# The number of tree species on Earth

Roberto Cazzolla Gatti[a,b,c], Peter B. Reich[d,e,f,1], Javier G. P. Gamarra[g], Tom Crowther[h], Cang Hui[i,j], Albert Morera[k,l], Jean-Francois Bastin[m], Sergio de-Miguel[k,l], Gert-Jan Nabuurs[n], Jens-Christian Svenning[o,p], Josep M. Serra-Diaz[q], Cory Merow[r], Brian Enquist[s], Maria Kamenetsky[t], Junho Lee[u], Jun Zhu[v], Jinyun Fang[w], Douglass F. Jacobs[a], Bryan Pijanowski[a], Arindam Banerjee[x], Robert A. Giaquinto[y], Giorgio Alberti[z,aa], Angelica Maria Almeyda Zambrano[bb], Esteban Alvarez-Davila[cc], Alejandro Araujo-Murakami[dd], Valerio Avitabile[ee], Gerardo A. Aymard[ff,gg], Radomir Balazy[hh], Chris Baraloto[ii], Jorcely G. Barroso[jj], Meredith L. Bastian[kk,ll], Philippe Birnbaum[mm,nn], Robert Bitariho[oo], Jan Bogaert[m], Frans Bongers[n], Olivier Bouriaud[pp], Pedro H. S. Brancalion[qq], Francis Q. Brearley[rr], Eben North Broadbent[ss], Filippo Bussotti[tt], Wendeson Castro da Silva[uu,vv], Ricardo Gomes César[qq], Goran Češljar[ww], Víctor Chama Moscoso[xx], Han Y. H. Chen[yy], Emil Cienciala[zz,aaa], Connie J. Clark[bbb], David A. Coomes[ccc], Selvadurai Dayanandan[ddd], Mathieu Decuyper[eee,fff], Laura E. Dee[ggg], Jhon Del Aguila Pasquel[hhh], Géraldine Derroire[iii], Marie Noel Kamdem Djuikouo[jjj], Tran Van Do[kkk], Jiri Dolezal[lll,mmm], Ilija Đ. Đorđević[ww], Julien Engel[nnn], Tom M. Fayle[ooo], Ted R. Feldpausch[ppp], Jonas K. Fridman[qqq], David J. Harris[rrr], Andreas Hemp[sss], Geerten Hengeveld[ttt], Bruno Herault[uuu,vvv,www], Martin Herold[xxx,yyy], Thomas Ibanez[zzz,aaaa], Andrzej M. Jagodzinski[bbbb], Bogdan Jaroszewicz[cccc], Kathryn J. Jeffery[dddd], Vivian Kvist Johannsen[eeee], Tommaso Jucker[ffff], Ahto Kangur[gggg], Victor N. Karminov[hhhh], Kuswata Kartawinata[iiii,jjjj], Deborah K. Kennard[kkkk], Sebastian Kepfer-Rojas[llll], Gunnar Keppel[mmmm], Mohammed Latif Khan[nnnn], Pramod Kumar Khare[oooo], Timothy J. Kileen[pppp], Hyun Seok Kim[qqqq,rrrr,ssss,tttt], Henn Korjus[gggg], Amit Kumar[uuuu], Ashwani Kumar[vvvv], Diana Laarmann[gggg], Nicolas Labrière[wwww], Mait Lang[gggg,xxxx], Simon L. Lewis[yyyy,zzzz], Natalia Lukina[hhhh], Brian S. Maitner[aaaaa], Yadvinder Malhi[bbbbb], Andrew R. Marshall[ccccc,ddddd], Olga V. Martynenko[eeeee], Abel L. Monteagudo Mendoza[fffff], Petr V. Ontikov[ggggg], Edgar Ortiz-Malavasi[hhhhh], Nadir C. Pallqui Camacho[iiiii], Alain Paquette[jjjjj], Minjee Park[a], Narayanaswamy Parthasarathy[kkkkk], Pablo Luis Peri[lllll], Pascal Petronelli[mmmmm], Sebastian Pfautsch[nnnnn], Oliver L. Phillips[yyyy], Nicolas Picard[g,ooooo], Daniel Piotto[ppppp], Lourens Poorter[n], John R. Poulsen[bbb], Hans Pretzsch[qqqqq], Hirma Ramírez-Angulo[rrrrr], Zorayda Restrepo Correa[sssss], Mirco Rodeghiero[ttttt,uuuuu], Rocío Del Pilar Rojas Gonzáles[vvvvv], Samir G. Rolim[wwwww], Francesco Rovero[xxxxx,yyyyy], Ervan Rutishauser[zzzzz], Purabi Saikia[aaaaaa], Christian Salas-Eljatib[bbbbbb,cccccc,dddddd], Dmitry Schepaschenko[eeeeee,ffffff], Michael Scherer-Lorenzen[gggggg], Vladimír Šebeň[hhhhhh], Marcos Silveira[iiiiii], Ferry Slik[jjjjjj], Bonaventure Sonké[kkkkkk], Alexandre F. Souza[llll], Krzysztof Jan Stereńczak[mmmmmm], Miroslav Svoboda[nnnnnn], Hermann Taedoumg[oooooo,pppppp], Nadja Tchebakova[eeeeee], John Terborgh[qqqqqq,rrrrrr], Elena Tikhonova[hhhh], Armando Torres-Lezama[ssssss], Fons van der Plas[tttttt], Rodolfo Vásquez[vvvvv], Helder Viana[uuuuuu,vvvvvv], Alexander C. Vibrans[wwwwww], Emilio Vilanova[xxxxxx], Vincent A. Vos[yyyyyy], Hua-Feng Wang[zzzzzz], Bertil Westerlund[aaaaaaa], Lee J. T. White[bbbbbbb,ccccccc,ddddddd], Susan K. Wiser[eeeeeee], Tomasz Zawiła-Niedźwiecki[fffffff], Lise Zemagho[kkkkkk], Zhi-Xin Zhu[gggggggg], Irié C. Zo-Bi[hhhhhhh], and Jingjing Liang[a,1]

Contributed by Peter B. Reich; received September 19, 2021; accepted December 7, 2021; reviewed by Robert Colwell and Fangliang He

**One of the most fundamental questions in ecology is how many species inhabit the Earth. However, due to massive logistical and financial challenges and taxonomic difficulties connected to the species concept definition, the global numbers of species, including those of important and well-studied life forms such as trees, still remain largely unknown. Here, based on global ground-sourced data, we estimate the total tree species richness at global, continental, and biome levels. Our results indicate that there are ∼73,000 tree species globally, among which ∼9,000 tree species are yet to be discovered. Roughly 40% of undiscovered tree species are in South America. Moreover, almost one-third of all tree species to be discovered may be rare, with very low populations and limited spatial distribution (likely in remote tropical lowlands and mountains). These findings highlight the vulnerability of global forest biodiversity to anthropogenic changes in land use and climate, which disproportionately threaten rare species and thus, global tree richness.**

biodiversity | forests | hyperdominance | rarity | richness

In 1994, Robert May (1) provided the optimistic observation that, by 2044, we would roughly know the current number of species on Earth. Half of that time period has already lapsed, and we are still far from that goal. Even for trees, which are among the largest and most widespread organisms on the planet (2–6), provide a wealth of ecosystem services for humans (7–9), and support much of terrestrial biodiversity (10), we still lack a fundamental understanding of how many species exist on our planet (3, 4, 11–13).

A growing body of evidence highlights details and mechanisms regarding the biogeographic patterns in tree species diversity, such as the number of species increasing consistently toward equatorial regions (14–16). With a manageable number of taxa, tree species in the higher latitudes are relatively well characterized. However, if hyperdominance of a small fraction of species in the tropics (17) is a general phenomenon, it would mean that these regions generally harbor a very large number of rare species, many of which are endemic. The contribution of rare species to ecosystem services may be relevant and is a topic of active research (18, 19), but it is challenging as most remain poorly documented (20–26). Therefore, estimating the number of tree species is essential to inform, optimize, and prioritize forest conservation efforts across the globe. Knowing

Reviewers: R.C., University of Connecticut; and F.H., University of Alberta.

Competing interest statement: M.L.B. is an employee of PNAS.

This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

[1]To whom correspondence may be addressed. Email: preich@umn.edu or albeca.liang@gmail.com.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2115329119/-/DCSupplemental.

Published January 31, 2022.



Downloaded by guest on February 10, 2022

ECOLOGY

## Significance

**Tree diversity is fundamental for forest ecosystem stability and services. However, because of limited available data, estimates of tree diversity at large geographic domains still rely heavily on published lists of species descriptions that are geographically uneven in coverage. These limitations have precluded efforts to generate a global perspective. Here, based on a ground-sourced global database, we estimate the number of tree species at biome, continental, and global scales. We estimated a global tree richness (≈73,300) that is ≈14% higher than numbers known today, with most undiscovered species being rare, continentally endemic, and tropical or subtropical. These results highlight the vulnerability of global tree species diversity to anthropogenic changes.**

diversity's extents will be useful in several ways. First, it can help us to infer the evolutionary mechanisms that have generated diversity, so that we can predict how those same mechanisms may play out in the future. Second, it may assist in assessment of which systems may be most resilient to global change. Third, if undetected species are mostly rare and rare species are more vulnerable to extinction risk, having a better grasp of those numbers is essential to managing for biodiversity preservation. Finally, with an understanding of total species pools, it is possible to quantify the impacts of regional conservation efforts while also improving the ability to predict extinctions, manage diversity hotspots, or collect germplasm (22, 23).

Because of the limited extent of data available, estimates of tree species diversity in large geographic domains still rely heavily on expert opinions and compiled published lists of species descriptions that are geographically uneven in coverage (24, 25). Although local specialists have been increasingly joining efforts to consolidate species lists in many domains, these limitations have precluded efforts to scale this information to generate a global perspective. Here, based on a ground-sourced global database numbering ∼64,100 species [a value similar to a prior enumeration of the total of known tree species of ∼60,000 (17)], we developed estimates of the number of tree species at biome, continental, and global scales. Specifically, by comparing species accumulation curves (SACs) of tree species across different spatial scales, we estimated the number of species that have not been recorded in the global data compilation used herein.

## Results and Discussion

**Global-, Continental-, and Biome-Level Patterns.** We compiled a comprehensive global occurrence dataset with 9,353 (100- × 100-km) grid cell samples (called "samples" or "sampling units" hereafter of ∼1°) (*Materials and Methods*) by combining an abundance-based tree species dataset (the Global Forest Biodiversity Initiative [GFBI]) (Fig. 1), based on forest plots worldwide and comprising ∼38 million trees for 28,192 species, with a large high-quality occurrence-based dataset (TREECHANGE) that includes forest plots and botanical vouchers (26) (*Materials and Methods*). It is important to note that despite the large number of grid cells, extensive data, and high mean global sample coverage (96.4%) (Table 1), sampling within grid cells in many regions of the world remains very sparse.

From this dataset, with a nonparametric estimator [Chao2 (27)], we calculated occurrence-based values of potential global and continental tree species richness (*Materials and Methods*, Fig. 2*A*, and Table 1). This estimator is sensitive to accurate quantification of the numbers of uniques and duplicates (below and *Materials and Methods*), and it is known that there are problems with false uniques in forest species richness datasets (24). These Chao2 values may thus represent an overestimate to the degree that tree species recorded in only one sampling unit have been mistakenly identified as unique. Therefore, we estimated the true number ($Chao2_{adj}$) of unique species (28) (*Materials and Methods*) by accounting for the relationships of uniques, duplicates, triplets, and quadruplets to constrain the estimated numbers of unique species and from this adjusted number, computed a more conservative estimate of global tree richness, which is ∼73,300 species ($Chao2_{adj}$). Based on the good performance of this estimator and its adjusted version reported in previous studies (9, 29–33), we considered the adjusted value ($Chao2_{adj}$) our most reasonable approximation to global tree species richness. We then derived SACs at global (Fig. 2*A*) and continental (Fig. 2*B*) scales. Moreover, for each continent, from the observed number of tree species, we also estimated the asymptotic richness at the within-continent biome-level extent (Fig. 3 and *SI Appendix*, Table S2).

