co-write taxonomic papers provide tangible benefits to Flora projects[32]. Finally, collecting effort is still low (fewer than 25 collections per 100 km^2 throughout much of New Guinea[2]; Fig. 1a) and land-use change is an increasing threat[33], so more botanical exploration is therefore urgently needed if unknown species are to be collected before they disappear. Considering that 2,812 new species have been published since 1970, and that larger and higher-diversity genera still need to be tackled, we estimate that in 50 years 3,000–4,000 species will be added to the number of vascular plants in New Guinea. Species discovery, however, will ultimately depend on enough experts being available to study the large number of collections that have been amassed in the past decades (Extended Data Fig. 2), including thousands of specimens that remain unidentified (Extended Data Table 4).

Knowledge on the flora of New Guinea has remained scattered for too long, which has limited basic and applied research in this highly diverse tropical wilderness area. Here, we have built an expert-verified checklist of New Guinea's 13,634 known vascular plant species and made it openly available to the global community. The checklist suggests that New Guinea is the most floristically diverse island in the world and that its high level of endemism is unmatched in tropical Asia. Our work

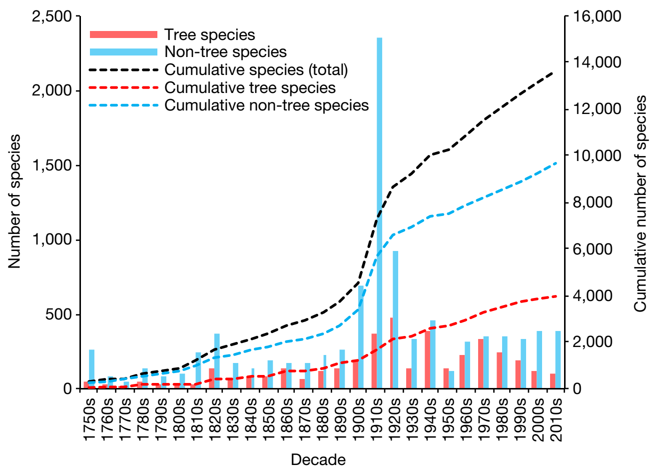


Fig. 4 | Species described per decade in New Guinea. The number of plant species (basionyms) described per decade from 1753 to 2019, grouped into tree species (red bars) and non-tree species (blue bars); and the cumulative number of verified species of trees (red dotted line), non-trees (blue dotted line) and total (black dotted line).

demonstrates that international collaborative efforts using verified digital data can rapidly synthesize biodiversity information. In doing so, such initiatives inform other branches of science and pave the way for the grand challenge of conserving New Guinea's rich flora.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2549-5.

1. Frodin, D. in *The Ecology of Papua, Part One* (eds Marshall, A. & Beehler, B.) 87–130 (Periplus, 2007).
2. Conn, B. J. in *Biodiversity and Terrestrial Ecosystems* (eds Peng, C.-I. & Chou, C.-H.) 123–156 (Institute of Botany, Academia Sinica, 1994).
3. Mittermeier, R. A. et al. Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conserv. Biol*. **12**, 516–520 (1998).
4. Paijmans, K. *New Guinea Vegetation* (Commonwealth Scientific and Industrial Research Organization in association with the Australian National Univ. Press, 1976).
5. Roos, M. C. et al. Species diversity and endemism of five major Malesian islands: diversity–area relationships. *J. Biogeogr*. **31**, 1893–1908 (2004).
6. Loh, J. & Harmon, D. A global index of biocultural diversity. *Ecol. Indic*. **5**, 231–241 (2005).
7. Cámara-Leret, R. & Dennehy, Z. Information gaps of indigenous and local knowledge for science-policy assessments. *Nat. Sustain*. **2**, 736–741 (2019).
8. Souter, G. *New Guinea: The Last Unknown* (Angus and Robertson, 1963).
9. Good, R. On the geographical relationships of the angiosperm flora of New Guinea. *Bull. Br. Mus. Nat. Hist*. **2**, 205–226 (1960).
10. Supriatna, J. et al. *The Irian Jaya Biodiversity Conservation Priority-Setting Workshop* (Conservation International, 1999).
11. Madagascar Catalogue. *Catalogue of the Plants of Madagascar* (Missouri Botanical Garden, St Louis, USA, accessed August 2019); http://www.tropicos.org/Project/Madagascar
12. Backer, C. A. & Bakhuizen van den Brink, R. C. Jr. *Flora of Java* Vol. 1–3 (Wolters Noordhoff, 1968).
13. Pelser, P. B., Barcelona, J. F. & Nickrent, D. L. (eds) *Co's Digital Flora of the Philippines* (accessed August 2019); www.philippineplants.org
14. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton Univ. Press, 1967).
15. Webb, C. O., Slik, J. F. & Triono, T. Biodiversity inventory and informatics in Southeast Asia. *Biodivers. Conserv*. **19**, 955–972 (2010).
16. Hoover, J. D. et al. Modeling hotspots of plant diversity in New Guinea. *Trop. Ecol*. **58**, 623–640 (2017).
17. Ulloa Ulloa, C. et al. An integrated assessment of the vascular plant species of the Americas. *Science* **358**, 1614–1617 (2017).
18. van Welzen, P. C., Slik, J. W. F. & Alahuhta, J. Plant distributions and plate tectonics in Malesia. *Biologiske Skrifter* **55**, 199–217 (2005).
19. Baldwin, S. L., Fitzgerald, P. G. & Webb, L. E. Tectonics of the New Guinea region. *Annu. Rev. Earth Planet. Sci*. **40**, 495–520 (2012).
20. Steinbauer, M. J. Topography-driven isolation, speciation and a global increase of endemism with elevation. *Glob. Ecol. Biogeogr*. **25**, 1097–1107 (2016).
21. ter Steege, H. et al. Towards a dynamic list of Amazonian tree species. *Sci. Rep*. **9**, 3501 (2019).
22. van Steenis, C. G. G. J. (ed.) *Flora Malesiana series I* Vol. I (Noordhoff-Kolff, 1950).
23. Boyle, B. et al. The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics* **14**, 16 (2013).
24. Lamanna, C. et al. Functional trait space and the latitudinal diversity gradient. *Proc. Natl Acad. Sci. USA* **111**, 13745–13750 (2014).
25. ter Steege, H. et al. The discovery of the Amazonian tree flora with an updated checklist of all known tree taxa. *Sci. Rep*. **6**, 29549 (2016).
26. Cardoso, D. et al. Amazon plant diversity revealed by a taxonomically verified species list. *Proc. Natl Acad. Sci. USA* **114**, 10695–10700 (2017).
27. Coode, M. J. E., Hinchcliffe, S. C. & Marsden, C. J. *Checklist of the Flowering Plants of N. E. Kepala Burung (Vogelkop), Irian Jaya, Indonesia* (Royal Botanic Gardens, Kew, 1997).
28. Basset, Y. et al. Quantifying biodiversity: experience with parataxonomists and digital photography in Papua New Guinea and Guyana. *Bioscience* **50**, 899–908 (2000).
29. Goodwin, Z. A., Harris, D. J., Filer, D., Wood, J. R. & Scotland, R. W. Widespread mistaken identity in tropical plant collections. *Curr. Biol*. **25**, R1066–R1067 (2015).
30. Goss, A. *The Floracrats: State-Sponsored Science and the Failure of the Enlightenment in Indonesia* (Univ. Wisconsin Press, 2011).
31. Conn, B. J. & Damas, K. Q. *Trees of Papua New Guinea* Vol. 1–3 (Xlibris, 2019).
32. Newman, M., Chayamarit, K. & Balslev, H. in *Tropical Plant Collections: Legacies from the Past? Essential Tools for the Future* (eds Friis, I. & Balslev, H.) 177–186 (Royal Danish Academy of Sciences and Letters, 2017).
33. Cámara-Leret, R. et al. The Manokwari Declaration: challenges ahead in conserving 70% of Tanah Papua's forests. *Forest Soc*. **3**, 148–151 (2019).

