Consistent patterns of common species across tropical tree communities

Tropical forests are a crucial component of the Earth system; they cover around 10% of the Earth's land surface but contribute approximately 33% of terrestrial net primary productivity. They account for around 40% of the carbon stored in live vegetation and are globally important carbon sinks. Tropical forests are also extraordinarily biodiverse, harbouring two-thirds of all known species and the majority of the world's biodiversity hotspots. Of note, as many tree species can be found in a single hectare of tropical forest as in the entire native Western European tree flora. Recent estimates suggest that there are approximately 37,900 named tropical tree species in the scientific literature, with potentially thousands more yet to be identified by scientists. This extraordinary diversity means that little is known about the biology of the vast majority of tropical tree species. Our understanding of tropical forest ecology, productivity and carbon storage and how they may respond to environmental change is hindered by this lack of knowledge. This limited understanding also curtails scientific input into land use, biodiversity, climate and other forest-related policy and management.

Our understanding of tropical forests may improve through a focus on the most common tree species. This is a promising avenue, given that species abundance distributions (SADs) showing a modest number of common species and much larger numbers of rare species have been documented across taxa globally. Indeed, analyses of tropical forest inventory data from Amazonia, Africa and Southeast Asia have shown that a relatively small number of common species comprise a majority of trees in the region. However, whether such patterns hold in other tropical forests is unknown, as there have been no comparable analyses for African or Southeast Asian tropical forests. Perhaps, given the substantial differences in total tree species richness, forest structure, contemporary climate and biogeographic and human-occupancy histories among continents, important contrasts in patterns of common species would be expected. Alternatively, if the same processes or mechanisms apply to all tropical forests, highly consistent patterns may be expected. Crucially, if a tractably modest number of common species do comprise the majority of tropical trees on Earth, this could open new ways of understanding tropical forests by investigating the ecology of the common species.

Cross-continental comparisons of common species patterns are complicated by unresolved differences in the results published Amazon forest studies. Estimates of hyperdominance—describing the minimum number of species required to account for 50% of all trees in a sample—range from 1.4% to 8.2% of the total number of species found in each of the Amazon forest datasets analysed. Therefore, here we: (1) investigate sample-related biases and standardize our sampling to enable meaningful comparisons among datasets; (2) test whether patterns of hyperdominance differ across Amazonia, Africa and Southeast Asia; (3) extrapolate our results to assess how many species comprise half of all Earth's tropical trees; (4) assess species abundance patterns, with differing classifications of 'common species' beyond hyperdominance; and (5) use resampling techniques to assess which sampled species are likely to be hyperdominant.

We analyse species abundance data from networks of inventory plots across three continents. We limit our analysis to closed canopy plots across three continents. We limit our analysis to closed canopy plots across three continents. We limit our analysis to closed canopy plots across three continents. We limit our analysis to closed canopy plots across three continents. We limit our analysis to closed canopy plots across three continents. We limit our analysis to closed canopy plots across three continents. We limit our analysis to closed canopy plots across three continents.
structurally intact old-growth tropical forests. For Amazonia, defined as the lowland Amazon Basin and Guiana Shield, we use the Amazon Tree Diversity Network and RAINFOR datasets (\(n = 1,097\) plots). For Africa, encompassing West, central and East Africa, we use the African Tropical Rainforest Observatory Network (AfriTRON)\(^6\), Central African Plot Network, and two smaller networks\(^7\) (\(n = 368\) plots). For Southeast Asia, defined as extending from Myanmar in the West to Sulawesi in the East, we use a tree diversity and carbon monitoring network (\(n = 103\) plots). We limit our analysis to trees with trunk diameter of at least 10 cm at breast height (1.3 m along the stem or above any buttresses or deformities), the widely used minimum size for inventorying tropical trees. The combined dataset includes 1,003,805 trees, of which 93.3% are identified to species (Fig. 1 and Extended Data Table 1).

**Consistent patterns of commonness**

The Africa, Amazonia and Southeast Asia datasets differ in the number and size of plots sampled and the number of trees sampled (Extended Data Table 1). We therefore excluded small plots (below 0.9 ha; Extended Data Fig. 1 and Methods) and used rarefaction—that is, repeated random subsampling of plots to comparable numbers of trees—to standardize sampling across the three datasets (Fig. 2).

