How many tropical forest tree species are there?

J. W. Ferry Slik[1], Víctor Arroyo-Rodríguez[2], Shin-Ichiro Aiba[3], Patricia Alvarez-Loayza[4], Luciana F. Alves[5], Peter Ashton[6], Patricia Balvanera[2], Meredith Bastian[7], Peter J. Bellingham[8], Eduardo van den Berg[9], Luis Bernacci[5], Polyanna da Conceição Bispo[10], Lilian Blanc[11], Katrin Böhning-Gaese[12], Pascal Boeckx[13], Frans Bongers[14], Brad Boyle[15], Matt Bradford[16], Francis Q. Brearley[17], Mireille Breuer Ndoundou Hockemba[18], Sarayudh Bunyavejchewin[19], Darley Calderado Leal Matos[20], Miguel Castillo-Santiago[21], Eduardo L. M. Catharino[22], Shauna-Lee Chai[23], Yukai Chen[24], Robert K. Colwell[25,26], Robin L. Chazdon[25], Connie Clark[4], David B. Clark[27], Deborah A. Clark[27], Heike Culmsee[28], Kipiro Damas[29], Handanakere S. Dattaraja[30], Gilles Dauby[31], Priya Davidar[32], Saara J. DeWalt[33], Jean-Louis Doucet[34], Alvaro Duque[35], Giselda Durigan[36], Karl Eichhorn[37], Pedro V. Eisenlohr[38], Eduardo Eler[39], Corneille Ewango[18], Nina Farwig[40], Kenneth J. Feeley[41], Leandro Ferreira[20], Richard Field[42], Ary T. de Oliveira Filho[43], Christine Fletcher[44], Olle Forshed[45], Geraldo Franco[36], Gabriella Fredriksson[46], Thomas Gillespie[47], Jean-François Gillet[34], Giriraj Amarnath[48], Daniel M. Griffith[49], James Grogan[50], Nimal Gunatilleke[51], David Harris[52], Rhett Harrison[53,54], Andy Hector[55], Jürgen Homeier[28], Nobuo Imai[56], Akira Itoh[57], Patrick A. Jansen[14,58], Carlos A. Joly[59], Bernardus H.J. de Jong[21], Kuswata Kartawinata[60], Elizabeth Kearsley[13], Daniel L. Kelly[61], David Kenfack[62], Michael Kessler[63], Kanehiro Kitayama[64], Robert Kooyman[65], Eileen Larney[66], Yves Laumonier[67], Susan Laurance[68], William Laurance[68], Michael J. Lawes[69], Ieda Leao do Amaral[39], Susan G. Letcher[70], Jeremy Lindsell[71], Xinghui Lu[72], Asyraf Mansor[73], Antti Marjokorpi[74], Emanuel H. Martin[75,76], Henrik Meilby[77], Felipe P. L. Melo[78], Dan Metcalfe[79], Vincent P. Medjibe[4], Jean Paul Metzger[80], Jerome Millet[81], D. Mohandass[82], Juan Carlos Montero[83], Márcio de Morisson Valeriano[10], Badru Mugerwa[84], Hidetoshi Nagamasu[64], Rueben Nilus[85], Susana Ochoa-Gaona[21], Onrizal[86], Navendu Page[30], Pia Parolin[87], Marc Parren[14], Narayanaswamy Parthasarathy[32], Ekananda Paudel[53], Andrea Permana[63], Maria T. F. Piedade[39], Nigel Pitman[4], Lourens Poorter[14], Axel Poulsen[88], John Poulsen[4], Jennifer Powers[89], Rama Chandra Prasad[90], Jean-Philippe Puyravaud[91], Jean-Claude Razafimahaimodison[66], Jan Reitsma[92], João Roberto dos Santos[10], Wilson Roberto Spironello[39], Hugo Romero-Saltos[93], Francesco Rovero[94], Andes Rozak[95], Kalle Ruokolainen[96], Ervan Rutishauser[97], Felipe Saiter[98], Philippe Saner[63], Braulio A. Santos[99], Fernanda Santos[20], Swapan Kumar Sarker[100], Manichanh Satdichanh[53], Christine B. Schmitt[101], Jochen Schöngart[102], Mark Schulze[103], Marcio Seiji Suganuma[80], Douglas Sheil[67,104], Eduardo da Silva Pinheiro[105], Plinio Sist[106], Tariq Stevart[107], Raman Sukumar[30], Fang Sun[108], Terry Sunderland[67], H.S. Suresh[30], Eizi Suzuki[3], Marcelo Tabarelli[78], Jangwei Tang[82], Natália Targhetta[102], Ida Theilade[77], Duncan W. Thomas[109], Peguy Tchouto[110], Johanna Hurtado[111], Renato Valencia[112], Johan van Valkenburg[113], Tran Van Do[114], Rodolfo Vasquez[107], Hans Verbeeck[13], Victor Adekunle[115], Simone A. Vieira[59], Campbell Webb[116], Timothy Whitfeld[117], Serge Wich[118,119], John Williams[120], Florian Wittmann[121], Hannsjoerg Wöll[122], Xiaobo Yang[123], C. Yves Adou Yao[124], Sandra Yap[125], Tsuyoshi Yoneda[3], Rakan A. Zahawi[111], Rahmad Zakaria[73], Runguo Zang[72]