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

Rodrigo Cámara-Leret[1,56]✉, David G. Frodin[1,57], Frits Adema[2], Christiane Anderson[3], Marc S. Appelhans[4], George Argent[5,57], Susana Arias Guerrero[2], Peter Ashton[1], William J. Baker[1], Anders S. Barfod[6], David Barrington[7], Renata Borosova[1], Gemma L. C. Bramley[1], Marie Briggs[1], Sven Buerki[8], Daniel Cahen[1], Martin W. Callmander[9], Martin Cheek[1], Cheng-Wei Chen[10], Barry J. Conn[11], Mark J. E. Coode[1], Iain Darbyshire[1], Sally Dawson[1], John Dransfield[1], Clare Drinkell[1], Brigitta Duyfjes[2], Atsushi Ebihara[12], Zacky Ezedin[13,14], Long-Fei Fu[15], Osia Gideon[16], Deden Girmansyah[17], Rafaël Govaerts[1], Helen Fortune-Hopkins[1], Gustavo Hassemer[18], Alistair Hay[19], Charlie D. Heatubun[1,20,21], D. J. Nicholas Hind[1], Peter Hoch[22], Peter Homot[23], Peter Hovenkamp[2,57], Mark Hughes[5], Matthew Jebb[24], Laura Jennings[1], Tiberius Jimbo[23], Michael Kessler[25], Ruth Kiew[26], Sandra Knapp[27], Penniel Lamei[23], Marcus Lehnert[28,29], Gwilym P. Lewis[1], Hans Peter Linder[25], Stuart Lindsay[30], Yee Wen Low[1,31,32], Eve Lucas[1], Jeffrey P. Mancera[33], Alexandre K. Monro[1], Alison Moore[1], David J. Middleton[31], Hidetoshi Nagamasu[34], Mark F. Newman[5], Eimear Nic Lughadha[1], Pablo H. A. Melo[35], Daniel J. Ohlsen[36,37], Caroline M. Pannell[1,38,39], Barbara Parris[40], Laura Pearce[1], Darin S. Penneys[41], Leon R. Perrie[42], Peter Petoe[1,6], Axel Dalberg Poulsen[5], Ghillean T. Prance[1], J. Peter Quakenbush[43], Niels Raes[2], Michele Rodda[31], Zachary S. Rogers[44], André Schuiteman[1], Pedro Schwartsburd[45], Robert W. Scotland[38], Mark P. Simmons[46], David A. Simpson[1,47], Peter Stevens[22], Michael Sundue[7], Weston Testo[48], Anna Trias-Blasi[1], Ian Turner[1,31], Timothy Utteridge[1], Lesley Walsingham[1], Bruce L. Webber[49,50], Ran Wei[51], George D. Weiblen[13,52], Maximilian Weigend[28], Peter Weston[19], Willem de Wilde[2], Peter Wilkie[5], Christine M. Wilmot-Dear[1], Hannah P. Wilson[5,53], John R. I. Wood[1,38], Li-Bing Zhang[22,54] & Peter C. van Welzen[2,55]