Rarefying to a common sample size of 77,587 stems, the size of the Asia dataset (equivalent to 150, 116 and 103 plots in Africa, Amazonia and Southeast Asia respectively), we find that 77 species (95% confidence interval: 62–92) in Africa comprise 50% of individual trees, compared with 174 species (95% confidence interval: 134–215) in Amazonia and 172 species (95% confidence interval: 125–217) in Southeast Asia (Table 1 and Fig. 2). However, the substantially lower number of hyperdominant species in Africa compared with Amazonia and Southeast Asia scales with the substantially lower number of total species. We find just 1,132 species in our standardized 77,587 tree sample in Africa, compared with 2,565 and 2,585 species in Amazonia and Southeast Asia, respectively for the same sample size. Consequently, percentage hyperdominance is statistically indistinguishable among the continents at 6.79% (95% confidence interval: 5.39%–8.20%), 6.80% (95% confidence interval: 5.24%–8.36%) and 6.65% (95% confidence interval: 4.59%–8.71%) in Africa, Amazonia and Southeast Asia, respectively (Table 1). This consistency is not affected by the aggregated spatial distribution of plots within each region (Extended Data Fig. 2) and holds true for analyses based solely on 1-ha plots (Methods). Thus, once sampling is standardized, there is marked pan-tropical consistency in the proportion of the total number of tree species accounted for by the most common species.

The consistency of commonness is not limited to defining common species as those that account for 50% of all individual trees in a dataset. The proportions of the total number of species required to account for thresholds between 10% and 90% of individual trees are also highly consistent across the rarefied data for the three continents (Fig. 3 and Extended Data Table 3). Thus, the data from the three continents appear to result from the same underlying statistical distribution.

Our rarefaction analysis shows that the number of hyperdominants, the total number of species and the percentage hyperdominance are dependent on sample size. This is because as plots—and therefore trees—are added to the sample, increasing numbers of rare species start to appear. Meanwhile, most common species have, by definition, already appeared, but their abundances increase. Thus, with increasing sample size, the number of hyperdominants increases, but at an ever-decreasing rate that tends towards saturation (Fig. 2 and Extended Data Fig. 3). The total number of species increases at a decreasing rate with increasing sample size, without apparent saturation. Therefore, as sample sizes increase, the percentage hyperdominance decreases gradually, but does not appear to saturate (Fig. 2 and Extended Data Fig. 3). This sample size dependence is likely to explain the published differences in percentage hyperdominance in Amazonian forests, which follow expectations given the sample size in each study.\(^6,20,21\)

Amazonia and Southeast Asia show remarkably similar patterns of commonness and diversity. The rarefaction curves of the number of species accounting for 50% of all trees (Fig. 2a), total number of species (Fig. 2b), percentage hyperdominance (Fig. 2c) and Fisher’s \(\alpha\)—the parameter of the log series distribution shown to best describe tropical tree species abundance distributions\(^21\) (Fig. 2d)—are almost identical between the two datasets. Furthermore, the numbers of species required to account for any threshold between 10% and 90% of trees in the respective rarefied samples of 77,587 trees are statistically indistinguishable (Table 1 and Extended Data Tables 2 and 3). This equivalence in overall tropical forest diversity patterns between these similarly species-rich regions is particularly striking given their very different biogeographic, climatic and anthropogenic histories, and the fact that Amazonia is one large contiguous region, whereas Southeast Asia is a series of islands and island-like regions.
In contrast to the similarity between Amazonia and Southeast Asia, our results provide sample size-corrected validation of the ‘odd-one-out’ observation of much lower tree species richness in Africa compared with Amazonia and Southeast Asia. Here we add a similar odd-one-out observation of a much lower number of common species in Africa than in Amazonia and Southeast Asia. However, in combination these two results lead to an almost identical percentage hyperdominance in the African, Amazonian and Southeast Asian rarefied data. This consistency extends to the proportion of species required to account for all thresholds between 10% and 90% of trees in the rarefied data (Fig. 3 and Extended Data Table 3). This pan-tropical invariance recasts the tropical forests of Africa from ‘odd’ in terms of species richness to statistically indistinguishable from those in Amazonia and Southeast Asia in terms of proportional patterns of abundance. Overall, using standardization by rarefaction, we find consistent patterns of species abundance across Africa, Amazonia and Southeast Asia.