[1] Faculty of Science, University Brunei Darusallam, Gadong, Brunei Darusallam. [2] Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México, Morelia, Mexico. [3] Kagoshima University, Kagoshima, Japan. [4] Center for Tropical Conservation, Duke University, Durham, North Carolina, USA. [5] Instituto Agronômico de Campinas, Campinas, Brazil. [6] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachussetts, USA. [7] Smithsonian's National Zoo, Washington DC, USA. [8] Landcare Research, Lincoln, New Zealand. [9] Universidade Federal de Lavras, Lavras, Brazil. [10] National Institute for Space Research-INPE, São José dos Campos, São Paulo, Brazil. [11] CIRAD, Belem, Brazil. [12] Biodiversity and Climate ResearchCentre (BiK-F), Senckenberg, Gesellschaft für Naturforschung, Frankfurt (Main), Germany. [13] Ghent University, Gent, Belgium. [14] Wageningen University, Wageningen, The Netherlands. [15] University of Arizona, Tucson, Arizona, USA. [16] CSIRO Land and Water, Tropical Forest Research Centre, Atherton, Australia. [17] Manchester Metropolitan University, Manchester, UK. [18] Wildlife Conservation Society - Congo Program, Nouabalé-Ndoki National Park, Republic of Congo. [19] Department of National Parks, Wildlife and Plant Conservation, Bangkok, Thailand. [20] Museu Paraense Emílio Goeldi, Belem, Brazil. [21] El Colegio de la Frontera Sur (ECOSUR), Mexico. [22] Instituto de Botânica de São Paulo, São Paulo, Brazil. [23] Alberta Innovates - Technology Futures, Vegreville, Alberta, Canada. [24] College of Life Sciences, Hainan Normal University, Haikou, China. [25] University of Connecticut, Storrs, Connecticut, USA, [26] University of Colorado, Boulder, Colorado, USA. [27] University of Missouri, St. Louis, Missouri, USA. [28] University of Göttingen, Göttingen, Germany. [29] PNG Forest Research Institute, Lae, Papua New Guinea. [30] Indian Institute of Science, Bangalore, India. [31] Universite Libre de Bruxelles, Brussels, Belgium. [32] Pondicherry University, Puducherry, India. [33] Clemson University, Clemson, South Carolina, USA. [34] Gembloux Agro-Bio Tech, Université de Liège, Gembloux, Belgium. [35] Universidad Nacional de Colombia, Medellin, Colombia. [36] Instituto Florestal de Sao Paulo, Sao Paulo, Brazil. [37] Eichhorn Ecology, Zeist, The Netherlands. [38] Universidade do Estado de Mato Grosso, Caceres, Brazil. [39] Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil. [40] Philipps-Universität Marburg, Marburg, Germany. [41] International Center for Tropical Botany, Florida International University, Miami, Florida, USA. [42] University of Nottingham, Nottingham, UK. [43] Instituto de Ciências Biológicas (ICB), Universidade Federal de Minas Gerais (UFMG), Pampulha, Belo Horizonte, Brazil. [44] Forest Research Institute Malaysia, Kuala Lumpur, Malaysia. [45] WWF Sweden, Stockholm, Sweden. [46] PanEco/SOCP, Medan, North Sumatra, Indonesia. [47] University of California, Los Angeles, California, USA. [48] International Water Management Institute (IWMI), Colombo, Sri Lanka. [49] Universidad Laica Eloy Alfaro de Manabi, Manta, Ecuador. [50] Mount Holyoke College, South Hadley, MA, USA. [51] University of Peradeniya, Peradeniya, Sri Lanka. [52] Royal Botanical Garden Edinburgh, Edinburgh, UK. [53] Centre for Mountain Ecosystem Studies (CMES), Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China. [54] China & World Agroforestry Centre (ICRAF), East and Central Asia Regional Office, Kunming, China. [55] University of Oxford, Oxford, UK. [56] Primate Research Institute, Kyoto University, Kanrin, Inuyama, Aichi, Japan. [57] Osaka City University, Osaka, Japan. [58] Smithsonian Tropical Research Institute, Panama. [59] Universidade Estadual de Campinas, Campinas, Brazil. [60] Herbarium Bogoriense, Research Center for Biology, Indonesian Institute of Sciences (LIPI), Cibinong, Indonesia. [61] Trinity College, The University of Dublin, Dublin, Ireland. [62] Smithsonian Tropical Research Institute, Washington DC, USA. [63] University of Zurich, Zurich, Switzerland. [64] Kyoto University, Kyoto, Japan. [65] Macquarie University, Sydney, Australia. [66] Centre ValBio, Ranomafana, Madagascar. [67] CIFOR, Bogor, Indonesia. [68] James Cook University, Cairns, Australia. [69] Research Institute for the Environment and Livelihoods, Charles Darwin University, Darwin, Australia. [70] Purchase College (SUNY), Purchase, New York, USA. [71] RSPB, Sandy, Bedfordshire, UK. [72] Institute of Forest Ecology, Environment and Protection, Chinese Academy of Forestry, Beijing, China. [73] Universiti Sains Malaysia, Georgetown, Penang, Malaysia. [74] Stora Enso Oyj, Helsinki, Finland. [75] Udzungwa Ecological Monitoring Centre, Mang'ula, Tanzania. [76] Sokoine University of Agriculture, Morogoro, Tanzania. [77] University of Copenhagen, Copenhagen, Denmark. [78] Universidade Federal de Pernambuco, Recife, Brazil. [79] CSIRO Land and Water, EcoSciences Precinct, Dutton Park, Australia. [80] Universidade de São Paulo, Sao Paulo, Brazil. [81] National University of Laos, Vientiane, Laos. [82] Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, China. [83] Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany. [84] Institute of Tropical Forest Conservation, Kampala, Uganda. [85] Sabah Forestry Department, Sandakan, Sabah, Malaysia. [86] University of Sumatera Utara, Medan, Sumatra, Indonesia. [87] University of Hamburg, Hamburg, Germany. [88] University of Oslo, Oslo, Norway. [89] Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, Minnesota, USA. [90] International Institute of Information Technology, Hyderabad, India. [91] Sigur Nature Trust, Masinagudi, India. [92] Bureau Waardenburg bv, Culemborg, The Netherlands. [93] Department of Biology, Yachay Tech University, Ciudad del Conocimiento, Urcuquí, Ecuador. [94] MUSE - Museo delle Scienze, Trento, Italy. [95] Cibodas Botanic Gardens, Indonesian Institute of Sciences (LIPI), Cianjur, Indonesia. [96] University of Turku, Turku, Finland. [97] Carboforexpert, Geneva, Switzerland. [98] Instituto Federal do Espírito Santo, Santa Teresa, Brazil. [99] Universidade Federal da Paraíba, João Pessoa, Brazil. [100] Shahjalal University of Science & Technology, Sylhet, Bangladesh. [101] Chair for Landscape Management, University of Freiburg, Freiburg, Germany. [102] Instituto Nacional de Pesquisas da Amazônia - Max Planck Project, Manaus, Brazil. [103] Oregon State University, Corvallis, OR, USA. [104] Norwegian University of Life Sciences, Aas, Norway. [105] Universidade Federal do Amazonas, Manaus, Brazil. [106] CIRAD, Monpellier, France. [107] Missouri Botanical Garden, St. Louis, Missouri, USA. [108] Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualien, Taiwan. [109] Department of