[1]Royal Botanic Gardens, Kew, London, UK. [2]Naturalis Biodiversity Center, Leiden, The Netherlands. [3]University of Michigan Herbarium, Ann Arbor, MI, USA. [4]Department of Systematics, Biodiversity and Evolution of Plants, Albrecht-von-Haller Institute of Plant Sciences, University of Göttingen, Göttingen, Germany. [5]Royal Botanic Garden Edinburgh, Edinburgh, UK. [6]Department of Bioscience, Aarhus University, Aarhus, Denmark. [7]The Pringle Herbarium, Department of Plant Biology, University of Vermont, Burlington, VT, USA. [8]Department of Biological Sciences, Boise State University, Boise, ID, USA. [9]Conservatoire et Jardin botaniques de la Ville de Genève, Chambésy, Switzerland. [10]Independent consultant, Keelung, Taiwan. [11]School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, Australia. [12]Department of Botany, National Museum of Nature and Science, Tsukuba, Japan. [13]Department of Plant and Microbial Biology, University of Minnesota, Saint Paul, MN, USA. [14]New Guinea Binatang Research Center, Madang, Papua New Guinea. [15]Guangxi Key Laboratory of Plant Conservation and Restoration Ecology in Karst Terrain, Guangxi Institute of Botany, Guangxi Zhuang Autonomous Region and Chinese Academy of Sciences, Guilin, China. [16]The Papua New Guinea University of Technology, Lae, Papua New Guinea. [17]Herbarium Bogoriense, Bogor, Indonesia. [18]Três Lagoas Campus, Federal University of Mato Grosso do Sul, Três Lagoas, Brazil. [19]Royal Botanic Garden Sydney, Sydney, New South Wales, Australia. [20]Fakultas Kehutanan, Universitas Papua, Manokwari, Indonesia. [21]Badan Penelitian dan Pengembangan Daerah Provinsi Papua Barat, Manokwari, Indonesia. [22]Missouri Botanical Garden, St Louis, MO, USA. [23]Papua New Guinea Forest Research Institute, Lae, Papua New Guinea. [24]The National Botanic Gardens of Ireland, Glasnevin, Ireland. [25]Department of Systematic and Evolutionary Botany, University of Zurich, Zurich, Switzerland. [26]Forest Research Institute Malaysia, Kepong, Malaysia. [27]Department of Life Sciences, Natural History Museum, London, UK. [28]Nees Institute for Biodiversity of Plants, University of Bonn, Bonn, Germany. [29]Institute of Biology, Herbarium, Martin Luther University Halle-Wittenberg, Halle, Germany. [30]Native Plants Centre, National Parks Board, Singapore, Singapore. [31]Singapore Botanic Gardens, National Parks Board, Singapore, Singapore. [32]School of Biological Sciences, University of Aberdeen, Aberdeen, UK. [33]Department of Biology, College of Arts and Sciences, University of the Philippines Manila, Manila, The Philippines. [34]The Kyoto University Museum, Kyoto University, Kyoto, Japan. [35]Institute of Biosciences, São Paulo State University, Rio Claro, Brazil. [36]Royal Botanic Gardens Victoria, South Yarra, Victoria, Australia. [37]School of BioSciences, The University of Melbourne, Parkville, Victoria, Australia. [38]Department of Plant Sciences, University of Oxford, Oxford, UK. [39]Marine Laboratory, Queen's University Belfast, Portaferry, UK. [40]Fern Research Foundation, Kerikeri, New Zealand. [41]University of North Carolina, Wilmington, NC, USA. [42]Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand. [43]Department of Biological Sciences, Western Michigan University, Kalamazoo, MI, USA. [44]Department of Animal and Range Sciences, New Mexico State University, Las Cruces, NM, USA. [45]Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, Brazil. [46]Department of Biology, Colorado State University, Fort Collins, CO, USA. [47]Department of Botany, School of Natural Sciences, Trinity College Dublin, Dublin, Ireland. [48]Department of Biology, University of Florida, Gainesville, FL, USA. [49]CSIRO Land and Water, Floreat, Western Australia, Australia. [50]School of Biological Sciences, University of Western Australia, Crawley, Western Australia, Australia. [51]Institute of Botany, Chinese Academy of Sciences, Beijing, China. [52]Bell Museum, University of Minnesota, Saint Paul, MN, USA. [53]Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK. [54]Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China. [55]Institute of Biology Leiden, University of Leiden, Leiden, The Netherlands. [56]Present address: Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland. [57]Deceased: David G. Frodin, George Argent, Peter Hovenkamp. ✉e-mail: rodrigo.camaraleret@ieu.uzh.ch

Methods

Study area
We defined the study area as the region encompassing the main island of New Guinea and the surrounding smaller islands that were connected to mainland New Guinea during the Last Glacial Maximum. We delimit it by selecting areas within a depth of −120 m of mainland New Guinea from the General Bathymetric Chart of the Oceans (http://www.gebco.net) (Extended Data Fig. 1). Accordingly, the study area spans a latitudinal range of 0.08° S to 10.66° S and a longitudinal range of 129.42° E to 150.21° E and excludes the Moluccas and Kai Islands to the west, Bougainville and the Solomon Islands to the east and the Micronesian islands to the north. Large islands in our study area include New Guinea, the Aru islands, the Raja Ampat islands, Biak, Yapen, New Britain, New Ireland and the Louisiade, Admiralty and Western islands.