At the global scale, we infer that there likely are ∼9,200 tree species yet to be discovered (Table 1), given the ∼64,000 species already encountered (3, 4, 34–37). Our estimates at continental scales (Fig. 2*B* and Table 1) show that roughly 43% of all Earth's tree species occur in South America, followed by Eurasia (22%), Africa (16%), North America (15%), and Oceania (11%). However, a lack of saturation (driven by the existence of high numbers of species uncommon in the landscape, incomplete sampling, or both), particularly in the South American accumulation curve (Fig. 2*B*), suggests that our estimates may still be incomplete accounts of continental and global tree species richness. More undiscovered species likely occur in South America than any other continent. Our findings are in general agreement with recent studies of Amazonian plant diversity, which suggested that there are many undiscovered species; moreover, different approaches to the problem arrived at different estimates of total numbers of known and unknown

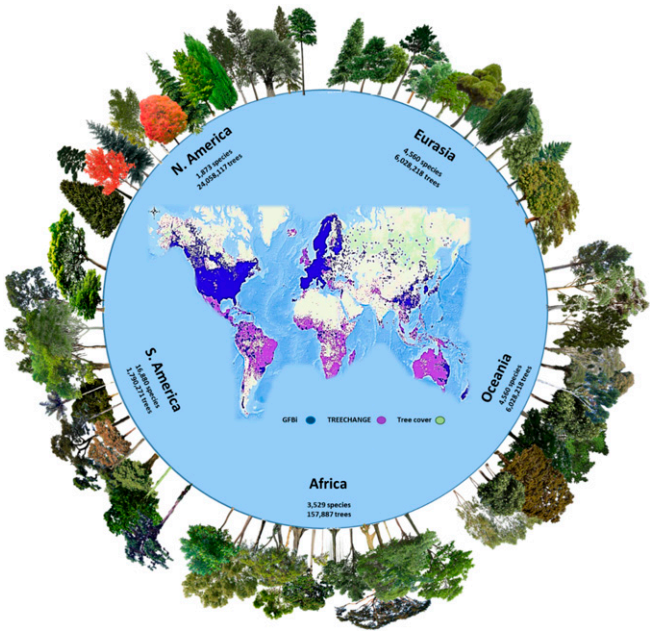


**Fig. 1.** The number of tree species and individuals per continent in the GFBI database. This dataset (blue points in the central map) was used for the parametric estimation and merged with the TREECHANGE occurrence-based data (purple points in the central map) to provide the estimates in this study. Green areas represent the global tree cover. GFBI consists of abundance-based records of ∼38 million trees for 28,192 species. Depicted here are some of the most frequent species recorded in each continent. Some GFBI and TREECHANGE points may overlap in the map.

Downloaded by guest on February 10, 2022

**Table 1. Observed, asymptotic, and adjusted tree species richness and sample coverage at continental and global scales (note that the global value is lower than the sum of the continental ones due to overlapping species among continents [Fig. 4] and due to independent estimators being run for each continent and globally)**

| Continent | No. of sampling units (~1° grid cells) | Species (observed) | Sample coverage, % | Chao2 (asymptotic) | 95% CI lower Chao2 (asymptotic) | 95% CI upper Chao2 (asymptotic) | $Chao2_{adj}$ | 95% CI lower $Chao2_{adj}$ | 95% CI upper $Chao2_{adj}$ | To be discovered (~$Chao2_{adj}$ − $S_{obs}$) | Hotspot biomes |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Global | 9,353 | 64,088 | 96.4 | 89,147 | 89,141 | 89,152 | **73,274** | 73,271 | 73,276 | **9,186** | |
| Africa | 1,575 | 10,441 | 96.0 | 14,031 | 14,028 | 14,033 | **11,875** | 11,874 | 11,877 | **1,434** | Tropical /subtropical moist and dry forests, mainly in the Congo River basin |
| Eurasia | 2,896 | 14,071 | 94.3 | 18,311 | 18,305 | 18,316 | **16,264** | 16,262 | 16,265 | **2,193** | Tropical/subtropical moist and dry forests, mainly in Southeast Asia |
| North America | 2,418 | 8,646 | 98.6 | 10,295 | 10,290 | 10,299 | **11,131** | 11,129 | 11,134 | **2,485** | Tropical subtropical moist and dry forests, mainly in Central America |
| South America | 1,461 | 27,186 | 95.0 | 46,738 | 46,729 | 46,747 | **31,112** | 31,110 | 31,115 | **3,926** | Tropical/subtropical forests, grasslands, savannas, and shrublands, mainly in the Amazon River basin and Andean high mountains |
| Oceania | 1,003 | 6,680 | 97.4 | 9,273 | 9,267 | 9,277 | **8,235** | 8,232 | 8,237 | **1,555** | Tropical/subtropical moist forests, mainly in northeast Australia and the Pacific Islands |

We also list some biomes that are hotspots of undiscovered species in each continent (*SI Appendix*, Table S2). Bold indicates the number of species to be discovered globally and continentally.

species, suggesting that, as a scientific community, we still have much more work to do to arrive at accurate estimates regionally, continentally, or globally (12, 24, 34–37). Additionally, a considerable number of species have likely not yet been encountered in each of the other four continents as well (Table 1), most likely in the species-rich and more poorly studied tropical regions within each (see below).

Our biome-level estimates of tree richness (Fig. 3 and *SI Appendix*, Table S2) provide a more detailed description of the distribution of species richness within continents and shed more light on South America's extremely high total tree diversity. As expected, the highest estimates of tree species in all continents are for the tropical and subtropical moist forest biome; for example, roughly half to two-thirds of all already known species occur in these forests on all five continents (*SI Appendix*, Table S2). Moreover, the hotspots of undiscovered species (Table 1 and *SI Appendix*, Table S2) may largely occur in these same species-rich and undersampled (37) regions. However, high numbers of known and unknown species also occur in other biomes, including tropical and subtropical dry forests, temperate forests, mangrove forests, and areas classified as nonforested biomes (e.g., lowland and montane grasslands,

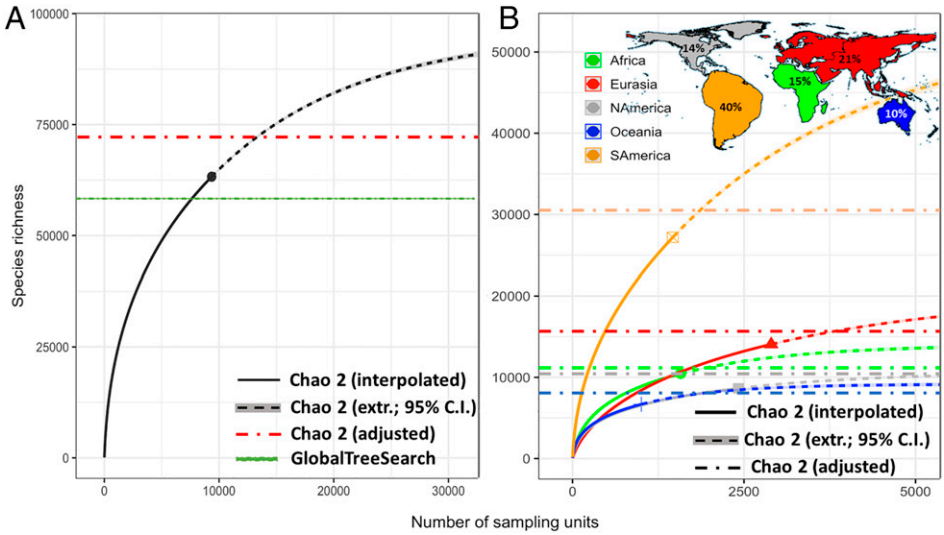


**Fig. 2.** Occurrence-based accumulation curves at global (*A*) and continental (*B*) scales. In *A*, nonparametric (interpolated) and asymptotic (extrapolated) species numbers from Chao2 (upper–lower 95% CI as shaded areas around the means; note that the CI shaded area is narrow because of the high number of sapling units), the $Chao2_{adj}$ estimate for the true number of singletons (red line) vs. the number of samples (1° grid cell ~100 × 100 km), and the number of species listed in GlobalTreeSearch (green line) are shown. In *B*, nonparametric (interpolated) and asymptotic (extr., extrapolated) estimates (upper–lower 95% CI as shaded areas around the means) and $Chao2_{adj}$ values for the true number of singletons (dashed lines) are displayed vs. the number of samples (1° grid cell ~100 × 100 km) within continents; the percentage of the global estimated richness in each continent is shown in the cartogram in *B*, *Inset* (total richness per continent is reported in Table 1).

ECOLOGY

Downloaded by guest on February 10, 2022

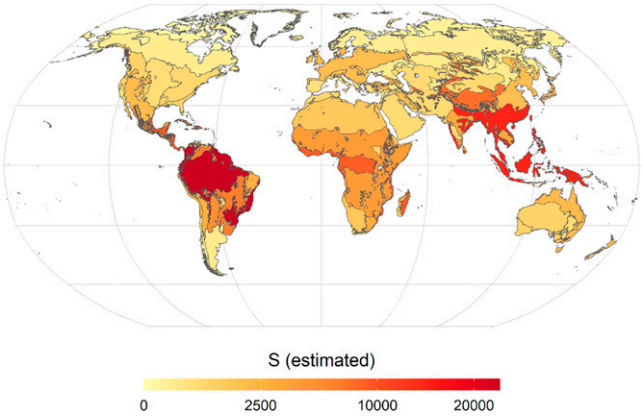


**Fig. 3.** Biome-level tree species richness estimates. The map shows the number of tree species estimated (*S* estimated from $Chao2_{adj}$) in terrestrial biomes of each continent as a color gradient from low richness (yellow) to high richness (red). More information is provided in *SI Appendix*, Table S2.

savanna, shrublands, deserts) but that include considerable areas of tree-rich, and often speciose, vegetation. The high total tree diversity in South America is dominated by the lowland wet tropics and subtropics, yet roughly one-third of all tree species on that continent are found only outside of that biome.

**Rarity in Forests Worldwide.** We also calculated indices of tree species rarity (percentages of singletons and doubletons) at continental and global scales (*SI Appendix*, Table S3) to help illuminate possible within-sample and among-sample abundance patterns. In fact, since the sample coverage deficit (1 − coverage = slope of the SAC at its right-hand end) is a statistically rigorous way of assessing the incompleteness of sampling (38), the proportion of singletons/uniques is, thus, strongly driven not only by long tails on the underlying species abundance/occurrence distribution but also, by sampling intensity/completeness.

Our most reliable abundance-based asymptotic richness estimates depend on the total number of observed species and the number of species with only one (singletons) or two (doubletons) individuals in each sample (which may represent measures of abundance-based rarity). Similarly, occurrence-based estimates depend on the number of species present in only one (unique) or two (duplicate) samples of each continent (which may represent measures of occurrence-based rarity). Rarity data within samples ($\alpha$; i.e., from the abundance-based dataset) provide an indication of the relative proportion of species that are rare at the landscape to small regional scale represented by individual grid cells (100 × 100 km). The global rarity value is 33%, with Africa (38%) and South America (37%) having the highest percentage of species rare within samples and North America (17%) and Eurasia (24%) having the lowest (*SI Appendix*, Table S3). It is important to note that our data do not mean that one-third of all species occur only once or twice in nature; instead, their rarity in our dataset suggests their rarity in nature but with unknown distributions of real occurrences. The ratio of singletons to doubletons within grid cells is higher in Africa and Oceania followed by South America and is quite low in North America and Eurasia.