The high species richness of tropical forests has long been recognized, yet there remains substantial uncertainty regarding the actual number of tropical tree species. Using a pan-tropical tree inventory database from closed canopy forests, consisting of 657,630 trees belonging to 11,371 species, we use a fitted value of Fisher's alpha and an approximate pan-tropical stem total to estimate the minimum number of tropical forest tree species to fall between ∼40,000 and ∼53,000, i.e. at the high end of previous estimates. Contrary to common assumption, the Indo-Pacific region was found to be as species rich as the Neotropics, with both regions having a minimum of ∼19,000-25,000 tree species. Continental Africa is relatively depauperate with a minimum of ∼4,500-6,000 tree species. Very few species are shared among the African, American and the Indo-Pacific regions. We provide a methodological framework for estimating species richness in trees that may help refine species richness estimates of tree dependent taxa.

Diversity estimation | Fishers's log-series | Pan-tropical | Spatial richness patterns | Tropical tree species richness

Significance statement People are fascinated by the amazing diversity of tropical forests and will be surprised to learn that robust estimates of the number of tropical tree species are lacking. We show that there are at least 40,000, but possibly more than 53,000 tree species in the tropics, in contrast to only 124 across temperate Europe. Almost all tropical tree species are restricted to their respective continents, while the Indo-Pacific region appears to be as species rich as tropical America, with each of these two regions being almost five times as rich in tree species as African tropical forests. Our study shows that most tree species are extremely rare, meaning that they are under serious risk of extinction at current deforestation rates.

Despite decades of biological inventories worldwide, we still do not know how many species exist and how they are distributed (1). Although global patterns of estimated vascular plant species richness and distribution have become more clear (2-5), no previous study has focused on trees as a distinct growth form. As a consequence, our estimation of the number of tree species in tropical forests still depends on untested expert opinions (6-8), rather than on an appropriate methodological framework and data set.

Given the importance of trees as key structural components of forest ecosystems, sources of timber and non-timber products, and providers of vital ecosystem services (9, 10), the lack of reliable estimates of the total number of tropical tree species represents a critical knowledge gap that has direct consequences for estimating the diversity of other tree dependent taxa (11). A classic example is Erwin's (6) estimate of the existence of 30 million arthropod species, which was based on observed host specificities of arthropods with individual tropical tree species combined with an estimate of the total number of tropical tree species. Global arthropod richness has subsequently been revised downward (7), but current estimates still suffer from the lack of information on the number of tropical tree species.

In recent decades, the number of tree inventory plots across the tropics has grown to such an extent that species richness estimation at the continental and pan-tropical scale can now be addressed using standardized species lists with abundance data. Prior estimates of plant richness at such broad scales have been based on analyses of incidence data obtained from herbarium collections and Flora treatments (2-5). However, these methods are highly sensitive to collecting biases and ignore valuable information on species' abundances (12). Abundance data enable extrapolation of richness from local to global scales using diversity estimators that fit the observed species rank-abundance data (13-15).

Results and Discussion

We estimate the number of tropical tree species from a standardized dataset of old-growth tropical forest tree inventories (including gymno- and angiosperms with diameters at breast height (dbh) ≥ 10 cm). This dataset contains tree species abundance data for 207 one-degree grid cells (locations) originally dominated by closed canopy forests across the tropics (Fig. 1). By calculating Fisher's-alpha (16) at the pan-tropical scale and combining this value with the estimated potential number of stems present within 500 km of each location, we arrive at a minimum number of tropical tree species of at least ∼40,000 and possibly more than ∼53,000 (Table 1), i.e. at the high end of current total estimates of 37,000 (7), 43,000 (8) and 50,000 (6), which are based on expert opinion.