Data compilation
An initial list of plant names for the study area was compiled from specimen data after four steps. First, we downloaded specimens within the extent of our study area from the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/occurrence/download/0064983-160910150852091, n = 394,821 records), Consortium of Pacific Herbaria (CPH, https://www.re3data.org/repository/r3d100012011, n = 30,188), Australasian Virtual Herbarium (AVH, http://avh.ala.org.au, n = 42,714) and Kew Herbarium Catalogue (http://apps.kew.org/herbcat/, n = 4,618). Second, we obtained herbarium specimen records from institutional repositories of the Naturalis Biodiversity Center (n = 189,382), Royal Botanic Gardens, Kew (n = 56,522) and University of Papua New Guinea (n = 17,929). Third, we downloaded type specimens from the Harvard University Herbaria (https://kiki.huh.harvard.edu/databases/specimen_index.html, n = 5,571), Natural History Museum (https://data.nhm.ac.uk, n= 1,325), New York Botanical Garden (http://sweetgum.nybg.org/science/collections, n = 1,236), Royal Botanic Garden Edinburgh (https://data.rbge.org.uk/search/herbarium, n = 1,200), Smithsonian National Museum of Natural History (https://collections.nmnh.si.edu/search/botany, n = 1,025), Missouri Botanical Garden (http://www.tropicos.org, n = 51) and Muséum National d'Histoire Naturelle (https://science.mnhn.fr/institution/mnhn/search, n = 32). Fourth, we obtained data curated by taxonomists for Orchidaceae (n = 12,830), Arecaceae (n = 3,684), Araliaceae (n = 1,713) and Cyatheaceae (n = 1,662). We manually unified headers and standardized entries for the fields of family, genus, species, collector name, collector number, date and elevation. Family circumscriptions were based on the Angiosperm Phylogeny Group IV[34] (angiosperms), on the Pteridophyte Phylogeny group (ferns and lycophytes)[35] and on a previous study (gymnosperms)[36]. All records from outside the study area were removed (that is, from Sumatra, Java, Borneo, Bali, Komodo, Flores, Moluccas, Solomon Islands and Bougainville). Names of collectors were verified using the Cyclopaedia of Malesian Collectors (http://www.nationaalherbarium.nl/FMCollectors/). Collectors' names that were absent from the Cyclopaedia of Malesian Collectors were reviewed by R.C.-L. and D.G.F., the latter an expert on the history of biological exploration in New Guinea[1].

We applied different quality filters to clean scientific names. First, fungi, lichens, algae, bryophytes and marine species (for example, sea grasses) were excluded. Second, doubtful species identifications (for example, 'cf.', 'sp. nov.', 'aff.', 'sp.') were classified to generic level. The list of genera was then used as the basis to query TNRS[23]. Misspelled genera were manually corrected and doubtful cases excluded. We removed all known hybrids from the analyses. The resulting list of 23,381 taxonomic names was submitted to TNRS for verification.

Expert review
From April to November 2018, 99 taxonomic, floristic and monographic experts (see author list) of the New Guinea flora reviewed the list of original names in their respective families of expertise (Supplementary Tables 1, 2). Each expert verified whether the original list of names was correctly resolved by TNRS, and included additional information about taxonomy (basionym name, basionym year), geographic range (native, endemic, distribution in Indonesia and/or Papua New Guinea) and life form (tree, herb, shrub, epiphyte, palm, etc.). When experts considered that a name that was accepted by TNRS was not correct, they wrote the correct name and cited the source(s) for these changes. Similarly, when experts considered a species not to be native, they were asked to write an explanation (for example, geographic error, taxonomic misidentification). Finally, experts also included names that had been missed from the original list (n = 1,263). To assess the discrepancy between TNRS and expert verification, we compared the total number of accepted species in both lists for 254 plant families by using a Wilcoxon signed-rank test. We then performed an independent comparison against a list of 13,073 accepted species names contained in POWO for the 'New Guinea' locality (http://www.plantsoftheworldonline.org, accessed 21 December 2019). POWO was launched in 2017 with an initial focus on tropical Africa, but aims to become a single point of access for authoritative information on all plant species by 2020. Accordingly, for the names in the POWO list that were missing in our checklist, experts assessed whether they were incorrect (that is, synonyms) and/or not native to the study area, and which names were correct and native. The former represent false presences in POWO; the latter represent species that experts overlooked and which were subsequently included in the final checklist.

Life forms and species discovery over time
To assess the percentage of tree and non-tree species within each family, we considered 'non-trees' to comprise the following life forms that lack distinct secondary wood growth or have multiple woody stems: herbs, epiphytes, shrubs, climbers, palms and tree ferns (Fig. 3). Families in which more than 50% of the species were trees were considered 'tree families'. To assess the rate at which species names in the checklist have been described and accepted, we compiled the year of publication of basionyms from the primary literature, the International Plant Names Index (https://www.ipni.org) and the Tropicos database (https://tropicos.org). To map collections of native species (Fig. 1a), we discarded duplicate records (that is, those with the same collector name, collector number, latitude and longitude) and records that lacked coordinates or that had coordinates within the sea. This resulted in a total of 153,979 unique records. A richness map using a 25 × 25-km grid was built in R[37] using commands from the libraries raster[38] and letsR[39] and artwork was designed using QGIS[40].

Reporting summary
Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability
The data that support the findings of this study are available within the Article and in Supplementary Tables 1–4.

Code availability
The R code used for calculations and analyses is available from the corresponding author on request.

34. APG IV. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* **181**, 1–20 (2016).
35. Pteridophyte Phylogeny Group. A community-derived classification for extant lycophytes and ferns. *J. Syst. Evol.* **54**, 563–603 (2016).
36. Christenhusz, M. J. et al. A new classification and linear sequence of extant gymnosperms. *Phytotaxa* **19**, 55–70 (2011).
37. R Core Team. R: a language and environment for statistical computing (R Foundation for Statistical Computing, 2019).