From the rarity data among samples (occurrence-based rarity), we estimated that South America accounts for the highest total number of rare species (~8,200) followed by Eurasia (~6,100) and Africa (~3,900). In Eurasia and North America, the percentage of species rare among grid cells was ~43%, and it was <40% in the other continents, with the lowest value in South America (~30%) (*SI Appendix*, Table S3). The ratio of singletons to doubletons among grid cells in North America (1.83) is the highest among continents; for all other continents, it is lower than 1.5. At a global scale, percentage abundance-based rarity is higher than occurrence-based rarity, while the ratio of singletons to doubletons shows the opposite trend. Since we were aware that the numbers (and proportions) of singletons/uniques and doubletons/duplicates (and their relative magnitudes) are very much a function not only of true rarity but also of sampling effort, in relation to true richness, we estimated all indices adjusting them for "true singletons/uniques" (*Materials and Methods*). However, our findings still confirm that most forests are likely to be dominated by just a few tree species (17) and include a long tail of rare species, which represents a consistent 30 to 40% of the overall tree richness in all continents. Although more species-rich regions (such as South America and Africa) have higher abundance-based rarity, North America and Eurasia (which contain more of a mix of biomes) showed higher occurrence-based rarity, and this finding could provide insights to better understand the biogeography of tree species on Earth.

Overall, almost a third of global tree richness on Earth is made up of rare species, which appear only once or twice in our samples. Thus, if the global forest system is dominated by a relatively modest number of abundant tree species, the global number of tree species strongly depends on those rarely detected (~35%) (*SI Appendix*, Table S3) and undetected species (some large fraction of the ~9,200 unobserved over the ~73,300 estimated) (Table 1) (34). These results highlight the vulnerability of global forest biodiversity to anthropogenic changes, particularly land use and climate, because the survival of rare taxa is disproportionately threatened by these pressures (16–19). The higher threats for rare species are an important concern if we consider that their functions in ecosystems, the services they provide, and the ecoevolutionary patterns of these hyperrare tree species are still poorly known (16–20, 25).

**Comparisons across Continents.** To better understand the biogeography of richness patterns across land masses, we also estimated species turnover among continents (Fig. 4). Specifically, we combined the data of the five continents to obtain the values of estimated tree species richness in all 31 possible intersections (*Materials and Methods*). The two continents that share the highest estimated numbers of tree species are North and South America (Fig. 4), which is not surprising since these continents are interconnected by land (since about 3 Mya) in a region where nearby species-rich tropical forests occur on both continents. Consistent with this pattern, the second-highest number of shared species is between Eurasia and Oceania (Fig. 4), which had a geological continuity through the Southeast Asian archipelago that is another hotspot of tree diversity. Overall, other than the highest number of rare species, South America also shows the highest estimated percentage (49%) of continental endemic species (Fig. 4), while Eurasia and Africa account together for almost another 32% of unique tree species in the world. The percentage of shared species estimated among all five continents is lower than 0.1 (Fig. 4).

To summarize our main findings, we estimated that the absolute number of tree species on Earth is considerably higher than previously reported, with 14.3% more species than currently known to science (3). By establishing a quantitative benchmark, this information could contribute to tree and forest conservation efforts and the future discovery of new trees and associated species in certain parts of the world. For instance, considering that we estimated that about 31,100 tree species are expected in South America ($Chao2_{adj}$ estimator) and those known to science are about 27,200 (Table 1), there might be about 3,900 tree species yet to be discovered in this continent, and most of them could be endemic (Fig. 4) and located in

Downloaded by guest on February 10, 2022

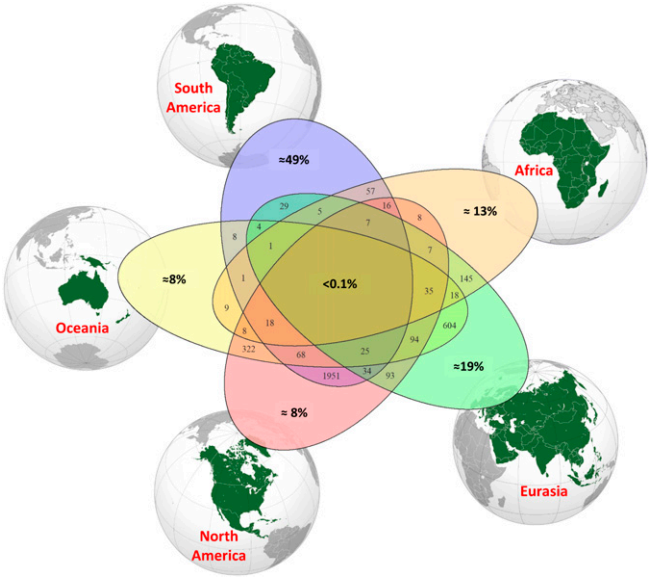


**Fig. 4.** Species richness partitioning among continents. Estimates of the percentage of continental endemic (bold percentage values close to each continental map are based on the $Chao2_{adj}$ estimator) (*Materials and Methods*) relative to the estimated richness per continent and shared species among continents (numbers in overlapping sets). In the center (bold percentage values at the intersection of all sets), the percentage of shared species among all five continents is shown.

diversity hotspots of the Amazon basin and the Andes–Amazon interface. This makes forest conservation of paramount priority in South America, especially considering the current tropical forest crisis from anthropogenic impacts such as deforestation, fires, and climate change. Similar arguments can be made about the prioritization of conservation of tropical and subtropical forests on other continents given the considerable numbers of likely undiscovered species on each and their likely rarity. For example, there are likely high numbers of undiscovered species in Central America and in Southeast Asia.

This study accelerates our science by estimating global tree richness with a more extensive dataset and more advanced statistical methods than previous attempts. However, both the underlying data and Chao richness estimators and adjustments are imperfect. We recognize several methodological issues that might have potentially biased our estimates and/or contributed to uncertainty. The first involves the uneven and unrepresentative distribution of the sampling areas in the globe and within continents, which is an issue despite the high–sample coverage metrics that we used. The second involves the possibility that some species might have been misclassified due to misidentification, failure to update taxonomic name changes, and misspellings, which could reduce accuracy in estimates of species numbers (24, 33, 34). There is compelling evidence of errors in most biodiversity datasets due to the inclusion of false uniques (24). For example, if two botanists in different parts of the same forest region encounter the same species of rare and unfamiliar tree, they may identify it differently or use different synonyms to identify it, biasing the count of uniques and the estimators. Therefore, because of the likely discrepancy between the actual proportion of uniques in a sample and the observed unique count included in our datasets, we estimated the true number of unique species (28) and from this adjusted number, computed and focus on a more conservative estimate of global tree richness, which is ∼73,300 species ($Chao2_{adj}$). There is also uncertainty about the accuracy of nonparametric estimators. Previous studies report that nonparametric estimators give lower values of tree species richness than parametric ones for the Amazon basin (34–36). However, our nonparametric estimate of tree species diversity in South American tropical forest biomes was higher than both parametric estimation and previous estimates in the Amazon (36). This might have resulted from previous studies being mainly based on Amazon lowlands, ignoring highlands. Thus, we examined sample completeness comparing continents but limiting their latitude to 23°N and S (tropical regions) (*SI Appendix*, Table S4). Results generally showed similar sample coverage at the grid-scale size used.

Future estimates of tree species richness in tropical, subtropical, and montane areas on all continents will be more accurate if an increased sample size is obtained (37), especially from areas poorly investigated. This begs the question on why South America alone could harbor >40% of all tree species. Compared with forest ecosystems on other continents, South America could have offered a larger continuous tropical forest area, a higher rate of speciation, a more robust mechanism of biodiversity maintenance, and reduced extinction rates [for instance, mild climates and the shortest period of human disturbance (39, 40)]. We also noticed that the SAC of South America continued to rise along the samples, whereas those of other continents start to level off, supporting the idea that undiscovered species numbers are likely high there, including in the Andean forests between 1,000- and 3,500-m altitude. A key challenge now is to install more plots in the Amazon-Andean transition zones, and to identify and monitor the trees within these plots.

Overall, our study points toward an estimated global tree richness (∼73,300) that is roughly 14% higher than numbers known today (3, 4), with many unknown species belonging to the tail of rare ones and often endemic to certain regions all across the globe. These results highlight the vulnerability of global tree species diversity to anthropogenic land use changes and to future climate (16–18). Losing regions of forest that contain these rare species will have direct and potentially long-lasting impacts on the global species diversity and their provisioning of ecosystem services (18–20). These results demonstrate both the lack of knowledge we still have about the tree species within our global forest systems and the value of approaches to help fill those gaps, which will be useful in providing fundamental insights about the diversity of life on our planet and its needed conservation.

## Materials and Methods

**Dataset and Sample Coverage.** We used the tree definition agreed on by IUCN's (International Union for Conservation of Nature) Global Tree Specialist Group (GTSG): "a woody plant with usually a single stem growing to a height of at least two meters, or if multi-stemmed, then at least one vertical stem five centimeters in diameter at breast height." A tree inventory abundance dataset from 105,749 forest plots, ∼38 million stems of 28,192 species, distributed across all five continents was compiled from the GFBI (https://gfbinitiative.net/) database. For the Tonga and Niue data in the GFBI dataset, the original source was the New Zealand National Vegetation Survey Databank. For the estimation of the total number of tree species worldwide, we further compiled an independent occurrence dataset that we combined with the GFBI data. The occurrence-based dataset (hereafter, TREECHANGE) consists of taxonomy and location of >6 million tree individuals. Being a major data infrastructure itself, this dataset represents species occurrence information and encompasses a huge variety of data—from ground-sourced forest plot data (similar to the GFBI). Supported by a large body of collaborating institutions all over the world, this dataset features extensive global coverage and has been used across many large-scale studies (26). A limitation of the TREECHANGE dataset is that its underlying datasets do not have a coherent and consistent design and sampling scheme, but as described below, it complements the calculation of the estimated total number of tree species worldwide based on GFBI data. We extracted taxonomic data and associated geographic coordinates from five main data aggregators of species occurrences: the Global Biodiversity Information Facility [accessed through rgbif R package (41)], the public domain of the

ECOLOGY

Downloaded by guest on February 10, 2022

Botanical Information and Ecological Network v.3 [accessed through the BIEN R package (42)], the Latin American Seasonally Dry Tropical Forest Floristic Network [DRYFLOR (43)], the RAINBIO database (44), and the Atlas of Living Australia [ALA; accessed through the ALA4 R package (45)]. The species list was initially extracted from a world tree species checklist [GlobalTreeSearch (46)]. We checked for taxonomic correctness using the Taxonomic Name Resolution online tool (47), following a quality assessment and control of the data using the workflow outlined in ref. 26. This workflow minimized common errors associated with occurrence data (43). GlobalTreeSearch uses the tree definition agreed on by IUCN's GTSG above mentioned.

For abundance-based analyses, we used the GFBI tree species dataset (at its original plot size), whose samples cover a total area of more than 73,000 ha (*SI Appendix*, Table S1). Then, to perform occurrence-based estimations, we compiled a larger and more comprehensive global dataset with 100- × 100-km sampling units (~1° grid cells) by combining the abundance-based data in the GFBI tree species dataset, which were converted in presence/absence occurrence data and pooled with the high-quality large occurrence-based TREECHANGE dataset. Globally, this yielded a dataset of 9,353 sampling units, with 696,063 occurrences. At the continental level, the combination of the two datasets to obtain a large occurrence-based dataset also yielded a number of sampling units somewhat comparable, in the sense of being a similar order of magnitude (Africa: 1,575; Eurasia: 2,896; North America: 2,418; South America: 1,461; Oceania: 1,003).