When the analysis was restricted to each of the three main tropical regions, we found that the Indo-Pacific had comparable tree species richness to that found in tropical America (Table 1). Moreover, these two regions show similar rates of species turnover for a given increase in geographical distance between locations (Fig. 2). This result contradicts the widely held view that the Neotropics are the most diverse and species-rich region for tropical trees (8, 15, 17, 18). This underestimation of Indo-Pacific tree species richness, and our inclusion of dry as well as moist and wet forests, may explain why some of the previous estimates (7, 8) are lower than ours. Nevertheless, the high species richness in the Indo-Pacific is understandable given the highly variable topography, complex geological history, steep environmental gradients, past and ongoing merging of several contrasting floras from Madagascar, India, Southeast Asia, and New Guinea-Australia (19, 20), as well as the large current and time integrated forest area (8).

Tropical continental Africa has a relatively depauperate tree flora, a finding consistent with earlier studies (21, 22). This region shows comparatively low species turnover, in other words, as sample area increases, the number of tree species increases at a much slower rate than in either the Indo-Pacific or the Neotropics (Fig. 2). The differences in species richness and spatial turnover, comparing continental Africa with the other tropical regions, cannot be explained solely by Africa's smaller forest area or lower environmental variability (Table 1; Fig. S2). Rather, these disparities support the hypothesis that African forests have

---

Significance

People are fascinated by the amazing diversity of tropical forests and will be surprised to learn that robust estimates of the number of tropical tree species are lacking. We show that there are at least 40,000, but possibly more than 53,000 tree species in the tropics, in contrast to only 124 across temperate Europe. Almost all tropical tree species are restricted to their respective continents, while the Indo-Pacific region appears to be as species rich as tropical America, with each of these two regions being almost five times as rich in tree species as African tropical forests. Our study shows that most tree species are extremely rare, meaning that they are under serious risk of extinction at current deforestation rates.

Reserved for Publication Footnotes

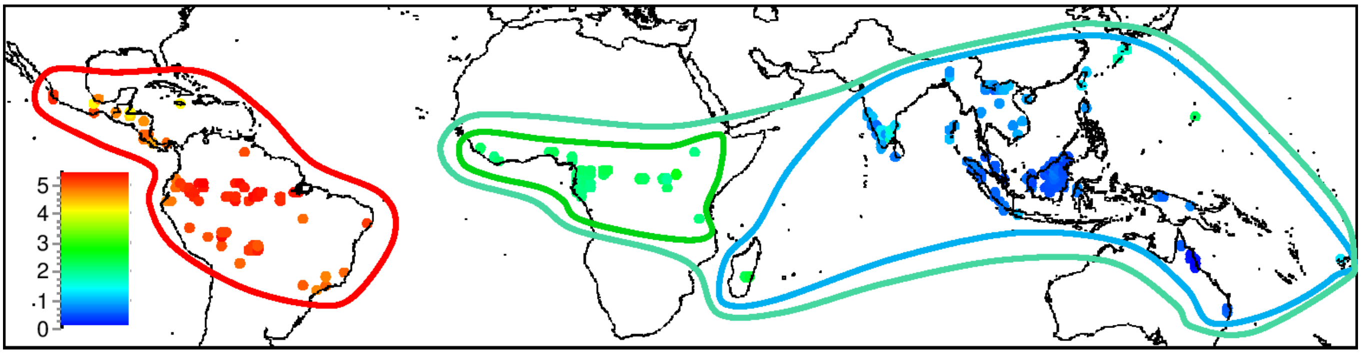

Fig. 1. Overview of sample locations and their floristic affinities (point colors correspond to scores on the first DCA-axis with similar colours indicating similar generic composition, while the lines indicate the floristic affinities as determined by cluster analysis).

Table 1. Species richness estimates for the tropics and the three main tropical sub-regions.

Region	Species observed	Stems observed	Unidentified stems (%)	Fisher's-alpha minimum	Fisher's- alpha maximum	Stems estimated	Species minimum	Species maximum
Africa	1376	117902	8.4	218.7	286.6	$3.4\text{x}10^{11}$	4626	5984
America	4375	116754	13.5	897.2	1203.4	$8.9\text{x}10^{11}$	18589	24580
Indo-Pacific	5672	422974	9.6	925.8	1225.2	$7.7\text{x}10^{11}$	19014	24819
Pan-tropical	11371	657630	10.1	1953.0	2607.7	$2.0\text{x}10^{12}$	40517	53345

Observed values represent the numbers in the original dataset. Unidentified stems were excluded from estimation of minimum Fisher's-alpha, but used to calculate maximum Fisher's-alpha. Estimated stems represent the number of stems predicted to occur within 500 km of each of the 207 sample locations.