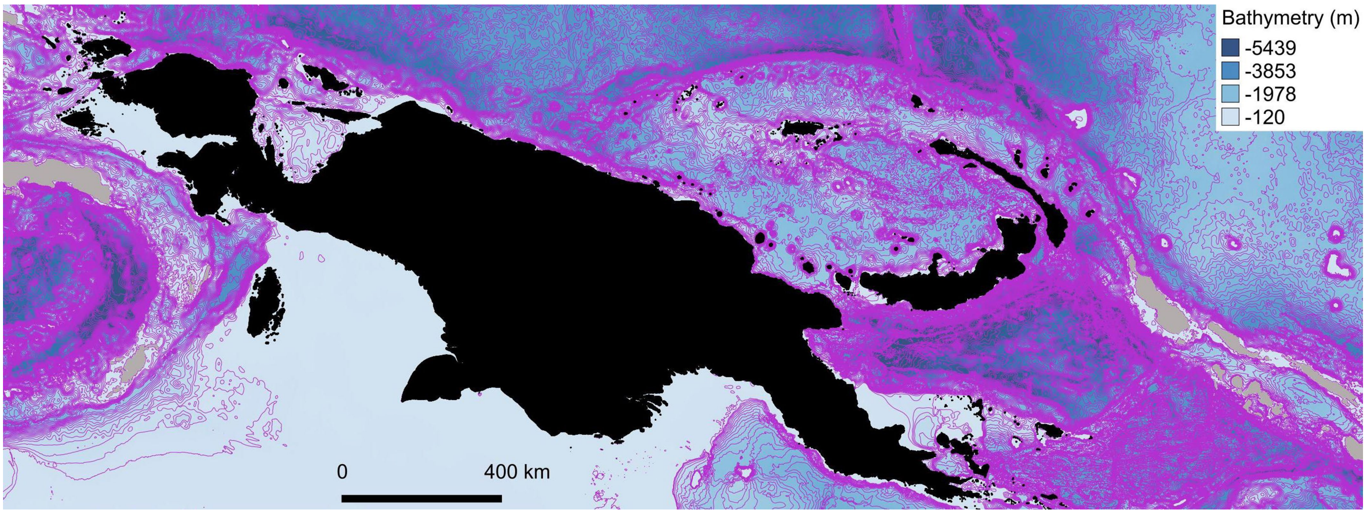
38. Hijmans, R. & van Etten, J. raster: Geographic data analysis and modeling. R package v.2.8-19. https://cran.r-project.org/web/packages/raster/index.html (2019).
39. Vilela, B. & Villalobos, F. letsR: a new R package for data handling and analysis in macroecology. *Methods Ecol. Evol*. **6**, 1229–1234 (2015).
40. QGIS Development Team. QGIS: a Free and Open Source Geographic Information System. http://qgis.osgeo.org (2020).

Acknowledgements We thank the collectors, herbarium and database curators and colleagues at our institutions for support and discussions; J. Bascompte for feedback on a previous draft; and I. Cámara Leret for the design of Fig. 3. R.C.-L. and P.C.v.W. received funding from the Royal Society International Exchanges (grant IE 170241, 'Building the New Guinea Research Team'); A.S.B. acknowledges financial support from the Carlsberg Foundation and the Danish Research Council; L.-F.F. acknowledges the National Natural Science Foundation of China (grant 31570307); G.H. acknowledges the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for a post-doctoral fellowship (process 153430/2018-4); S.K. was supported by the US National Science Foundation (grant DEB-1457366); Y.W.L is funded by the National Parks Board, Singapore under a postgraduate research scholarship; M.L. received funding from the Deutsche Forschungsgemeinschaft (LE 1826/3-1) and Synthesys (GB-TAF-6305); P.H.A.M. acknowledges the São Paulo Research Foundation – FAPESP (grant 2015/13112-7 and 2018/09379-6); D.S.P. received funding from the US National Science Foundation (grant DEB-1754667); the contributions of G.D.W. to the project were supported by the US National Science Foundation (grants DEB-0515678, DEB-0816749 and DEB-0841885); and M.H. and H.P.W. received funding from the Edinburgh Botanic Garden (Sibbald) Trust and the M. L. MacIntyre Begonia Trust. The Royal Botanic Garden Edinburgh is supported by the Scottish Government's Rural and Environmental Science and Analytical Services Division.

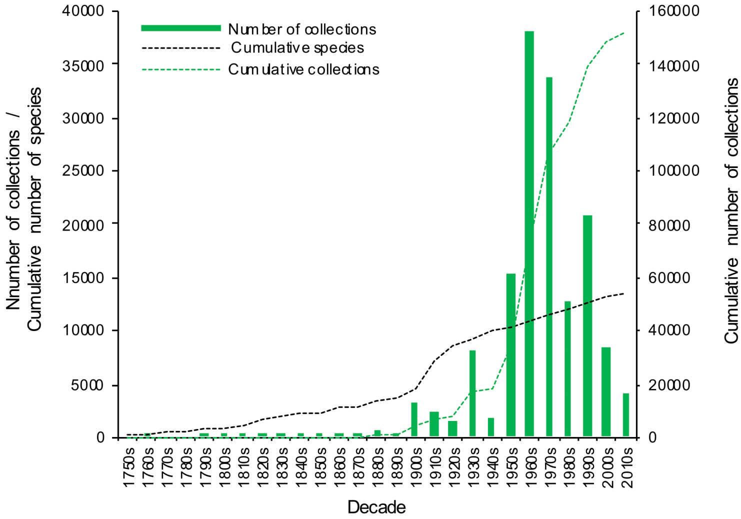
Author contributions R.C.-L. conceived the study, analysed the data and wrote a first draft of the paper. All authors verified taxonomic data and contributed to revisions.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2549-5.
Correspondence and requests for materials should be addressed to R.C.-L.
Peer review information *Nature* thanks Timothy Baker, Lars Chatrou, Thomas Givnish, Vojtech Novotny and Hans ter Steege for their contribution to the peer review of this work.



Extended Data Fig. 1 | Delimitation of the study area of New Guinea. The study area (black islands) includes islands within a depth of −120 m of mainland New Guinea according to the General Bathymetric Chart of the Oceans (http://www.gebco.net). Purple lines depict seafloor depth starting at −120 m.



Extended Data Fig. 2 | Collection effort and discovery of the New Guinea flora through time. The number of plant collections that have been digitized (green bars), the cumulative total number of collections (green dotted line) and the cumulative number of plant species (basionyms) described over time (black dotted line).