To ensure that our estimations of species richness were not biased by differences in sample coverage (e.g., an estimate of the total probability of occurrence of the species observed in the sample, taking into account species present but not detected) among continents, we estimated the inventory completeness (as defined by ref. 48) for the complete database and for each continent separately using the Chao–Shen sample coverage estimator (38, 48), which is a bias-reduced estimator of sample completeness:

$$C_n = 1 - \frac{f_1}{n}\left[\frac{(n-1)f_1}{(n-1)f_1 + 2f_2}\right],$$

where $f_1$ and $f_2$ are the numbers of singletons and doubletons (for abundance-based data) or the species occurred in only one (uniques) and in two (duplicates) 100- × 100-km (~1°) samples (for occurrence-based data), respectively; $n$ is the total number of individuals (for abundance-based data) or occurrences (for occurrence-based data) in the sample; and $C_n$ is the proportion of the total number of individuals (for abundance-based data) or occurrences (for occurrence-based data) in an assemblage (observed and not observed) that belong to the species represented in the sample (49, 50).

Because estimates of species richness can be strongly dependent on differences in inventory completeness, we checked whether sample coverage was similar in all five continents. Since all continents showed a similar proportion of sample coverage (all >94%), both from occurrence- (Table 1) and abundance-based data (*SI Appendix*, Table S1), we confirmed that our global estimate—based on global sample coverage of 96.44% (occurrence data) and 99.97% (abundance data)—was not disproportionately influenced by any specific continent. However, the slightly lower occurrence-based sample coverage of South America and Eurasia, with 95 and 94.26%, respectively, and the clustered distribution of some plots could explain the nonsaturating trend of their accumulation curves compared with the other continents (Fig. 2*B*). We also note that sample completeness at finer scales would be lower in all continents.

We selected the continental scale for our estimates, together with the common study frames of biomes (51), because nonparametric species richness estimators perform better when samples are collected in a continuous incremental area without relevant landmass separation such as oceans (31–33).

For instance, working at a global biome-level only would ensure that the current climatic conditions are similar, but this approach to estimate species richness, taken alone, would reduce the information implied in the estimates because they would be affected by several factors. 1) Within each across-ocean biome, there are still important ecological and evolutionary differences that would affect the estimates at the global biome level [in fact, conventional levels of ecological hierarchical organization are not scale dependent (52), whereas species richness estimates are]. 2) With nonparametric estimates based on SACs, it is better to ensure a continuity of sampled areas (e.g., continuous terrestrial lands) (53, 54). 3) The ecological conditions that have shaped the evolutionary patterns (phylogeny and diversity) of tree species on Earth were much different when continents were conglomerate in Pangea (55) and then slowly shifted away (i.e., during this long geological time, biomes were much different to current ones) (56–58).

Therefore, other than estimating global tree species richness at a global biome level (Fig. 3 and *SI Appendix*, Table S2), we analyzed continental richness to also account for evolutionary changes in response to the biome main variables (latitude, climate, solar radiation, etc.), which shaped current tree diversity. Adding the figures at a continental (and a continental biome) level, we ensured that our estimates are based on the 135-My biogeographical and temporal continuity of the five main vegetated landmasses, which is an implied assumption of the estimators. This approach also allows a better discussion of the results for species turnover among continents (Fig. 4), which might be a result of their connections in Laurasia and Gondwana and the following continental drift.

**Species Richness Estimators.** We initially computed a parametric estimate of species richness on the abundance-based data for 28,192 species from the GFBI dataset (*SI Appendix*, Table S1). In particular, we considered the Fisher's $\alpha$ for abundance data (calculated from http://groundvegetationdb-web.com/ground_veg/home/diversity_index).

We found that the abundance-based Fisher's $\alpha$ underestimated the absolute species richness because our global (*SI Appendix*, Table S1) Fisher estimate was close but lower than the observed number of species in our occurrence-based dataset (64,100 from GFBI + TREECHANGE). Because this parametric estimator assumes a log-series distribution of abundances (59), we performed a goodness-of-fit test and evaluated it with a Kolmogorov–Smirnov test of whether our global and abundance data fit a log-series distribution. Since all datasets (global: $D = 0.1$, $P = 1$; Africa: $D = 0.1$, $P = 1$; Eurasia: $D = 0.5$, $P = 0.17$; North America: $D = 0.2$, $P = 0.99$; Oceania: $D = 0.2$, $P = 0.99$; South America: $D = 0.2$, $P = 0.99$) follow a log-series distribution, we calculated the $\alpha$-values. At a global level, we obtained a Fisher's $\alpha$-value of 3,040 (*SI Appendix*, Table S1).

We used this value and the most recent estimates on the global number of trees by Crowther et al. (60) to estimate the global number of species from Fisher's classical equation (61):

$$S = \alpha \ln\left(1 + \frac{N}{\alpha}\right),$$

where $N$ is the total number of trees and $\alpha$ is the Fisher's $\alpha$-parameter. This yielded an estimate of 62,624 to 62,915 species (lower–upper bootstrap 95% CI) from the $3.04 \pm 0.19 \times 10^{12}$ (±95% CI) global tree stems calculated by Crowther et al. (60). Although Fisher's parametric approach stands on the very strict assumption of infinite log-series species abundance distributions, giving rise to overestimation of hyperrarity (62, 63), it estimated slightly less than the observed number of species in our occurrence-based dataset (using the $\alpha$-value derived from our abundance-based dataset). We thus did not further employ this estimate. Instead, with the larger occurrence dataset composed of GFBI (converted to presence/absence) and TREECHANGE data, we then calculated the Chao2 index, which is a lower-bound estimator and considered one of the most reliable and less affected by bias among all nonparametric indices (27, 64–66). The values of the estimators from the samples to plot the curves shown in Fig. 2 were randomized, interpolated, and extrapolated with the package iNext in R (67).

The Chao2 estimator (bias corrected) is calculated by the following formula:

$$Chao2 = S_{obs} + \left(\frac{m-1}{m}\right)\left(\frac{Q_1(Q_1-1)}{2(Q_2+1)}\right),$$

where $S_{obs}$ are the actual numbers of species observed in the samples ($m$) and $Q_1$ and $Q_2$ are the species that appear in only one (unique) and two (duplicate) sampling units, respectively (27, 29). The 95% CI (CI bias corrected) of this index can be calculated by the formula

$$\text{Lower } 95\% \text{ Bound} = S_{obs} + T/K; \text{ Upper } 95\% \text{ Bound} = S_{obs} + TK,$$

$$\text{where } T = Chao2 - S_{obs} \text{ and } K = \exp\left\{1.96\left[\log\left(1 + \frac{\hat{var}(\hat{S}_{Chao2})}{T^2}\right)\right]^{1/2}\right\}.$$

This estimation yielded a global value of 89,147 ± 1,101.5 species (Chao2 ± 95% CI) (Table 1). We are well aware that some studies provide different preferred estimators (68–70). However, many analyses, including simulation-based experiments, encourage the use of Chao2 to minimize bias (a summary is in ref. 71). This is the reason we considered the Chao2 index (based on occurrence data) our more useful estimator. Nonetheless, this estimator is sensitive to accurate quantification of the numbers of uniques and duplicates, and it is known that there are problems with false uniques in forest species richness datasets (24). Our Chao2 values may, thus, represent an overestimate to the degree that tree species recorded in only one sampling unit have been mistakenly identified as unique. Therefore, to check the reliability of our nonparametric estimates, we calculated the true number of uniques ($\hat{Q}_1$) (28) in each continent and at a global scale to understand whether our values were influenced by the number of "falsely unique species." This estimation of the true number of uniques is calculated with the formula adapted from ref. 28 for

Downloaded by guest on February 10, 2022

incidence-based data:

$$\hat{Q}_1 = \left(\frac{T-1}{T}\right)\frac{2Q_2^2}{3Q_3} + \left(\frac{T-1}{T}\right)2Q_2\left(\frac{Q_2}{2Q_3} - \frac{Q_3}{4Q_4}\right),$$

where $\hat{Q}_1$ is the estimated true number of uniques; $T$ is the number of sampling units (map cells); and $Q_2$, $Q_3$, and $Q_4$ are observed duplicates, triplicates, and quadruplicates.

At the global level, the estimate of the true number of uniques is 13,162 compared with the observed 24,768. At the continental level, the number of estimated uniques was much lower than the observed one in South America (4,888 vs. 13,110) and somewhat lower in Eurasia (3,424 vs. 5,806), Africa (2,192 vs. 3,466), and Oceania (1,444 vs. 2,208), but it was slightly higher in North America (2,460 vs. 2,360). We then used the adjusted number of uniques in the Chao2 equation (see above) to calculate the $\text{Chao2}_{\text{adj}}$ estimates, $\hat{S}_{adjChao2}$ (27–29).

We also calculated tree species rarity at continental and global scales for abundance (abundance-based rarity; i.e., based on the number of adjusted singletons [*S1*] and doubletons [*S2*]) and occurrence (occurrence-based rarity; i.e., based on the adjusted number of unique species and the number of duplicate ones). We defined the number of rare species as the sum of adjusted singletons and doubletons. We also computed an index of rarity importance using our occurrence-based dataset as the proportion of rare species over total richness and an $S1_{adjusted}/S2$ ratio, which is the proportion of singletons over doubletons.

**Continental Biodiversity Partitioning.** We estimated the number of species shared among continents and unique to each continent using the Chao2 estimator (Fig. 4) from the occurrence-based data, and we represented them in a Venn diagram. We combined the observations of species richness for the five continents ($n = 5$) in all possible $2^5 - 1 = 31$ combinations.

First, we calculated asymptotic species richness (Chao2) from occurrences observed in each continent; then, we intersected (creating a unique presence/absence binary entry for each species) the observed occurrences per each pair, triplet, quadruplet, and all five of continents (obtaining the occurrences of all the observed species in each combination of continents) and calculated the asymptotic species richness (Chao2) per each pair, triplet, quadruplet, and quintuplet continents. Therefore, a total of 31 estimates were obtained by the Chao2 index and plotted in a Venn diagram with the R package VennDiagram (72).

Additional cross-checks of the data pooling approach and *q1*/*q2* relationship are in *SI Appendix*, *SI Methods*.