Scaling to the study region

Next, we estimate commonness patterns in each of our three study regions: Africa, Amazonia and Southeast Asia. We extrapolate log series fits to the empirical Africa, Amazonia and Southeast Asia data-sets (Extended Data Fig. 4), including a correction to account for the clumped spatial occurrence of species, to the total number of trees with trunk diameter of at least 10 cm in each study region. We estimate that just 104 species (95% confidence interval: 101–107) account for 50% of the 113 billion trees in Africa’s closed canopy tropical forests (Table 2). We also estimate that just 299 species (95% confidence interval: 295–304) account for 50% of the 344 billion trees in Amazonia’s closed canopy tropical forest, and 278 (95% confidence interval: 268–289) account for 50% of the 129 billion trees in Southeast Asia’s closed canopy tropical forests (Table 2). Our results from Amazonia match those derived using a different extrapolation approach.

Our extrapolations again outline consistent percentage hyperdominance: just 2.2% of African, 2.2% Amazonian and 2.3% of Southeast Asian species account for 50% of all trees with trunk diameters of at least 10 cm in each region (Table 2). The dominant proportions of total species required to account for 10% to 90% of trees are also very similar across continents (Fig. 3 and Extended Data Table 3). The lower percentage dominance values from the extrapolated data compared...
with those from the rarefied data are consistent with the pattern, described above, of many more rare species being added as the number of trees increases while many fewer common species are added (Fig. 2). Overall, the extrapolated results show that there is a tractable number of common species in tropical forests in Africa, Amazonia and Southeast Asia.

### Scaling to the tropics

We next estimate the number of common tropical tree species on Earth by multiplying the pan-tropical proportion of common species by the total number of tropical tree species on Earth. Our results suggest a pan-tropical hyperdominant percentage of 2.24% (Table 2). However, our extrapolations cannot provide an estimate of the total number of tropical tree species because we do not—for this study—have data from all tropical regions, including a lack of data from Central America, New Guinea and Micronesia. Furthermore, there is no consensus estimate of the total number of tropical tree species on Earth.

A compilation of lists of species known to science suggests a total of 60,065 tree species globally. Tropical forest biomes likely comprise 63% of this list (E. Beech, personal communication, 2021), implying that there are around 37,900 known tropical tree species. This minimum estimate does not account for species that are yet to be identified and described by scientists. An alternative extrapolation method estimated that there are 46,900 species for the closed canopy tropical forest biome (range 40,500–53,300 species), implying that there are 9,000 yet-to-be-identified species. This is in agreement with a recent global study suggesting that there are around 9,200 tree species remaining yet to be formally named, almost all in the tropics. Thus, together, these studies suggest there are likely to be approximately 47,000 tropical tree species in the world’s closed canopy tropical forests.

Our best estimate is that 1,053 tree species (2.24% of 47,000 species) account for half of Earth’s 800 billion trees with trunk diameters of at least 10 cm found in the closed-canopy tropical forest biome. Although the true number may be lower or higher, the conclusion that a tractable number of species dominate tropical forests is clear. Some of these species are likely to be extraordinarily common: our best estimate is that just 61 species account for 80 billion individual trees (0.13% of 47,000 species). At the other end of the spectrum, we estimate that the rarest approximately 39,500 species account for just 80 billion trees, or 10% of individuals. Meanwhile, the other 90% of all trees are estimated to belong to just 7,487 species (15.93% of 47,000 species). Thus, these results open the possibility of focusing efforts on understanding the biology of a tractable number of species in tropical forests to approximate the whole stand.