Extended Data Table 1 | The 31 plant families in New Guinea that have more than 100 species, arranged in descending order of native species

Family	Native species	Endemic species	% Endemic
Orchidaceae	2856	2464	86.3
Rubiaceae	784	668	85.2
Ericaceae	438	431	98.4
Poaceae	376	113	30.1
Myrtaceae	352	256	72.7
Arecaceae	292	257	88
Fabaceae	289	79	27.3
Apocynaceae	284	196	69
Cyperaceae	254	24	9.4
Lauraceae	242	194	80.2
Euphorbiaceae	241	157	65.1
Polypodiaceae	235	108	46
Gesneriaceae	228	218	95.6
Melastomataceae	228	188	82.5
Thelypteridaceae	203	140	69
Zingiberaceae	202	191	94.6
Phyllanthaceae	202	138	68.3
Moraceae	201	96	47.8
Pandanaceae	200	175	87.5
Urticaceae	196	158	80.6
Asteraceae	171	109	63.7
Sapindaceae	149	100	67.1
Rutaceae	143	100	69.9
Malvaceae	141	74	52.5
Myristicaceae	134	115	85.8
Primulaceae	133	113	85
Araliaceae	131	107	81.7
Elaeocarpaceae	127	108	85
Araceae	111	73	65.8
Dryopteridaceae	110	52	47.3
Cyatheaceae	110	92	83.6

Extended Data Table 2 | The 20 most-diverse plant genera in New Guinea, arranged in descending order of native species

Genus	Family	Native species	Endemic species	% Endemic
Bulbophyllum	Orchidaceae	658	599	91
Dendrobium	Orchidaceae	614	528	86
Syzygium	Myrtaceae	207	174	84.1
Ficus	Moraceae	179	93	52
Rhododendron	Ericaceae	171	169	98.8
Psychotria	Rubiaceae	146	144	98.6
Glomera	Orchidaceae	143	138	96.5
Freycinetia	Pandanaceae	140	127	90.7
Phreatia	Orchidaceae	138	128	92.8
Vaccinium	Ericaceae	135	135	100
Taeniophyllum	Orchidaceae	130	125	96.2
Cyrtandra	Gesneriaceae	112	108	96.4
Crepidium	Orchidaceae	110	102	92.7
Myristica	Myristicaceae	99	89	89.9
Saurauia	Actinidiaceae	93	91	97.8
Medinilla	Melastomataceae	91	82	90.1
Macaranga	Euphorbiaceae	91	76	83.5
Begonia	Begoniaceae	90	86	95.6
Timonius	Rubiaceae	89	80	89.9
Oberonia	Orchidaceae	87	81	93.1

Extended Data Table 3 | The 61 endemic genera to New Guinea, their number of species and availability of sequences in GenBank

Genus	Family	Species	GenBank sequences
Aistopetalum	Cunoniaceae	2	no
Anakasia	Araliaceae	1	no
Annesijoa	Euphorbiaceae	1	no
Anthorrhiza	Rubiaceae	9	yes
Antiaropsis	Moraceae	2	yes
Archboldiodendron	Pentaphylacaceae	1	no
Basisperma	Myrtaceae	1	no
Brachionostylum	Asteraceae	1	no
Brassiophoenix	Arecaceae	2	yes
Buergersiochloa	Poaceae	1	yes
Calycacanthus	Acanthaceae	1	yes
Chaetostachydium	Rubiaceae	3	no
Chimaerochloa	Poaceae	1	yes
Chlaenandra	Menispermaceae	1	no
Cyrtandropsis	Gesneriaceae	14	no
Decatoca	Ericaceae	1	no
Distrianthes	Loranthaceae	2	no
Dolianthus	Rubiaceae	13	yes
Dransfieldia	Arecaceae	1	yes
Dryadorchis	Orchidaceae	5	no
Eleutherostylis	Malvaceae	1	no
Fittingia	Primulaceae	9	no
Gibbsia	Urticaceae	2	yes
Gjellerupia	Opiliaceae	1	no
Gymnophragma	Acanthaceae	1	no
Hartleya	Stemonuraceae	1	no
Holochlamys	Araceae	1	yes
Hulemacanthus	Acanthaceae	2	no
Ischnea	Asteraceae	4	yes
Jadunia	Acanthaceae	2	yes
Kairoa	Monimiaceae	4	yes
Kairothamnus	Picrodendraceae	1	no
Lagenocypsela	Asteraceae	2	no
Lamiodendron	Bignoniaceae	1	yes
Macrococculus	Menispermaceae	1	yes
Magodendron	Sapotaceae	2	yes
Manjekia	Arecaceae	1	yes
Maschalodesme	Rubiaceae	2	no
Novaguinea	Asteraceae	1	no
Opocunonia	Cunoniaceae	1	yes
Pachystylus	Rubiaceae	1	no
Papuacalia	Asteraceae	17	yes
Papuaea	Orchidaceae	1	no
Papuanthes	Loranthaceae	1	no
Papuasicyos	Cucurbitaceae	7	yes
Papuodendron	Malvaceae	2	yes
Paramyristica	Myristicaceae	1	no
Piora	Asteraceae	1	no
Pseudobotrys	Cardiopteridaceae	2	no
Rhadinopus	Rubiaceae	2	no
Rheopteris	Pteridaceae	1	yes
Ruthiella	Campanulaceae	4	no
Sepikea	Gesneriaceae	1	no
Sericolea	Elaeocarpaceae	15	yes
Siphonandrium	Rubiaceae	1	no
Sommieria	Arecaceae	1	yes
Thylacophora	Zingiberaceae	1	no
Thysanosoria	Lomariopsidaceae	1	no
Urceodiscus	Cucurbitaceae	1	yes
Wallaceodoxa	Arecaceae	1	yes
Xylonymus	Celastraceae	1	no