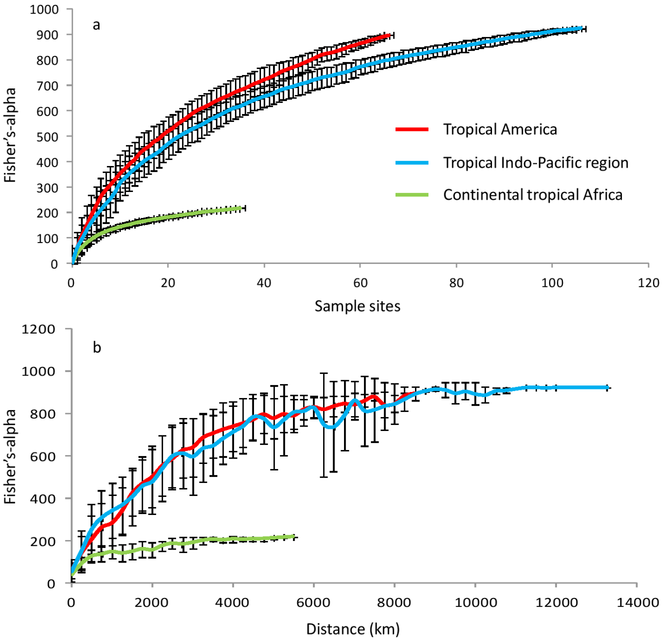

Fig. 2. Increase in Fisher's-alpha with (a) increasing numbers of locations (average of 50 replicates per region with random input order of locations), i.e. regional diversity and (b) increasing distance around locations (based on 50 replicates per region each with a randomly selected starting location), i.e. species turnover. Error bars indicate standard deviation among location re-orderings. Fisher's-alpha can decline if the number of stems added to the sample increases disproportionally to the number of new species detected.

experienced severe extinction events due to repeated shrinkage of forest area during the Pleistocene (19, 23). When these forests expanded to their present size, they could only be repopulated by a severely depleted species pool derived from a limited number of refugia. In contrast, tropical America retained considerable forest cover and equatorial forests of the Indo-Pacific may even have expanded during the same period (19, 20, 23, 24).

We provide the first survey-based minimum estimate of tropical tree species richness and its distribution. We acknowledge, however, that the current estimate is just a first step in an ongoing effort. Estimates of species richness will become more refined and increasingly accurate as forest surveys continue to expand. This study highlights the usefulness and critical importance of forest surveys, and we emphasize once more the existence of large numbers of tree species with exceptionally small population sizes, which may necessitate novel conservation approaches for effective preservation of current tree diversity (25, 26).

Materials and Methods

Data set. Tree inventory data (gymnosperms and angiosperms only; trunk diameter at breast height ≥ 10 cm) from old-growth forest plots without signs of recent human disturbance were compiled from across the tropics and subtropics (i.e., within 30° north and south of the Equator). Individual trees from the inventories were pooled within their respective one degree grid cells (henceforth called *locations*). Species names were standardized using 'The Plant List' <www.theplantlist.org>, 'Taxonomic Name Resolution Service' <http://tnrs.iplantcollaborative.org/TNRSapp.html> and 'The Asian Plant Synonym Lookup' <http://phylodiversity.net/fslik/synonym_lookup.htm>. Unknown taxa were not used in diversity and composition analyses. We removed all locations with fewer than 250 identified individuals to minimize effects of sample size, resulting in 207 retained locations (Dataset S1), each showing a reasonably high sample coverage (0.96 ± 0.3, mean ± SE) that did not differ among geographic areas (Table S1), thus suggesting that our estimations of species richness were not biased by differences in sample coverage among regions.

Phylogeographic analyses. To identify the main tropical regions for species richness comparison, we performed 'Minimum Variance Clustering' with 'Squared Euclidean Distances' on square-root transformed relative abundance data at the genus level (Fig. S1). These analyses were conducted at the genus level because virtually no overlap existed between continents at the species level. We subsequently ran a Detrended Component Analysis to visualize these floristic patterns across the tropics (Fig. 1). To assess if difference in species richness and turnover among regions are related to differences in environmental variability among regions, we performed a Principal Component Analysis (PCA) of locations versus environmental variables (climatic data (27) and edaphic data (28)) (Fig. S2).

Sample coverage and non-parametric estimators of species richness. Because estimates of species richness can be strongly dependent on differences in inventory completeness (29, 30), we estimated the inventory completeness for the complete database and for each region separately using the sample coverage estimator recommended by Chao & Shen (31), which is a reduced-bias estimator of sample completeness:

$$\hat{C}_n = 1 - \frac{f_1}{n}\left[\frac{(n-1)f_1}{(n-1)f_1 + 2f_2}\right]$$

where f_1 and f_2 are the number of species represented by one (singletons) and two (doubletons) individuals in the sample, respectively, n is the total number of individuals in the sample, and $\hat{C}_n$ is the proportion of the total number of individuals in an assemblage (observed and not observed) that belong to the species represented in the sample. Sample coverage was almost identical in all regions (Table S1), suggesting that our regional comparisons of species richness were not biased by differences in sample coverage among regions, nor was our pan-tropical estimate disproportionately influenced by any one region (30).

Estimates of sample coverage indicated that more than 90% of all the trees present in the tropics belonged to species included in our recorded samples (Table S1). We also computed estimates of species richness based on several non-parametric estimators (Table S2). In particular, we considered nine of the estimators available in the software SPADE (31), and they estimated that, on average, we have recorded between 79% and 91% of the species present in all regions (Table S2). However, consistent with ter Steege et al. (15), we found that these methods underestimated the species richness, as estimates for each tropical region (1,539 to 1,680 species in Africa; 4,959 to 5,540 in America; and 6,232 to 6,784 species in Asia) were between one-half and one-third of previous estimates based on expert opinion and available floras (6-8). In fact, recent estimates suggest that there are approximately 16,000 tree species in the Amazon (15), so having fewer than 5,540 tree species in the whole Neotropics is highly unlikely (Table S2). As ter Steege et al. (15) argued, the failure of these non-parametric methods to yield plausible estimates arises from fact that these estimators are designed to estimate the expected number of species at a local scale, based on samples that are fully representative of the area sampled (15, 32). However, like ter Steege et al. (15), we are attempting to estimate the number of species for the whole tropics, including areas that have been poorly sampled.