### Identifying the most common species

Our analyses showing that 104, 299 and 278 common species account for 50% of the trees in our African, Amazonian and Southeast Asian study regions, respectively, do not yield a list of named species. To assess which named species are likely to be hyperdominant, we use a subsampling procedure similar to the rarefaction methodology above. We randomly subsample from approximately 10,000 trees per subsample (drawn by plot) and increase the size of the subsample in 10,000-tree increments until the size of each regional dataset is reached, and repeat this process 100 times. For each sampled increment of 10,000 trees we then calculate the proportion of random subsamples in which each species qualifies as hyperdominant (Supplementary Table 1). We then assign the species to one of four groups:

1. Both hyperdominant in the full data and hyperdominant in the majority of subsamples even at very small sample sizes. These 50, 95 and 105 species in our Africa, Amazonia and Southeast Asia datasets, respectively, represent 3.5%, 2.1% and 4.1% of sampled species in each dataset. These species are likely to be geographically widespread and abundant.
2. Both hyperdominant in the full data and hyperdominant in the majority of subsamples, but at the smallest sample sizes only occasionally hyperdominant. These 32, 129 and 67 species in our Africa, Amazonia and Southeast Asia datasets, respectively, represent 2.3%, 2.9% and 2.6% of sampled species in each dataset. These species are likely to be geographically widespread but not always abundant.
3. Not quite hyperdominant in the full data, but hyperdominant in a substantial proportion of subsamples. These 102, 339 and 200 species in our Africa, Amazonia and Southeast Asia datasets, respectively, represent 7.2%, 7.5% and 7.7% of sampled species in each dataset. These species are probably locally abundant but not necessarily geographically widespread.
4. Not hyperdominant in the full data and almost never hyperdominant in the subsamples. These 1,232, 3,929 and 2,213 species in our Africa, Amazonia and Southeast Asia datasets, respectively, represent 87%, 87.5% and 85.6% of sampled species in each dataset. These species are probably neither geographically widespread nor abundant.

---

Table 2 | Extrapolated tree species hyperdominance results for African, Amazonian, Southeast Asian tropical forests at the regional scale

<table>
<thead>
<tr>
<th>Region</th>
<th>Number of hyperdominants</th>
<th>Total species</th>
<th>Hyperdominant percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazonia</td>
<td>299 [295, 304]</td>
<td>13,826 [13,615, 14,036]</td>
<td>2.16</td>
</tr>
<tr>
<td>Southeast Asia</td>
<td>278 [268, 289]</td>
<td>11,963 [11,451, 12,475]</td>
<td>2.32</td>
</tr>
</tbody>
</table>

*Calculated as the sum of the number of hyperdominants and total species across the three major tropical forest regions with hyperdominance percentage derived therefrom. Prediction intervals (in brackets) combine uncertainty from the standard error of predicted means and the residual s.d. of the regression of the bias correction fit.
We suggest that if all trees in a region were sampled, the hyperdominant species would be drawn from the first three groups, which are listed in Supplementary Table 2. This candidate list of 1,119 hyperdominant species contains 184 species in Africa, 563 species in Amazonia and 372 species in Southeast Asia, with no species appearing on more than one region's list. Thus, the list of species that are likely candidates for hyperdominance is manageable small.

There is uncertainty in our candidate hyperdominant list owing to the limitations of the underlying samples of plots across the landscape. Specifically, some species that always have low local abundance but are geographically widespread and lack habitat restrictions may require larger sample sizes for their hyperdominance to become clear. Similarly, species that combine low local abundance and habitat specificity may be missed. If the distribution and extent of specialist habitat is great enough to result in hyperdominance of specialists but is not sufficiently captured in our sampling, such species might not appear in our candidate list. By contrast, some species in our candidate hyperdominant list will not be true hyperdominants. Of particular note, some apparently common species may actually comprise a group of cryptic species, with none of these cryptic species being hyperdominant by itself28–33. However, the striking similarity in species abundance patterns across the Africa, Amazonia and Southeast Asia datasets, despite differing sampling intensity on each continent, suggests that these potential limitations do not substantially affect the overall patterns found. We therefore expect a high overlap between our list of candidate hyperdominant species and eventual elucidation of the actual hyperdominants of these three regions and the pan-tropics.

Our list of 1,119 candidate hyperdominant species represents a tractable number of species on which to prioritize autecological research. Indeed, given their commonness, ecological data already exists for many of these species: 95% have some autecological data recorded in a large global database34; 83% have at least 10 different types of measurement, typically including their growth form, maximum height, wood density and aspects of leaf chemistry. This indicates that these species are already relatively well known. Therefore, only limited additional data may be required to open new approaches to better understanding tropical forests through their most common tree species, including how they may react to today’s era of rapid global environmental change.