GenBank sequences at https://www.ncbi.nlm.nih.gov/genbank/

Extended Data Table 4 | Number of New Guinea specimens and unidentified specimens, and percentage of unidentified specimens, for larger vascular plant genera held at BISH, BRI, CANB, L, LAE and NSW

	TOTAL			BISH			BRI			CANB			L			LAE			NSW		
Genus	S	U	%U	S	U	%U	S	U	%U	S	U	%U	S	U	%U	S	U	%U	S	U	%U
Aglaia	2353	1271	54	87	86	99	341	196	57	268	239	89	824	76	9	727	574	79	106	100	94
Alpinia	695	514	74	1	0	0	11	2	18	6	1	17	524	417	80	145	89	61	8	5	63
Archidendron	302	29	10	3	0	0	20	0	0	14	0	0	183	13	7	77	14	18	5	2	40
Ardisia	446	188	42	20	5	25	46	14	30	20	7	35	166	60	36	167	93	56	27	9	33
Asplenium	578	56	10	0	0	0	7	6	86	8	2	25	464	14	3	86	31	36	13	3	23
Begonia	1195	341	29	15	0	0	313	192	61	15	10	67	395	83	21	431	40	9	26	16	62
Beilschmiedia	712	171	24	6	0	0	71	45	63	30	14	47	488	61	13	104	48	46	13	3	23
Bulbophyllum	3992	1701	43	4	1	25	139	85	61	951	527	55	1774	496	28	1116	585	52	8	7	88
Casearia	786	265	34	78	45	58	198	99	50	10	5	50	243	94	39	254	21	8	3	1	33
Cryptocarya	3331	1226	37	54	41	76	1088	437	40	168	28	17	977	218	22	957	490	51	87	12	14
Cyrtandra	2091	1401	67	86	12	14	410	363	89	26	16	62	959	839	87	577	147	25	33	24	73
Dendrobium	5383	1133	21	89	12	13	532	123	23	500	388	78	2461	506	21	1720	95	6	81	9	11
Diospyros	1848	365	20	23	5	22	486	121	25	74	28	38	616	169	27	624	41	7	25	1	4
Dysoxylum	2520	910	36	51	18	35	632	248	39	67	33	49	853	67	8	845	527	62	72	17	24
Elaeocarpus	4097	308	8	175	1	1	1084	6	1	151	61	40	1026	58	6	1578	134	8	83	48	58
Elatostema	2785	1735	62	38	22	58	281	33	12	458	377	82	682	424	62	1094	768	70	232	111	48
Endiandra	1084	199	18	29	0	0	266	141	53	69	8	12	340	40	12	340	8	2	40	2	5
Euodia	1141	411	36	29	18	62	221	107	48	33	17	52	282	78	28	555	181	33	21	10	48
Ficus	20109	1979	10	205	71	35	3116	548	18	3733	125	3	3881	91	2	8941	1108	12	233	36	15
Garcinia	3046	1716	56	46	4	9	723	497	69	126	14	11	1164	646	55	923	546	59	64	9	14
Glochidion	1524	335	22	19	4	21	93	5	5	50	3	6	811	192	24	520	123	24	31	8	26
Guioa	1013	162	16	34	1	3	210	31	15	117	31	26	209	3	1	403	90	22	40	6	15
Hedyotis	1190	398	33	35	6	17	214	70	33	113	15	13	321	165	51	485	134	28	22	8	36
Hoya	1954	459	23	14	4	29	338	59	17	345	120	35	372	65	17	859	205	24	26	6	23
Litsea	2378	1031	43	35	12	34	556	390	70	86	23	27	956	152	16	704	440	63	41	14	34
Macaranga	3934	993	25	39	5	13	880	348	40	168	10	6	1545	217	14	1250	406	32	52	7	13
Medinilla	3698	1238	33	76	13	17	853	458	54	279	47	17	1063	115	11	1329	570	43	98	35	36
Mussaenda	1785	465	26	64	17	27	258	106	41	189	22	12	381	133	35	825	171	21	68	16	24
Myristica	4270	838	20	196	87	44	350	94	27	322	83	26	1445	29	2	1770	435	25	187	110	59
Pandanus	960	384	40	7	7	100	278	62	22	1	0	0	364	29	8	304	281	92	6	5	83
Piper	5259	1402	27	149	21	14	1068	300	28	394	50	13	1092	408	37	2354	599	25	202	24	12
Pouteria	4204	875	21	51	4	8	1970	683	35	152	15	10	706	12	2	1240	151	12	85	10	12
Psychotria	6168	2132	35	208	64	31	711	2	0	892	248	28	1171	567	48	2971	1184	40	215	67	31
Rhododendron	2904	148	5	35	14	40	57	24	42	221	5	2	1357	3	0	1150	97	8	84	5	6
Riedelia	722	431	60	26	13	50	56	24	43	0	0	0	465	295	63	161	90	56	14	9	64
Saurauia	5290	2376	45	249	176	71	838	587	70	790	564	71	525	107	20	2625	781	30	263	161	61
Schefflera	2162	497	23	55	3	5	382	130	34	228	43	19	383	19	5	1058	284	27	56	18	32
Solanum	2868	172	6	62	6	10	513	28	5	379	17	4	486	16	3	1326	97	7	102	8	8
Syzygium	4034	1540	38	145	54	37	552	213	39	514	205	40	1293	422	33	1226	561	46	304	85	28
Vaccinium	1116	31	3	1	0	0	6	1	17	6	0	0	1021	4	0	73	24	33	9	2	22

S, specimens; U, unidentified specimens; %U, percentage of unidentified specimens.
Herbarium acronyms: BISH, Bishop Museum; BRI, Queensland Herbarium; CANB, Australian National Herbarium; L, Naturalis; LAE, Papua New Guinea Forest Research Institute; NSW, Royal Botanic Gardens and Domain Trust.