Selection of species-abundance distribution (SAD) model to predict species richness. Several models have been proposed to describe observed species–abundance distributions (SADs) within a community (33). SAD models allow for an understanding of the abundance structure of biological communities, and can be useful for predicting unsampled portions of communities. The fit of SAD models to the data depends, among other factors, on community evenness and sampling intensity (33, 34). For instance, extremely uneven SADs are predicted by the geometric series (35), while unusually even SADs are predicted by the broken stick model (36). The log series (16) and log normal (37) models are intermediate, differing in the assumed proportions of rare species: Fisher's log series assumes very high proportions of rare species, while the log normal model assumes very low proportions of rare species (33). Regarding sampling intensity, complete surveys usually follow log-normal types of SADs, whereas incomplete sampled communities usually deviate from log-normality (34).

Here we used Fisher's log series to estimate the expected number of species within each region. The log series distribution is one of the most frequently used and thoroughly investigated models of the relationship between species richness and the relative abundance of species (33), and it has been successfully used to estimate the number of species at different spatial scales (15, 25), though not previously at an intercontinental scale. To fit and compare the log series model with other commonly used SAD models (i.e., log normal, Broken-stick and Pareto [power-law] distributions) we used maximum likelihood tools with the *sads* package for R 3.0.3 (38). We ranked the models from the best to the worst based on Akaike's information criterion (AIC). The set of models with a difference in AIC (Δ) < 2 can be considered to have equivalently strong empirical support and similar plausibility (39).

Graphical comparison of the models showed that the log series provided the best fit to our data (Fig. S3); this result was confirmed by the AIC analyses, in which the log series model had the strongest support (Table S3). The visual analysis of the frequency of species in octaves of abundance (so-called Preston plots) predicted by each SAD model also supported the idea that log series fit the data well and was the best model to predict the proportion of rare species (Figs. S4 - S7). Therefore, we can conclude that the log series is an appropriate SAD model to fit our data.

Application of the Fisher's log-series. Fisher's-alpha values can be used to extrapolate species richness of a defined region if the number of individuals is known. Extrapolations with Fisher's-alpha, however, rely on two assumptions: (i) species abundances are distributed in a log-series manner, and (ii) the plant community is homogeneous at the scale of the sampling. The first assumption held true in our case and is generally valid even for small sample sizes in tropical forests (13). The second assumption may be an adequate approximation at small spatial scales for most lowland tropical forests (15, 40) but is unquestionably violated at larger spatial scales due to increasingly biogeographically, environmentally, and spatially structured plant communities. In such cases a large number of randomly placed sample sites are needed to capture this heterogeneity, which might adequately approximate the homogeneity assumption at a large spatial scale. Despite the large number of plots used in this study, the second assumption was not completely met at the continental scale, as shown by the incomplete leveling off of the Fisher's-alpha curves (Fig. 2a). This means that our regional Fisher's-alpha values, and thus our species richness estimates remain minimum estimates of the true values.

Lower-bound estimation of species richness. For each one degree grid cell located within 500 km of our 207 locations (excluding major water bodies), we calculated the total number of stems as the inverse distance-weighted average of the stem density observed in the five nearest locations, multiplied by the size of each grid cell (Fig. S8). Latitudinal change in grid cell size was taken into account. Because our analysis was focused on original (potential) tree species richness, we ignored recent deforestation. Uncertainty in this spatial extrapolation was assessed with a jackknife approach using 100 runs. Jackknifing removes each data point in turn and re-computes the spatial surface based on the remaining points. The differences between the original data values and the cross-validated values indicate the prediction accuracy of the surface model (Fig. S8).

The lower-bound of species richness at the pan-tropical and continental scales was then calculated using Fisher's log series, with Fisher's-alpha and total number of stems as input variables at each spatial scale. Since the total number of stems was upscaled by several orders of magnitudes, one might expect that this extreme extrapolation could potentially cause a large error in our species richness estimates. However, a simple sensitivity test, in which we varied stem numbers between 0.1 and 1.9 times the estimated values for each spatial scale (a range much larger than the observed error in our stem number predictions), showed that this extreme perturbation in stem number resulted in only a 3-11% difference in number of species predicted (Table S4).

Upper-bound estimation of species richness. Our species richness estimates are near the true value if the unidentified species (individual recorded, but not identified) in the tree inventories, which we excluded from the analyses, follow the same rank-abundance pattern as the identified species. However, if rare species are disproportionately represented in the unidentified category, our analyses may have underestimated the true Fisher's-alpha and thus the species richness estimates. Because we did not know the relative abundance of rare species in the category of unidentified individuals, we selected a dataset of multiple tree inventories with a total of 10,647 individual trees from eastern Borneo for which this information was available. All individuals in this dataset were either identified to species or only to morpho-species (unidentified). A Mann-Whitney W test showed that, not surprisingly, rare species were disproportionately represented in the unidentified category ($df = 1$, $N = 1103$, $W = 194798$, $p < 0.0001$).

Using logistic regression of species square root transformed abundance versus identification status (identified versus unidentified) we predicted the probability that a tree species would be classified as identified. This 'classification probability' was then assigned to each individual tree belonging to that species. After adding a random number between zero and one to the 'classification probability' of each individual tree, to account for the fact that even individuals belonging to common species may remain unidentified, we sorted the whole list of individuals from high to low. We produced ten such sorted lists, each time varying the random number added to the 'classification probability' of an individual tree. For each list we could then count the number of species present within any level of identified individuals and calculate a minimum Fisher's-alpha. The average of these ten minimum Fisher's-alpha values could then be compared to the actual Fisher's-alpha observed for the whole list. Dividing the actual Fisher's-alpha by the minimum Fisher's-alpha gives an inflation factor with which observed Fisher's-alpha values can be multiplied to estimate the upper boundary of Fisher's-alpha for any percentage of identified species. The inflation factor (Y) showed a power-function relationship with ratio of identified individuals (X) given by: $Y = 1.2237 \times X^{-0.767}$. Using the appropriate inflation factor, we calculated the maximum expected species richness at all spatial scales (Table 1).