**Data Availability.** Anonymized numeric data have been deposited in GFBI, https://gfbinitiative.net/data/.

**ACKNOWLEDGMENTS.** We thank the following agencies, initiatives, teams, and individuals for data collection and other technical support: the GFBI for establishing the data standards and collaborative framework; the US Department of Agriculture, Forest Service, Forest Inventory and Analysis Program; the University of Alaska Fairbanks; the SODEFOR, Ivory Coast; University Félix Houphouët-Boigny (Ivory Coast); the Queensland Herbarium; and past Queensland Government Forestry and Natural Resource Management Departments and staff for data collection for over seven decades. We thank Javier Eduardo Silva Espejo, Rahman Laskar, Salam Dilip, Bijit, Bironjoy, and Samar; Badru Mugerwa and Emmanuel Akampurira together with a team of field assistants (Valentine and Lawrence); all persons who made the Third Spanish Forest Inventory possible, especially the main coordinator J. A. Villanueva (IFN3); the French National Forest Inventory (NFI; NFI campaigns [raw data 2005 and following annual surveys] were downloaded by the GFBI at https://inventaire-forestier.ign.fr/dataIFN/, site accessed on 1 January 2015); the Italian Forest Inventory (NFI campaigns raw data 2005 and following surveys were downloaded by the GFBI at https://www.inventarioforestale.org/, site accessed on 27 April 2019); Swiss National Forest Inventory, Swiss Federal Research Institute WSL, and Federal Office for the Environment FOEN, Switzerland; Coordination for the Improvement of Higher Education Personnel of Brazil (CAPES) Grant 88881.064976/2014-01; Rafael Ávila and Sharon van Tuylen, Instituto Nacional de Bosques, Guatemala for facilitating Guatemalan data; the National Focal Center for Forest condition monitoring of Serbia, Institute of Forestry, Belgrade, Serbia; the Thünen Institute of Forest Ecosystems (Germany) for providing NFI data; the Food and Agriculture Organization of the United Nations and the United Nations High Commissioner for Refugees undertaking the Safe Access to Fuel and Energy project; Russian Science Foundation Project 21-46-07002 for the plot data collected in the Krasnoyarsk region; and the Amazon Forest Inventory Network (RAINFOR), the African Tropical Rainforest Observation Network, and the ForestPlots.net initiative for their contributions from Amazonian and African forests. The Natural Forest plot data were collected between January 2009 and March 2014 by the LUCAS Programme for the New Zealand Ministry for the Environment, as provided by the New Zealand National Vegetation Survey Databank. All Tropical Ecology Assessment and Monitoring (TEAM) data are provided by the TEAM Network, a collaboration between Conservation International, the Missouri Botanical Garden, the Smithsonian Institution, and the Wildlife Conservation Society, and partially funded by these institutions, the Gordon and Betty Moore Foundation, and other donors, with thanks to all current and previous TEAM site managers and other collaborators who helped with collecting data; the people of the Redidoti, Pierrekondre, and Cassipora villages who were instrumental in assisting with the collection of data and sharing local knowledge of their forest; and the dedicated members of the field crew of the Kabo 2012 census. This research was supported in part through computational resources provided by Information Technology at Purdue, West Lafayette, Indiana. We also thank Robert K. Colwell and Fangliang He for insightful comments and recommendations during the revisions of this manuscript. This work is supported in part by US Department of Agriculture National Institute of Food and Agriculture McIntire Stennis Projects 1017711 and 1016676; a faculty start-up grant from the Department of Forestry and Natural Resources, Purdue University; US NSF Biology Integration Institutes Program NSF-DBI-2021898; Key Project of National Key Research and Development Plan, China Grant 2017YFC0504005; São Paulo Research Foundation Grant 2014/14503-7; Proyecto FONACIT Grant 1998003436 and UNELLEZ Grant 23198105; EU, Sumforest–REFORM, Risk Resilient Forest Management, FKZ Grant 2816ERA02S; German Research Foundation, KROOF Tree and stand-level growth reactions on drought in mixed versus pure forests of Norway spruce and European beech Grant PR 292/12-1; Bavarian State Ministry for Food, Agriculture and Forestry, W07 long-term yield experiments, Grant 7831-26625-2017 and Project E33; the Deutsche Forschungsgemeinschaft Priority Program 1374 Biodiversity Exploratories; the International Tropical Timber Organization Project PD 53/00 Rev.3 (F); the State Forest Management Centre, Estonia, and the Environmental Investment Centre, Estonia; Natural Sciences and Engineering Research Council of Canada Discover Grant Project Grants RGPIN-2014-04181 and STPGP428641; European Structural Funds by FEDER 2014 to 2020 Grant GY0006894; European Investment Funds by FEDER/COMPETE/POCI-Operacional Competitiveness and Internacionalization Programme under Project POCI-01-0145-FEDER-006958 and National Funds by FCT–Portuguese Foundation for Science and Technology Project UIDB/04033/2020. ICNF-Instituto de Conservação da Natureza e Florestas. 6° Inventário Florestal Nacional; Chilean research Grants Fondecyt No. 1191816 and FONDEF No. ID19 10421; Vietnam National Foundation for Science and Technology Development Grant NAFOSTED-106-NN.06-2016.10; German Research Foundation Grant FOR 1246; the Project LIFE+ ForBioSensing PL "Comprehensive monitoring of stand dynamics in Białowieża Forest supported with remote sensing techniques" cofunded by Life Plus Contract LIFE13 ENV/PL/000048 and National Fund for Environmental Protection and Water Management in Poland Contract 485/2014/WN10/OP-NM-LF/D; National Natural Scientific Foundation of China Grants 31660055 and 31660074; the Polish State Forests National Forest Holding (2016); National Science Center (Poland) Grant 2011/02/A/NZ9/00108; the Dutch Ministry of Economic Affairs for funding the Dutch National Forest Inventory; US NASA Grant 11-TE11-0100; the TEAM/Conservation International project for funding the data collection and the Instituto Nacional de Pesquisas da Amazônia; the Ministère des Forêts, de la Faune et des Parcs du Québec (Canada); the exploratory plots of FunDivEUROPE received funding from European Union Seventh Framework Programme FP7/2007-2013 under Grant 265171; DBT, Government of India through the project "Mapping and quantitative assessment of geographic distribution and population status of plant resources of Eastern Himalayan region" (Sanction Order BT/PR7928/NDB/52/9/2006 dated 29 September 2006); financial support from the Natural Sciences and Engineering Research Council of Canada (to S.D.); Czech Science Foundation Standard Grant 19-14620S and European Research Council Advanced Grant 669609; RFBR Grant 16-05-00496; the project implementation demonstration object on the transformation of declining spruce forests into ecologically more stable multifunctional ecosystems, ITMS Grant 26220220026, supported by the Research & Development Operational Program funded by the ERDF; the Swedish NFI, Department of Forest Resource Management, Swedish University of Agricultural Sciences SLU; National Research Foundation of South Africa Grants 89967 and 109244 and the South African Research Chair Initiative; the University Research Committee of the University of the South Pacific and New Colombo Plan funding through the Department of Foreign Affairs and Trade of the Australian government; the TEAM project in Uganda supported by the Moore Foundation and the Buffett Foundation through Conservation International and Wildlife Conservation Society; the COBIMFO project funded by Belgian Science Policy Office Contract SD/AR/01A; the German Federal Ministry of Education and Research Grant FKZ 01LL0908AD for the project Land Use and Climate Change Interactions in the Vu Gia Thu Bon River basin, Central Vietnam; Programme Tropenbos Côte d'Ivoire Project 04/97-1111a du Complément d'Inventaire de la Flore dans le Parc National de Taï; Danish Council for Independent Research | Natural Sciences TREECHANGE Grant 6108-00078B (to J.-C.S.)and VILLUM FONDEN Grant 16549 (to J.-C.S.); ERC Advanced Grant 291585 ("T-FORCES") and a Royal Society Wolfson Research Merit Award (to O.L.P.); RAINFOR plots supported by the Gordon and Betty Moore Foundation and the UK Natural Environment Research Council (NERC), notably NERC Consortium Grants AMAZONICA (NE/F005806/1), TROBIT (NE/D005590/1), and BIO- RED (NE/N012542/1); Fundação de Amparo à Pesquisa e Inovação de Santa Catarina, FAPESC

ECOLOGY

Downloaded by guest on February 10, 2022

Grant 2016TR2524, Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq Grant 312075/2013-8; "Investissement d'Avenir" grant managed by Agence Nationale de la Recherche (CEBA, reference ANR-10-LABX-25-01); CIFOR's Global Comparative Study on REDD+ funded by the Norwegian Agency for Development Cooperation (Norad), the Australian Department of Foreign Affairs and Trade, the European Union, the International Climate Initiative of the German Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety, and the CGIAR Research Program on Forests, Trees and Agroforestry, and donors to the CGIAR Fund; the Nature and Biodiversity Conservation Union under the project entitled "Biodiversity under Climate Change: Community Based Conservation, Management and Development Concepts for the Wild Coffee Forests" funded by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety through the International Climate Initiative; the Conselho Nacional de Desenvolvimento Científico e Tecnológico; the institutional project "EXTEMIT - K" CZ.02.1.01/0.0/0.0/15_003/0000433 financed by OP RDE; EC DG VIII Grants BZ-5041 (ECOSYN), NWO-WOTRO (W84-204), and GTZ; AfriTRON network plots funded by the local communities and NERC, ERC, the European Union, the Royal Society, and Leverhume Trust; BOLFOR (Proyecto de Manejo Forestal Sostenible Bolivia); the Global Environment Research Fund Grants F-071 and D-1006, JSPS KAKENHI Grant JP17K15289; the National Institute of Biology (now the Research Center for Biology), Indonesian Institute of Sciences, Indonesia IFBN Project Contract 4000114425/15/NL/FF/gp funded by ESA; NSF Grants DBI-1565046, DEB-0424767, DEB-0639393, and DEB-1147429; NASA Terrestrial Ecology Program; Swiss National Science Foundation Grants 130720 and 147092; Projects D/9170/07, D/018222/08, D/023225/09, and D/032548/10 funded by the Spanish Agency for International Development Cooperation (Agencia Española de Cooperación Internacional para el Desarrollo) and Fundación Biodiversidad in cooperation with the Universidad Mayor de San Simón, the Manejo Forestal en las Tierras Tropicales de Bolivia project, and Compañía Industrial Maderera Ltda.; the Agency for Economic and Environmental Development of the north province of New Caledonia (Projects Ecofor & Cogefor, 2011 to 2016); Russian Science Foundation Grant 16-17-10284 The accumulation of carbon in forest soils and forest succession status; the Norwegian Ministry of Food and Agriculture; a grant from the Royal Society and the Natural Environment Research Council (United Kingdom; to S.L.L.); the Spanish Agency for International Development Cooperation (Agencia Española de Cooperación Internacional para el Desarrollo) and Fundación Biodiversidad, in cooperation with the governments of Syria and Lebanon; COBIMFO Project, Federal Science Policy, Belgium; Consejo Nacional de Ciencia y Tecnología, Mexico; Comisión Nacional Forestal, Mexico; BEF-China Project FOR 891 funded by the German Research Foundation; WWF Russell Train Fellowship Grant ST54 to PMU; Wildlife Conservation Society DRC Program under CARPE Funding; the Seoul National University Big Data Institute through the Data Science Research Project 2016, R&D Program for Forest Science Technology Projects 2013069C10-1719-AA03 and S111215L020110 funded by Korea Forest Service (the Korea Forestry Promotion Institute); Department of Biotechnology, Government of India Grant BT/PR12899/NDB/39/506/2015 (dated 20 June 2017) and Science and Engineering Research Board, Government of India Grant YSS/2015/000479 (dated 12 January 2016); Tropenbos International-Suriname; the Institute for World Forestry, University of Hamburg; REMBIOFOR Project "Remote sensing based assessment of woody biomass and carbon storage in forests" funded by the National Centre for Research and Development, Warsaw, Poland, under BIOSTRATEG Program Agreement BIOSTRATEG1/267755/4/NCBR/2015; Project "Environmental and genetic factors affecting productivity of forest ecosystems on forest and post-industrial habitats" (2011 to 2015) Grant OR/2717/3/11 funded by the General Directorate of State Forests, Warsaw, Poland; Project "Carbon balance of the major forest-forming tree species in Poland" (2007 to 2011) Grant 1/07 funded by the General Directorate of State Forests, Warsaw, Poland; the research professorship for "Ecosystem-based sustainable development" funded by Eberswalde University for Sustainable Development; Pontificia Universidad Católica del Ecuador supported the fieldwork census in Yasuni National Park; the National Forest Programme of the National Institute of Agricultural Research; São Paulo Research Foundation Grants #2014/14503-7, 2017/05662-2, and 03/12595-7; MAUA group supported by FAPEAM-PRONEX Grant 1600/2006; CNPq/FAPEAM-PELD Grant 403792/2012-6; ATTO Project Grants MCTI-FINEP 1759/10 and BMBF 01LB1001A; Czech Science Foundation Standard Grants 17-07378S and 17-19376S; the Long-Term Research Development Project RVO 67985939 of Institute of Botany of the Czech Academy of Sciences; Slovak Research and Development Agency Project APVV-20-0168; the Strategic Science Investment Fund of the New Zealand Ministry for Business, Innovation and Employment; FAPESP Grants 2014/14503-7 and 2017/05662-2; and CNPq Universal Grant 479599/2008-4. The Digital Environment for Enabling Data-Driven Science Project is funded by NSF Grant CIF21 DIBBs: EI: 1724728; Natural Forest plot data collected between January 2002 and March 2007 by the LUCAS programme for the Ministry for the Environment; and Human Modified Tropical Forests Programme of NERC Grant NE/K016377/1. P.B.R. acknowledges funding support from US NSF Long-Term Ecological Research Grant DEB-1831944 and Biological Integration Institutes Grant NSF-DBI-2021898.