nature research

Corresponding author(s): Rodrigo Cámara-Leret

Last updated by author(s): Jun 8, 2020

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 our Editorial Policies 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	We gathered a list of vascular plant species names by downloading species collections available for our study area from the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/occurrence/download/0064983-160910150852091), Consortium of Pacific Herbaria (CPH, https://www.re3data.org/repository/r3d100012011), Australasian Virtual Herbarium (AVH, http://avh.ala.org.au) and the Kew Herbarium Catalogue (http://apps.kew.org/herbcat/). We also downloaded type specimens from the Harvard University Herbaria (https://kiki.huh.harvard.edu/databases/specimen_index.html), the Natural History Museum (https://data.nhm.ac.uk), New York Botanical Garden (http://sweetgum.nybg.org/science/collections), Royal Botanic Garden Edinburgh (https://data.rbge.org.uk/search/herbarium), Smithsonian National Museum of Natural History (https://collections.nmnh.si.edu/search/botany), Missouri Botanical Garden (http://www.tropicos.org) and the Muséum national d'Histoire naturelle (https://science.mnhn.fr/institution/mnhn/search).
Data analysis	The list of taxonomic names obtained from public repositories was submitted to the Taxonomic Names Resolution Service (http://tnrs.iplantcollaborative.org) to resolve names. An independent contrast against a list of 13,073 accepted species names contained in Plants of the World Online for the 'New Guinea' locality helped confirm the robustness of the expert review. Other calculations and analyses were performed with R version 3.5.2 and map artwork was designed using QGIS v 2.18.9.

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 and 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

Data and materials availability: The data that support the findings of this study are available as Supplementary Tables 1-4, and Extended Data Tables 1-4.

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	We quantified the total number of native vascular plant species in the New Guinea region by combining a data-driven approach and expert knowledge. After gathering a list of names from public repositories, we resolved the names of 23,381 taxa using two contrasting approaches: 1. A 'Big data' approach typically used in macroecological studies, i.e., using the Taxonomic Name Resolution Service software (TNRS: (http://tnrs.iplantcollaborative.org). 2. An expert approach, based on the knowledge of 99 taxonomic, floristic and monographic experts of the flora of New Guinea. Experts reviewed the list of original names in their respective families of expertise and verified whether the original list of names was correctly resolved by TNRS, and included additional information about taxonomy (basionym name, basionym year), geographic range (native, endemic, distribution) and life form (tree, herb, shrub, epiphyte, palm, etc.). When experts considered that a name that was accepted by TNRS was not correct, they wrote the correct name. Similarly, when experts considered a species note to be native, they were asked to write an explanation (e.g. geographic error, taxonomic misidentification). Finally, experts also included names that had been missed from the original list. To assess the discrepancy between the TNRS and expert verification, we compared the total number of accepted species in both lists for 254 plant families using a Wilcoxon-signed rank test. Additionally, we performed an independent contrast against a list of 13,073 accepted species names contained in Plants of the World Online for the 'New Guinea' locality (POWO; http://www.plantsoftheworldonline.org, accessed December 21, 2019). POWO was launched in 2017 with an initial focus on tropical Africa, but aims to become a single point of access for authoritative information on all plant species by 2020. Accordingly, for the names in the POWO list that were missing in our checklist, experts assessed if they were incorrect (i.e., synonyms) and/or not native to the study area, and which names were correct and native. The former represent false presences in POWO; the latter represent species that experts overlooked and which were subsequently included in the final checklist.
Research sample	We focused our analysis on all the native vascular plant species occurring in the New Guinea region, which we delimt as the area spanning a latitudinal range of -0.08 to -10.66 S and a longitudinal range of 129.42 to 150.21 E.
Sampling strategy	No statistical method was used for sample size calculation. The sample size in our case reflects the data that is available from public or institutional repositories for our study area of New Guinea.
Data collection	Data collection was based on natural history collections available for the study area in: 1. Public repositories, including the Global Biodiversity Information Facility (http://www.gbif.org), Consortium of Pacific Herbaria (http://www.pacificherbaria.org), Australasian Virtual Herbarium (http://avh.chah.org.au), Kew Herbarium Catalogue (http://apps.kew.org/herbcat/), Harvard University Herbaria (https://kiki.huh.harvard.edu/databases/specimen_index.html), the Natural History Museum (https://data.nhm.ac.uk), New York Botanical Garden (http://sweetgum.nybg.org/science/collections), Royal Botanic Garden Edinburgh (https://data.rbge.org.uk/search/herbarium), Smithsonian National Museum of Natural History (https://collections.nmnh.si.edu/search/botany), Missouri Botanical Garden (http://www.tropicos.org) and the Muséum national d'Histoire naturelle (https://science.mnhn.fr/institution/mnhn/search). 2. Institutional repositories of the Naturalis Biodiversity Center, Royal Botanic Gardens Kew and University of Papua New Guinea. 3. Databases curated by taxonomists for the families Orchidaceae, Arecaceae, Araliaceae and Cyatheaceae.
Timing and spatial scale	We used all collections available through time for the study area.
Data exclusions	We pre-established different filters to restrict our analyses on vascular plants and clean scientific names: we excluded fungi, lichens, algae, bryophytes and marine species (e.g.,sea grasses); doubtful species identifications (e.g., 'cf.', 'sp. nov.', 'aff.', 'sp.') were classified to generic level; Misspelled genera were manually corrected and doubtful cases excluded; we removed all known hybrids from the analyses.
Reproducibility	As this is not an experimental study, replication was not conducted. All data used in our analysis is made available as Supplementary Tables. Comparison of names among datasets was done using Excel and the R computer programming language and the analysis code is available on request from the corresponding author.

Randomization	Our study was not experimental but based on taxonomic knowledge and on published literature of the vascular plants of New Guinea. Thus, no randomization was required.
Blinding	Investigators were not blinded during data acquisition or analysis because our study is not experimental.

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 and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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