Acknowledgements We like to thank Michael Fine and one anonymous reviewer for useful comments provided. The following funding agencies are acknowledged: Arnold Arboretum of Harvard University (USA); Belgian Science Policy (SD/AR/01A); Botanic Garden Meise (Belgium); Brazilian Science Council; Conselho Nacional de Desenvovimento Científico e Tecnológico of Brazil (MAUA (PRONEX-FAPEAM) (1600/2006), Universal (479599/2008-4)); British Ecological Society (4709/5747); Coordination of Improvement of Higher Education Personnel (Brazil); DEFRA Darwin Initiative (UK); Deland Award for student research, Arnold Arboretum (USA); Department of Biotechnology-National Remote Sensing Agency – India; Department of National Parks, Wildlife and Plant Conservation (Thailand); El Consejo de Ciencia y Technologia (No. 33851-B) (Mexico); Federal Ministry of Education and Research (Germany); Forest Department Sarawak (Malaysia); Fulbright Program (USA); Fundação de Amparo a Pesquisa de Minas Gerais (Brazil); German Agency for International Cooperation (GIZ) (Projects #08.7860.3-001.00 & 13.1432.7-001.00); German Science Foundation (DFG, grants CRC 552 and CU127/3-1, HO 3296/2-2, HO3296/4-1 and RU 816) (Germany); Grant Agency of the Czech Republic, project 14-36098G (Czech Republic); Intitut National pour L'etude et la Recherche Agronomiques (Democratic Republic Congo); Instituto Nacional de Pesquisas de Amazonia / Max-Planck Institute for

Chemistry (Brazil/Germany); Large-Scale Biosphere Atmosphere Experiment in Amazonia (Brazil); Missouri Botanical Garden (USA); Museo delle Scienze (Italy); National Institute for Environmental Studies (Japan); National Science Foundation (DEB-0075334, DEB-0515678; LTEB1357177, LTREB/DEB1357112, NSF DEB-0424767, NSF DEB-0639393, NSF DEB-1147429, NSF DEB-1350125, NSF-DEB-1053237, NSF DEB-0841885) (USA); NERC Human-Modified Tropical Forests programme (UK); Operation Wallacea; PAPIIT-DGAPA-UNAM (IB-200812, IN-204215) (Mexico); Pontificia Universidad Catolica del Ecuador (Ecuador); Rockefeller Foundation (USA); Royal Forest Department (Thailand); Royal Society South-East Asia Rainforest Research Programme (Project No. RS243) (UK); Rufford small grant foundation (UK); Saint Louis Zoo (USA); São Paulo Research Foundation (FAPESP 03/12595-7 - COTEC/IF 41.065/2005 - IBAMA/CGEN 093/2005) (Brazil); Smithsonian Tropical Research Institute (USA); Society of Systematic Biologists (UK); Systematics Association (UK); Swiss National Science Foundation; Tropenbos International; University of Minnesota (USA); USDA Forest Service, International Institute of Tropical Forestry (USA); USAID (USA); World Wildlife Fund (USA). Data & comment acknowledgements: Part of the data in this publication were provided by the Tropical Ecology Assessment and Monitoring (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. The Center for Tropical Forest Science network has received major support from the Smithsonian Institution – particularly the Smithsonian Tropical Research Institute, the Arnold Arboretum of Harvard University, the National Science Foundation (multiple grants), the Rockefeller Foundation, the John Merck Fund, the John D. and Catherine T. MacArthur Foundation, the Andrew W. Mellon Foundation, the Frank Levinson Family Foundation, the HSBC Climate Partnership, the Bromley Charitable Trust, John Swire & Sons Pty Ltd. and Small World Institute Fund. We also thank David Newbery, Thorsten Wiegand, George Weiblen and Vojtech Novotny, Punchi Manage Saranga Amila Ruwan, Stuart Davies, Miriam van Heist, Terese Hart, Helen Murphy, Kazuki Miyamoto, Sylvester Tan, Edmund Tanner, Mauricio Alvarez, Ana Andrade, Phourin Chhang, George Chuyong, Indiana Coronado, Chang-Fu Hsieh, Shawn Lum, Jean-Remy Makana, Cao Min, Xiaoxue Mo, Meyner Nusalawo, Atila Oliveira, Nigel Pitman, Lars Schmidt, Lee Sing Kong, Takuo Yamakura, Nicole Zweifel..