Author contributions: J.L. developed the original idea, coordinated tree-level data compilation with the GFBI and TREECHANGE teams, and provided facility support. R.C.G. conceptualized the study, performed the data and statistical analysis, implemented the computations, and made the figures and tables. R.C.G., J.L., and P.B.R made major contributions to data interpretation and writing the manuscript, with R.C.G. taking the lead role throughout, J.L. co-developing the initial draft version, and P.B.R. co-developing the original and final submitted versions. J.G.P.G, T.C., J-F. B., C.H. and P.B.R. contributed to the analytical methods, interpretation of the results, and the writing. A.M. and R.A.G. contributed to the R coding, computations, data analysis, and mapping. S.d-M. and A.B. supported and supervised R coding, data analysis and mapping. J.L., J.G.P.G., T.C., S.d.-M., G.-J.N., J.-C.S., J.M.S.-D., C.M., B.E., J.F., G.A., A.M.A.Z., E.A.-D., A.A.-M., V.A., G.A.A., R. Balazy, C.B., J.G.B., M.L.B., P.B., R. Bitariho, J.B., F. Bongers, O.B., P.H.S.B., F.Q.B., E.N.B., F. Bussotti, W.C.d.S., R.G.C., G.Č., V.C.M., H.Y.H.C., E.C., C.J.C., D.A.C., S.D., M.D., J.D.A.P., G.D., M.N.K.D., T.V.D., J.D., I.Đ.Đ., J.E., T.M.F., T.R.F., J.K.F., D.J.H., A.H., G.H., B.H., M.H., T.I., A.M.J., B.J., K.J.J., V.K.J., T.J., A. Kangur, V.N.K., K.K., D.K.K., S.K.-R., G.K., M.L.K., P.K.K., T.J.K., H.S.K., H.K., Amit Kumar, Ashwani Kumar, D.L., N. Labrière, M.L., S.L.L., N. Lukina, B.S.M., Y.M., A.R.M., O.V.M., A.L.M.M., P.V.O., E.O.-M., N.C.P.C., A.P., M.P., N. Parthasarathy, P.L.P., P.P., S.P., O.L.P., N. Picard, D.P., L.P., J.R.P., H.P., H.R.-A., Z.R.C., M.R., R.D.P.R.G., S.G.R., F.R., E.R., P.S., C.S.-E., D.S., M.S.-L., V.Š., M. Silveira, F.S., B.S., A.F.S., K.J.S., M. Svoboda, H.T., N.T., J.T., E.T., A.T.-L., F.v.d.P., R.V., H.V., A.C.V., E.V., V.A.V., H.-F.W., B.W., L.J.T.W., S.K.W., T.Z.-N., L.Z., Z.-X.Z., and I.C.Z.-B. contributed data; R.C.G., P.B.R., J.L., J.G.P.G., T.C., C.H., A.M., J.-F.B., S.d.-M., G.-J.N., J.-C.S., J.M.S.-D., C.M., B.E., M.K., J.Z., J.F., D.F.J., B.P., A.B., R.A.G., G.A., A.M.A.Z., E.A.-D., A.A.-M., V.A., G.A.A., R. Balazy, C.B., J.G.B., M.L.B., P.B., R. Bitariho, J.B., F. Bongers, O.B., P.H.S.B., F.Q.B., E.N.B., F. Bussotti, W.C.d.S., R.G.C., G.Č., V.C.M., H.Y.H.C., E.C., C.J.C., D.A.C., S.D., M.D., L.E.D., J.D.A.P., G.D., M.N.K.D., T.V.D., J.D., I.Đ.Đ., J.E., T.M.F., T.R.F., J.K.F., D.J.H., A.H., G.H., B.H., M.H., T.I., A.M.J., B.J., K.J.J., V.K.J., T.J., A. Kangur, V.N.K., K.K., D.K.K., S.K.-R., G.K., M.L.K., P.K.K., T.J.K., H.S.K., H.K., Amit Kumar, Ashwani Kumar, D.L., N. Labrière, M.L., S.L.L., N. Lukina, B.S.M., Y.M., A.R.M., O.V.M., A.L.M.M., P.V.O., E.O.-M., N.C.P.C., A.P., M.P., N. Parthasarathy, P.L.P., P.P., S.P., O.L.P., N. Picard, D.P., L.P., J.R.P., H.P., H.R.-A., Z.R.C., M.R., R.D.P.R.G., S.G.R., F.R., E.R., P.S., C.S.-E., D.S., M.S.-L., V.Š., M. Silveira, F.S., B.S., A.F.S., K.J.S., M. Svoboda, H.T., N.T., J.T., E.T., A.T.-L., F.v.d.P., R.V., H.V., A.C.V., E.V., V.A.V., H.-F.W., B.W., L.J.T.W., S.K.W., T.Z.-N., L.Z., Z.-X.Z., and I.C.Z.-B. wrote the paper.

[a]Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907; [b]Department of Biological, Geological, and Environmental Sciences, Alma Mater Studiorum University of Bologna, Bologna 40126, Italy; [c]Biological Institute, Tomsk State University, Tomsk 634050, Russia; [d]Department of Forest Resources, University of Minnesota, St. Paul, MN 55108; [e]Institute for Global Change Biology and School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109; [f]Hawkesbury Institute for the Environment, Western Sydney University, Penrith 2753, Australia; [g]Forestry Department, FAO, Rome 00153, Italy; [h]Institute of Integrative Biology, ETH Zurich 8092 Zurich, Switzerland; [i]Centre for Invasion Biology, Department of Mathematical Sciences, Stellenbosch University, Stellenbosch 7602, South Africa; [j]Mathematical Biology Unit, African Institute for Mathematical Sciences, Muizenberg 7945, South Africa; [k]Department of Crop and Forest Sciences, University of Lleida 25198 Lleida, Spain; [l]Joint Research Unit CTFC–AGROTECNIO–CERCA 25280 Solsona, Spain; [m]TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, University of Liege, Gembloux 5030, Belgium; [n]Forest Ecology and Forest Management Group, Wageningen University and Research, Wageningen 6700 AA, The Netherlands; [o]Center for Biodiversity Dynamics in a Changing World (BIOCHANGE), Department of Biology, Aarhus University DK-8000 Aarhus C, Denmark; [p]Section for Ecoinformatics and Biodiversity, Department of Biology, Aarhus University DK-8000 Aarhus C, Denmark; [q]AgroParisTech, INRAE, Silva, Université de Lorraine 5400 Nancy, France; [r]Department of Ecology and Evolutionary Biology, University of Connecticut, Mansfield, CT 06268; [s]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721; [t]Department of Population Health Sciences, University of Wisconsin–Madison, Madison, WI 53704; [u]Statistics Program, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia; [v]Department of Statistics, University of Wisconsin–Madison, Madison, WI 53706; [w]Institute of Botany, CAS, Peking University, Beijing 100871, China; [x]Department of Computer Science, University of Illinois at Urbana–Champaign, IL 61801; [y]Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55108; [z]Department of Agricultural, Food, Environmental and Animal Sciences, University of Udine 33100, Italy; [aa]Faculty of Science and Technology, Free University of Bolzano-Bozen, Bolzano 39100, Italy; [bb]Spatial Ecology and Conservation Lab, Department of Tourism, Hospitality, and Events Management, University of Florida, Gainesville, FL 32611; [cc]Escuela de Ciencias Ambientales, Universidad Nacional (Colombia) Abierta y a Distancia, Bogotá 2102, Colombia; [dd]Museo de Historia Natural Noel Kempff Mercado, Universidad Autónoma Gabriel Rene Moreno, Casilla 2489, Santa Cruz, Bolivia; [ee]European Commission, Joint Research Centre 21027 Ispra, Italy; [ff]Compensation International Progress S. A.–Ciprogress Greenlife, Bogotá, Colombia; [gg]UNELLEZ-Guanare, Programa de Ciencias del Agro y el Mar, Herbario Universitario (PORT), Mesa de Cavacas, estado Portuguesa 3323, Venezuela; [hh]Department of Geomatics, Forest Research Institute, Sękocin

Downloaded by guest on February 10, 2022

Stary, Raszyn 05-090, Poland; [ii]International Center for Tropical Botany, Department of Biological Sciences, Florida International University, Miami, FL 33133; [jj]Forest Science Laboratory, Multidisciplinary Center, Universidade Federal do Acre, Cruzeiro do Sul 69920-900, Brazil; [kk]Proceedings of the National Academy of Sciences, U.S.A., Washington, DC 20001; [ll]Department of Evolutionary Anthropology, Duke University, Durham, NC 27708-06080; [mm]AMAP, Université de Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier 34090, France; [nn]Institut Agronomique néo-Calédonien, Equipe Sol & Végétation 98800 Noumea, New Caledonia; [oo]Institute of Tropical Forest Conservation, Mbarara University of Science and Technology, Kabale, Uganda; [pp]Integrated Center for Research, Development and Innovation in Advanced Materials, Nanotechnologies, and Distributed Systems for Fabrication and Control, University Stefan cel Mare of Suceava 720229 Suceava, Romania; [qq]Department of Forest Sciences, "Luiz de Queiroz" College of Agriculture, University of São Paulo, Piracicaba 13418-900, Brazil; [rr]Department of Natural Sciences, Manchester Metropolitan University, Manchester M1 5GD, United Kingdom; [ss]Spatial Ecology and Conservation Lab, School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611; [tt]Department of Agriculture, Alimentation, Environment and Forestry, Università degli Studi di Firenze, Firenze 50144, Italy; [uu]Amazonia Green Landscape Protection and Governance Programme, SOS Amazônia, Rio Branco 69905-082, Brazil; [vv]Laboratory of Botany and Plant Ecology, Center of Biological and Nature Sciences, Federal University of Acre, Rio Branco 69920-900, Brazil; [ww]Department of Spatial Regulation, GIS and Forest Policy, Institute of Forestry 11030 Belgrade, Serbia; [xx]Universidad Nacional de San Antonio Abad del Cusco, Cusco 08000, Peru; [yy]Faculty of Natural Resources Management, Lakehead University, Thunder Bay, ON, Canada P7B 5E1; [zz]Institute of Forest Ecosystem Research 254 01 Jilove u Prahy, Czech Republic; [aaa]Global Change Research Institute of the Czech Academy of Sciences 603 00 Brno, Czech Republic; [bbb]Nicholas School of the Environment, Duke University, Durham, NC 27708; [ccc]Department of Plant Sciences in the Conservation Research Institute, University of Cambridge, Cambridge CB2 3EA, United Kingdom; [ddd]Quebec Center for Biodiversity Sciences, Centre for Sustainability Research, CSFG and Biology Department, Concordia University, Montreal, ON QC H3G 1M8, Canada; [eee]Laboratory for Geoinformation Science and Remote Sensing, Department of Environmental Sciences, Wageningen University and Research, Wageningen 6700 AA, The Netherlands; [fff]World Agroforestry (ICRAF), Nairobi 00100, Kenya; [ggg]Ecology and Evolutionary Biology, University of California, Santa Barbara, CA 93106; [hhh]Instituto de Investigaciones de la Amazonia Peruana, Iquitos, Peru; [iii]Cirad, UMREcoFoG (Agroparistech, CNRS, INRAE, Université des Antilles, Université de Guyane) 97 310 Kourou, French Guiana; [jjj]Department of Plant Science, Faculty of Science, University of Buea, Buea, Cameroon; [kkk]Silviculture Rearch Institute, Vietnamese Academy of Forest Sciences, Hanoi, Vietnam; [lll]Institute of Botany of the Czech Academy of Science 25243 Průhonice, Czech Republic; [mmm]Faculty of Science, University of South Bohemia 37005 České Budějovice, Czech Republic; [nnn]AMAP, IRD, CIRAD, CNRS, INRAE, Université de Montpellier, Montpellier Cedex 5 F-34398, France; [ooo]School of Biological and Behavioural Sciences, Queen Mary University of London, London E1 4NS, United Kingdom; [ppp]College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4PY, United Kingdom; [qqq]Department of Forest Resource Management, Swedish University of Agricultural Sciences, Uppsala 750 07, Sweden; [rrr]Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, United Kingdom; [sss]Department of Plant Systematics, University of Bayreuth 95440 Bayreuth, Germany; [ttt]Biometris & Forest and Nature Policy Group, Wageningen University and Research, Wageningen 6700 AA, The Netherlands; [uuu]CIRAD, UPR Forêts et Sociétés, Yamoussoukro, Côte d'Ivoire; [vvv]Forêts et Sociétés, Univ Montpellier, CIRAD, Montpellier 34000, France; [www]Institut National Polytechnique Félix Houphouët-Boigny, Yamoussoukro, Côte d'Ivoire; [xxx]Laboratory for Geoinformation Science and Remote Sensing, Department of Environmental Sciences, Wageningen University and Research, Wageningen 6700 AA, The Netherlands; [yyy]Section 1.4 Remote Sensing and Geoinformatics, Helmholtz GFZ German Research Centre for Geosciences, Telegrafenberg, Potsdam 14473, Germany; [zzz]Institut Agronomique néo-Calédonien, Equipe Sol & Végétation 98800, New Caledonia; [aaaa]Department of Biology, University of Hawai'i at Hilo, Hilo, HI 96720; [bbbb]Institute of Dendrology, Polish Academy of Sciences, Kornik 62-035, Poland; [cccc]Białowieża Geobotanical Station, Faculty of Biology, University of Warsaw, Warsaw 17-230, Poland; [dddd]Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA, United Kingdom; [eeee]Department of Geosciences and Natural Resource Management, University of Copenhagen 1165, Denmark; [ffff]School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, United Kingdom; [gggg]Institute of Forestry and Rural Engineering, Estonian University of Life Sciences, Tartu 51006, Estonia; [hhhh]Center for Forest Ecology and Productivity of the Russian Academy of Sciences, Moscow 119991, Russia; [iiii]Integrative Research Center, The Field Museum, Chicago, IL 60605; [jjjj]Herbarium Bogoriense, Research Center for Biology, Indonesian Institute of Sciences/Lembaga Ilmu Pengetahuan Indonesia, Cibinong Science Center, Cibinong 16912 Indonesia; [kkkk]Department of Physical and Environmental Sciences, Colorado Mesa University, Grand Junction, CO 81501-3122; [llll]Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen 1017, Denmark; [mmmm]UniSA STEM and Future Industries Institute, University of South Australia 5001 Adelaide, Australia; [nnnn]Department of Botany, Dr. Harisingh Gour Vishwavidyalaya (A Central University), Sagar 470003, India; [oooo]Department of Botany, Dr. Harisingh Gour Central University, Madhya Pradesh Sagar-470003, India; [pppp]Museo de Historia natural Noel Kempff Mercado, Santa Cruz, Bolivia; [qqqq]Department of Agriculture, Forestry and Bioresources, Seoul National University, Seoul 08826, South Korea; [rrrr]Interdisciplinary Program in Agricultural and Forest Meteorology, Seoul National University, Seoul 08826, South Korea; [ssss]National Center for AgroMeteorology, Seoul 08826, South Korea; [tttt]Institute of Future Environmental and Forest Resources, Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul 08826, South Korea; [uuuu]Department of Geoinformatics, Central University of Jharkhand, Ranchi 835205, Jharkhand, India; [vvvv]Department of Botany, Dr. Harisingh Gour Vishwavidyalaya (A Central University), Sagar 470003, India; [wwww]Laboratoire Évolution et Diversité Biologique, UMR 5174 (CNRS/IRD/UPS) 31062 Toulouse Cedex 9, France; [xxxx]Tartu Observatory, University of Tartu, Tartu 61602, Estonia; [yyyy]School of Geography, University of Leeds, Leeds LS2 9JT, United Kingdom; [zzzz]Department of Geography, University College London, London WC1E 6BT, United Kingdom; [aaaaa]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721; [bbbbb]Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, United Kingdom; [ccccc]Forest Research Institute, University of the Sunshine Coast, Sippy Downs QLD 4556, Australia; [ddddd]Department of Environment and Geography, University of York YO10 5NG, United Kingdom; [eeeee]All-Russian Institute of Continuous Education in Forestry 141200 Pushkino, Moscow region, Russia; [fffff]Universidad Nacional de San Antonio Abad del Cusco, Cusco, Peru; [ggggg]FSBI "ROSLESINFORG," 141208 Ivanteevka, Moscow region, Russia; [hhhhh]Escuela de Ingeniería Forestal, Instituto Tecnologico de Costa Rica, Cartago 30101, Costa Rica; [iiiii]Universidad Nacional de San Antonio Abad del Cusco, Cusco, Peru; [jjjjj]Département des sciences biologiques, Centre for Forest Research, Université du Québec à Montreal, Montreal, QC, Canada H3C 3P8; [kkkkk]Department of Ecology and Environmental Sciences, Pondicherry University, Puducherry 605 014, India; [lllll]Instituto Nacional de Tecnologia Agropecuaria–Universidad Nacional de la Patagonia Austral–CONICET, Río Gallegos CP 9400, Argentina; [mmmmm]Cirad, UMREcoFoG (Agroparistech, CNRS, INRAE, Université des Antilles, Université de Guyane) 97 310 Kourou, French Guiana; [nnnnn]Urban Ecosystem Research, School of Social Science, Western Sydney University, Penrith, NSW 2751, Australia; [ooooo]Groupement d'Intérêt Public (GIP) Ecofor, Paris 75116, France; [ppppp]Laboratory of Tropical Dendrology and Forestry, Training Center in Agroforestry Sciences, Federal University of Southern Bahia, Ilheus 45613-204, Brazil; [qqqqq]School of Life Sciences, Chair of Forest Growth and Yield Science, Technical University of Munich 85354 Freising, Germany; [rrrrr]Instituto de Investigaciones paa el Desarrollo Forestal, Universidad de Los Andes, Mérida 5101, Venezuela; [sssss]Escuela Ambiental, Facultad de Ingeniería, Universidad de Antioquia, Medellin, Colombia; [ttttt]Agriculture Food Environment Centre (C3A), University of Trento, San Michele all'Adige 38122, Italy; [uuuuu]Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige 38010, Italy; [vvvvv]Herbario Selva Central (HOXA), Jardín Botánico de Missouri, Oxapampa, Pasco Mz.E-6, Peru; [wwwwww]Laboratory of Tropical Dendrology and Forestry, Training Center in Agroforestry Sciences, Federal University of Southern Bahia, Ilheus 45613-204, Brazil; [xxxxx]Department of Biology, University of Florence 50019 Sesto Fiorentino, Italy; [yyyyy]MUSE-Museo delle Scienze 38122 Trento, Italy; [zzzzz]InfoFlora, Conservatoire et Jardin Botanique de Genève 1292 Chambesy, Switzerland; [aaaaaa]Department of Environmental Sciences, Central University of Jharkhand, Ranchi 835205, India; [bbbbbb]Centro de Modelacion y Monitoreo de Ecosistemas, Universidad Mayor, Universidad de La Frontera, Santiago, Chile; [cccccc]Vicerrectoría de Investigación y Postgrado, Universidad de La Frontera, Temuco 4811230, Chile; [dddddd]Departamento de Silvicultura y Cons. de la Naturaleza, Universidad de Chile, Santiago 8820808, Chile; [eeeeee]Sukachev Institute of Forest of the SB Russian Academy of Sciences 660036 Krasnoyarsk, Russia; [ffffff]International Institute for Applied Systems Analysis, Laxenburg A-2361, Austria; [gggggg]Faculty of Biology, Geobotany, University of Freiburg 79104 Freiburg, Germany; [hhhhhh]National Forest Centre 96001 Zvolen, Slovakia; [iiiiii]Laboratory of Botany and Plant Ecology, Center of Biological and Nature Sciences, Federal University of Acre, Rio Branco 69920-900, Brazil; [jjjjjj]Environmental and Life Sciences, Faculty of Science, Universiti Brunei Darussalam BE1410 Gadong, Brunei; [kkkkkk]Plant Systematic and Ecology Laboratory, Department of Biology, Higher Teachers' Training College, University of Yaounde I, Yaounde, Cameroon; [llllll]CB, Departamento de Ecologia, Universidade Federal do Rio Grande do Norte, Natal 59072-970, Brazil; [mmmmmm]Department of Geomatics, Forest Research Institute, Sękocin Stary, Raszyn 05-090, Poland; [nnnnnn]Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague 16521 Prague, Czech Republic; [oooooo]Department of Plant Biology, Faculty of Sciences, University of Yaounde I, Yaounde, Cameroon; [pppppp]Bioversity International, IITA Regional Office in Cameroon, Yaounde, Cameroon; [qqqqqq]Department of Biology and Florida Museum of Natural History, University of Florida, Gainesville, FL 32611; [rrrrrr]School of Science and Engineering, James Cook University, Cairns QLD 4878, Australia; [ssssss]Instituto de Investigaciones paa el Desarrollo Forestal, Universidad de Los Andes, Mérida 5101, Venezuela; [tttttt]Systematic Botany and Functional Biodiversity, Institute of Biology, Leipzig University 04103 Leipzig, Germany; [uuuuuu]Centre for the Research and Technology of Agro-Environmental and Biological Sciences, CITAB, University of Trás-os-Montes and Alto Douro, UTAD, Quinta de Prados 5000-801 Vila Real, Viseu, Portugal; [vvvvvv]Department of Ecology and Sustainable Agriculture, Agricultural High School, Polytechnic Institute of Viseu 3500-606 Viseu, Portugal; [wwwwwww]Department of Forest Engineering, Universidade Regional de Blumenau, Blumenau-Santa Catarina 89030, Brazil; [xxxxxx]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720; [yyyyyy]Universidad Autónoma del Beni, Riberalta, Beni 2W3Q+VHJ, Bolivia; [zzzzzz]Key Laboratory of Tropical Biological

ECOLOGY

Downloaded by guest on February 10, 2022

Resources of Ministry of Education, School of Life and Pharmaceutical Sciences, Hainan University, Haikou 570228, China; [aaaaaaa]Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå 750 07, Sweden; [bbbbbbb]Ministere des Eaux, des Forets, de la Mer et de l'Environnement chargé du Plan Climat et Objectifs de Development Durable, Estuaire Libreville, Gabon; [ccccccc]Institut de Recherche en Ecologie Tropicale, Libreville, Gabon; [ddddddd]Faculty of Natural Sciences, University of Stirling FK9 4LA, United Kingdom; [eeeeeee]Manaaki Whenua–Landcare Research, Lincoln 7640, New Zealand; [fffffff]State Forests–Coordination Centre for Environmental Projects, Warsaw 02-362, Poland; [ggggggg]Key Laboratory of Tropical Biological Resources of Ministry of Education, School of Life and Pharmaceutical Sciences, Hainan University, Haikou 570228, China; and [hhhhhhh]Département FOREN, Institut National Polytechnique Félix Houphouët-Boigny, Yamoussoukro BP 1093, Côte d'Ivoire