1. Pimm SL, et al. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science* 344: 1246752.
2. Kier G, et al. (2005) Global patterns of plant diversity and floristic knowledge. *J Biogeogr* 32: 1107-1116.
3. Mutke J, Barthlott W (2005) Patterns of vascular plant diversity at continental to global scales. *Biol Skr* 55: 521-531.
4. Kreft H, Jetz W (2007) Global patterns and determinants of vascular plant diversity. *Proc Natl Acad Sci USA* 104: 5925-5930.
5. Joppa LN, Roberts DL, Pimm SL (2011) How many species of flowering plants are there? *Proc Roy Soc B* 278: 554–559.
6. Erwin TL (1982) Tropical forests: their richness in Coleoptera and other arthropod species. *Coleopts Bull* 36: 74-75.
7. Odegaard F (2000) How many species of arthropods - Erwins estimate revisited. *Biol J Linnean Soc* 71: 583-597.
8. Fine PVA, Ree RH (2006) Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *Am Nat* 168: 796-804.
9. Pan Y, Birdsey R, Fang J, et al. (2011) A large and persistent carbon sink in the world's forests. *Science* 333: 988-93.
10. Chazdon RL (2008) Beyond deforestation: restoring forests and ecosystem services on degraded lands. *Science* 320: 1458-1460.
11. Basset Y, et al. (2012) Arthropod diversity in a tropical forest. *Science* 338: 1481–1484.
12. Feeley KJ, Silman MR (2011) Keep collecting: Accurate species distribution modelling requires more collections than previously thought. *Divers Distrib* 17: 1132-1140.
13. Condit R, et al. (1996) Species-area and species-individual relationships for tropical trees: a comparison of three 50-ha plots. *J Ecol* 84: 549-562.
14. Xu H, Liu S, Li Y, Zang R, He F (2012) Assessing non-parametric and area-based methods for estimating regional species richness. *J Veg Sci* 23: 1006-1012.
15. ter Steege H, et al. (2013) Hyperdominance in the Amazonian tree flora. *Science* 342: 1243092.
16. Fisher RA, Corbet AS, Williams CB (1943) 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.
17. Gentry AH (1988) Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Ann Mo Bot Gard* 75: 1-34.
18. Gentry AH (1992) Tropical forest biodiversity: distributional patterns and their conservation significance. *Oikos* 63: 19-28.
19. Morley RJ (2000) *Origin and Evolution of Tropical Rain Forests.* Wiley, Chichester, UK.
20. de Bruyn M et al. (2014) Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Syst Biol* 10.1093/sysbio/syu047.
21. Richards PW (1973) Africa, the 'odd man out'. *Tropical Forest Ecosystems of Africa and South America: a Comparative Review,* eds Meggers BJ, Ayensu ES, Duckworth WD (Smithsonian Institution Press, Washington, DC, USA), pp 21–26.
22. Parmentier I, et al. (2007) The odd man out? Might climate explain the lower tree alpha-diversity of African rain forests relative to Amazonian rain forests? *J Ecol* 95: 1058-1071.
23. Anhuf D, et al. (2006) Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeogr Palaeocl* 239: 510-527.
24. Cannon CH, Morley RJ, Bush ABG (2009) The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. *Proc Natl Acad Sci USA* 106: 11188-11193.
25. Hubbell SP (2013) Tropical rain forest conservation and the twin challenges of diversity and rarity. *Ecol Evol* 3: 3263-3274.
26. Howe HF (2014) Diversity storage: Implications for tropical conservation and restoration. *Global Ecol Conserv* 2: 349-358.
27. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25: 1965-1978.
28. FAO (2002) *Terrastat, Global Land Resources GIS Models and Databases for Poverty and Food Insecurity Mapping*. Land and Water Digital Media Series 20.
29. Brose U, Martinez ND, Williams RJ (2003) Estimating species richness: Sensitivity to sample coverage and insensitivity to spatial patterns. *Ecology* 84: 2364–2377.
30. Chao A, Jost L (2012) Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* 93: 2533–2547.
31. Chao A, Shen TJ (2010) *Program SPADE: species prediction and diversity estimation. Program and user's guide*. CARE, Hsin-Chu, Taiwan.
32. Chao A, Colwell RK, Lin C-W, Gotelli NJ (2009) Sufficient sampling for asymptotic minimum species richness estimators. *Ecology* 90: 1125–1133.
33. McGill B et al. (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol Lett* 10: 995–1015.
34. Ulrich W, Ollik M, Ugland KI (2010) A meta-analysis of species–abundance distributions. *Oikos* 119: 1149–1155
35. May RM (1975) *Patterns of species abundance and diversity*. *Ecology and evolution of communities*, eds Cody ML, Diamond JM Cambridge Univ. Press, Cambridge, pp 81–120.
36. MacArthur R (1957) On the relative abundance of bird species. *Proc Natl Acad Sci USA* 43: 293–295.
37. Preston FW (1948) The commonness and rarity of species. *Ecology* 29: 254–283.
38. Prado PI, Miranda MD (2014) *Package 'sads': Maximum Likelihood Models for Species Abundance Distributions*. CRAN http://piklprado.github.io/sads.
39. Burnham KP, Anderson DR (2002) *Model selection and multimodel inference, A practical Information-Theoretic approach*. Second Edition. Springer Science + Bussines Media Inc. USA.
40. Pitman NCA, et al. (2001) Dominance and distribution of tree species in Upper Amazonian terra firme forests. *Ecology* 82: 2101-2117.

Figure 1. Overview of sample locations and their floristic affinities (point colors correspond to scores on the first DCA-axis with similar colours indicating similar generic composition, while the lines indicate the floristic affinities as determined by cluster analysis).

Figure 2. Increase in Fisher's-alpha with (a) increasing numbers of locations (average of 50 replicates per region with random input order of locations), i.e. regional diversity and (b) increasing distance around locations (based on 50 replicates per region each with a randomly selected starting location), i.e. species turnover. Error bars indicate standard deviation among location re-orderings. Fisher's-alpha can decline if the number of stems added to the sample increases disproportionally to the number of new species detected.

==EndOfDocument==