1. R. M. May, Conceptual aspects of the quantification of the extent of biological diversity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **345**, 13–20 (1994).
2. R. Cazzolla Gatti, A century of biodiversity: Some open questions and some answers. *Biodiversity* **18**, 175–185 (2017).
3. E. Beech, M. Rivers, S. Oldfield, P. P. Smith, Global Tree Search: The first complete global database of tree species and country distributions. *J. Sustain. For.* **36**, 454–489 (2017).
4. H. Qian, T. Deng, H. Sun, Global and regional tree species diversity. *J. Plant Ecol.* **12**, 210–215 (2018).
5. K. J. Gaston, Global patterns in biodiversity. *Nature* **405**, 220–227 (2000).
6. P. Crane, Can we save the charismatic megaflora? *Oryx* **49**, 377–378 (2015).
7. E.-D. Schulze, "The carbon and nitrogen cycle in forest ecosystems" in *Carbon and Nitrogen Cycling in European Forest Ecosystems*, E.-D. Schulze, Ed. (Ecological Studies [Analysis and Synthesis], Springer, Berlin, 2000), **vol. 142**, pp. 3–13.
8. A. R. Townsend, G. P. Asner, C. C. Cleveland, The biogeochemical heterogeneity of tropical forests. *Trends Ecol. Evol.* **23**, 424–431 (2008).
9. J. Liang *et al.*, Positive biodiversity-productivity relationship predominant in global forests. *Science* **354**, aaf8957 (2016).
10. P. Balvanera *et al.*, Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett.* **9**, 1146–1156 (2006).
11. B. A. Hawkins, M. Rodríguez, S. G. Weller, Global angiosperm family richness revisited: Linking ecology and evolution to climate. *J. Biogeogr.* **38**, 1253–1266 (2011).
12. S. P. Hubbell *et al.*, Colloquium paper: How many tree species are there in the Amazon and how many of them will go extinct? *Proc. Natl. Acad. Sci. U.S.A.* **105** (suppl. 1), 11498–11504 (2008).
13. B. J. Cardinale *et al.*, Biodiversity loss and its impact on humanity. *Nature* **486**, 59–67 (2012).
14. G. C. Stevens, The latitudinal gradient in geographical range: How so many species coexist in the tropics. *Am. Nat.* **133**, 240–256 (1989).
15. Z. Wang, J. H. Brown, Z. Tang, J. Fang, Temperature dependence, spatial scale, and tree species diversity in eastern Asia and North America. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13388–13392 (2009).
16. B. J. Enquist *et al.*, The commonness of rarity: Global and future distribution of rarity across land plants. *Sci. Adv.* **5**, eaaz0414 (2019).
17. H. ter Steege *et al.*, Hyperdominance in the Amazonian tree flora. *Science* **342**, 1243092 (2013).
18. R. P. Leitão *et al.*, Rare species contribute disproportionately to the functional structure of species assemblages. *Proc. Biol. Sci.* **283**, 20160084 (2016).
19. L. E. Dee *et al.*, When do ecosystem services depend on rare species? *Trends Ecol. Evol.* **34**, 746–758 (2019).
20. S. Fauset *et al.*, Hyperdominance in Amazonian forest carbon cycling. *Nat. Commun.* **6**, 6857 (2015).
21. J. F. Bastin *et al.*, The extent of forest in dryland biomes. *Science* **356**, 635–638 (2017).
22. A. D. Barnosky *et al.*, Has the Earth's sixth mass extinction already arrived? *Nature* **471**, 51–57 (2011).
23. T. M. Brooks *et al.*, Global biodiversity conservation priorities. *Science* **313**, 58–61 (2006).
24. D. Cardoso *et al.*, Amazon plant diversity revealed by a taxonomically verified species list. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10695–10700 (2017).
25. B. R. Scheffers, L. N. Joppa, S. L. Pimm, W. F. Laurance, What we know and don't know about Earth's missing biodiversity. *Trends Ecol. Evol.* **27**, 501–510 (2012).
26. J. M. Serra-Diaz, B. J. Enquist, B. Maitner, C. Merow, J. C. Svenning, Big data of tree species distributions: How big and how good? *For. Ecosyst.* **4**, 30 (2017).
27. A. Chao, Nonparametric estimation of the number of classes in a population. *Scand. J. Stat.* **11**, 265–270 (1984).
28. C. H. Chiu, A. Chao, Estimating and comparing microbial diversity in the presence of sequencing errors. *PeerJ* **4**, e1634 (2016).
29. A. Chao, Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* **43**, 783–791 (1987).
30. S. Oldfield, C. Lusty, A. MacKinven, *The World List of Threatened Trees* (World Conservation Press, Cambridge, United Kingdom, 1998).
31. N. J. Gotelli, R. K. Colwell, "Estimating species richness" in *Biological Diversity: Frontiers in Measurement and Assessment*, A. E. Magurran, B. J. McGill, Eds. (Oxford University Press, Oxford, United Kingdom, 2011), pp. 39–54.
32. K. J. Iknayan, M. W. Tingley, B. J. Furnas, S. R. Beissinger, Detecting diversity: Emerging methods to estimate species diversity. *Trends Ecol. Evol.* **29**, 97–106 (2014).
33. A. Chao, R. K. Colwell, C. W. Lin, N. J. Gotelli, Sufficient sampling for asymptotic minimum species richness estimators. *Ecology* **90**, 1125–1133 (2009).
34. H. Ter Steege *et al.*, The discovery of the Amazonian tree flora with an updated checklist of all known tree taxa. *Sci. Rep.* **6**, 29549 (2016).
35. H. Ter Steege *et al.*, Towards a dynamic list of Amazonian tree species. *Sci. Rep.* **9**, 3501 (2019).
36. J. W. Slik *et al.*, An estimate of the number of tropical tree species. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7472–7477 (2015). Correction in: *Proc. Natl. Acad. Sci. U.S.A.* **112**, E4628–E4629 (2015).
37. H. Ter Steege *et al.*, Biased-corrected richness estimates for the Amazonian tree flora. *Sci. Rep.* **10**, 10130 (2020).
38. A. Chao, L. Jost, Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology* **93**, 2533–2547 (2012).
39. N. C. Pitman, P. M. Jørgensen, R. S. Williams, S. León-Yánez, R. Valencia, Extinction-rate estimates for a modern neotropical flora. *Conserv. Biol.* **16**, 1427–1431 (2002).
40. O. R. Wearn, D. C. Reuman, R. M. Ewers, Extinction debt and windows of conservation opportunity in the Brazilian Amazon. *Science* **337**, 228–232 (2012).
41. S. Chamberlain, K. Ram, V. Barve, D. Mcglinn, rgbif: Interface to the Global "Biodiversity" Information Facility "API" (R Package Version 0.9.8, 2016). https://cran.r-project.org/web/packages/rgbif/index.html. Accessed 15 March 2020.
42. B. S. Maitner *et al.*, The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods Ecol. Evol.* **9**, 373–379 (2018).
43. K. Banda-R *et al.*; DRYFLOR, Plant diversity patterns in neotropical dry forests and their conservation implications. *Science* **353**, 1383–1387 (2016).
44. G. Dauby *et al.*, RAINBIO: A mega-database of tropical African vascular plants distributions. *PhytoKeys* **74**, 1–18 (2016).
45. P. Flemons, B. Raymond, P. Brenton, L. Belbin, Atlas of living Australia. https://www.ala.org.au/. Accessed 12 February 2019.
46. BGCI, GlobalTreeSearch online database. https://www.bgci.org/global_tree_search.php. Accessed 19 June 2017.
47. B. Boyle *et al.*, The taxonomic name resolution service: An online tool for automated standardization of plant names. *BMC Bioinformatics* **14**, 16 (2013).
48. A. Chao *et al.*, Quantifying sample completeness and comparing diversities among assemblages. *Ecol. Res.* **35**, 292–314 (2020).
49. A. Chao, S. M. Lee, Estimating the number of classes via sample coverage. *J. Am. Stat. Assoc.* **87**, 210–217 (1992).
50. C. H. Chiu, Y. T. Wang, B. A. Walther, A. Chao, An improved nonparametric lower bound of species richness via a modified good-turing frequency formula. *Biometrics* **70**, 671–682 (2014).
51. J. P. Mellard, P. Audoye, M. Loreau, Seasonal patterns in species diversity across biomes. *Ecology* **100**, e02627 (2019).
52. T. F. Allen, T. W. Hoekstra, The confusion between scale-defined levels and conventional levels of organization in ecology. *J. Veg. Sci.* **1**, 5–12 (1990).
53. N. J. Gotelli, R. K. Colwell, Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.* **4**, 379–391 (2001).
54. A. E. Magurran, *Measuring Biological Diversity* (Blackwell Science, Oxford, United Kingdom, 2004).
55. K. Mao *et al.*, Distribution of living Cupressaceae reflects the breakup of Pangea. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7793–7798 (2012).
56. K. Mao, G. Hao, J. Liu, R. P. Adams, R. I. Milne, Diversification and biogeography of Juniperus (Cupressaceae): Variable diversification rates and multiple intercontinental dispersals. *New Phytol.* **188**, 254–272 (2010).
57. J. H. Whiteside, D. S. Grogan, P. E. Olsen, D. V. Kent, Climatically driven biogeographic provinces of Late Triassic tropical Pangea. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8972–8977 (2011).
58. M. D. Crisp, L. G. Cook, Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. *New Phytol.* **192**, 997–1009 (2011).
59. A. E. Magurran, *Ecological Diversity and Its Measurement* (Princeton University Press, Princeton, NJ, 1988).
60. T. W. Crowther *et al.*, Mapping tree density at a global scale. *Nature* **525**, 201–205 (2015).
61. R. A. Fisher, A. S. Corbet, C. B. Williams, The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **12**, 42–58 (1943).
62. S. P. Hubbell, Estimating the global number of tropical tree species, and Fisher's paradox. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7343–7344 (2015).
63. I. Volkov, J. R. Banavar, S. P. Hubbell, A. Maritan, Patterns of relative species abundance in rainforests and coral reefs. *Nature* **450**, 45–49 (2007).
64. R. K. Colwell, J. A. Coddington, Estimating terrestrial biodiversity through extrapolation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **345**, 101–118 (1994).
65. B. A. Walther, S. Morand, Comparative performance of species richness estimation methods. *Parasitology* **116**, 395–405 (1998).

Downloaded by guest on February 10, 2022

66. A. Chao, R. K. Colwell, Thirty years of progeny from Chao's inequality: Estimating and comparing richness with incidence data and incomplete sampling. *SORT Stat. Oper. Res. Trans.* **41**, 3–54 (2017).
67. T. C. Hsieh, K. H. Ma, A. Chao, iNEXT: Interpolation and Extrapolation for Species Diversity (R Package Version 2.0.20, 2020). http://chao.stat.nthu.edu.tw/wordpress/software_download/. Accessed 15 March 2020.
68. U. Brose, N. D. Martinez, R. J. Williams, Estimating species richness: Sensitivity to sample coverage and insensitivity to spatial patterns. *Ecology* **84**, 2364–2377 (2003).
69. H. Wagner, O. Wildi, Realistic simulation of the effects of abundance distribution and spatial heterogeneity on non-parametric estimators of species richness. *Ecoscience* **9**, 241–250 (2002).
70. W. E. Kunin *et al.*, Upscaling biodiversity: Estimating the species–area relationship from small samples. *Ecol. Monogr.* **88**, 170–187 (2018).
71. H. Xu *et al.*, Assessing non-parametric and area-based methods for estimating regional species richness. *J. Veg. Sci.* **23**, 1006–1012 (2012).
72. H. Chen, P. C. Boutros, VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35 (2011).



Downloaded by guest on February 10, 2022