Discussion

Charles Darwin wrote in The Origin of Species that “rarity is the attribute of a vast number of species of all classes and in all countries”35. If this is the case, then common species are themselves rare. Our results concur: despite their formidable diversity, the trees in tropical forests fit the ‘rare is common, common is rare’ pattern36 which has been documented in many other taxa37–39. Beyond this, our analyses reveal highly consistent patterns of commonness across three major tropical forest regions. Notably, despite substantial inter-continental variation in biogeographic history, contemporary environment, forest structure and species composition, we have found an emergent property of the tropical forest system. For the trees that structure tropical forests, a consistent ~2.2% of the total species pool accounts for 50% of all individual trees in Africa, Amazonia and Southeast Asia. This consistency is all the more notable given relatively lower tree species richness of African tropical forests compared with Amazonian and Southeast Asian forests, probably owing to higher extinction rates in African forests, with evidence of major losses of African species at the Oligocene–Miocene boundary40, and contractions of rainforest area due to drier conditions during repeated glacial–interglacial cycles over the past 2.6 million years39.

We find common diversity patterns despite the very different histories of human occupancy in Amazonian, African and Southeast Asian tropical forests41. The relatively recent arrival of humans in Amazonia approximately 20,000 years ago has been linked to greater Pleistocene extinctions, in contrast to much longer human occupancy in the tropical forests of Africa and Southeast Asia42. Some have also suggested that Amazonian forest composition was altered by humans through the incipient domestication of tree species, increasing the abundance of a small number of favoured species43. Others have reported large areas of deforestation associated with the African Iron Age44. How can such different human histories result in near-identical patterns of tree species dominance? The most parsimonious explanation is that the system tends to return to a state with a similar species abundance pattern.

Nevertheless, consistent patterns of commonness do not necessarily imply the same causal mechanisms. The ubiquity of the broad ‘rare is common, common is rare’ pattern in ecology, which is also found in non-biological complex systems45, means inferences as to the cause of this broad pattern are challenging46–48. Although combinatoric methods49 and models that maximize the entropy of information50,51 both produce the ubiquitous ‘reverse lazy J’ pattern, empirical observations show fewer common species and more rare species than expected by statistical controls alone52. Similarly, neutral models produce the same broad pattern, but produce too few individuals of the most common Amazonian tree species53. This suggests that biological mechanisms influence tree community assembly to produce a consistent proportion of common species across continents.

Recent analyses have revealed that the same few families contribute most of the species richness in Africa and Amazonia54, which when combined with analyses showing that more diverse families have more common species55, may indicate a role for deep evolutionary mechanisms driving the patterns we find. Yet, considering the substantially smaller regional species pool in Africa compared with Amazonia and Southeast Asia, one might expect differing continental patterns of species dominance if evolutionary drivers were the primary mechanism, not the highly consistent patterns that we find. Similarly, if environmental filtering were a key mechanism, the different contemporary environments, with Africa much drier on average than the other two continents56, and Southeast Asia consisting of scattered island-like areas of forest compared with the contiguous forested region of Amazonia, would also imply differing continental patterns of species dominance, not the near-identical patterns that we find. These constraints limit the potential mechanisms that could apply across our three-continent context.

One potential cross-continental mechanism is dispersal limitation, where the dispersal capabilities of species result in some suitable habitat patches remaining unoccupied. Another mechanism is density- or distance-dependent mortality, which appears widespread across tropical forests48. Here, specialist species-specific natural enemies such as pathogens and herbivores reduce seed or juvenile conspecific survival rates near conspecific adults or in areas of high juvenile conspecific density, thereby reducing competitive exclusion and contributing to the maintenance of high tree species richness in tropical forests57. It is possible that common species have largely evaded density- and/or distance-dependent mortality. Analyses showing that species abundance can be either high or low within given genera58 support this hypothesis. Further progress on putative mechanisms can be made, for example, by exploring whether ecological or functional traits differ between common and rare species, and assessing the consistency of any differences among tropical continents59. Although deducing mechanisms is complex, the identification of a tractable number of common species in tropical forests will facilitate progress in understanding of tropical forests beyond species abundance distributions.

Refining our results, particularly the naming of common species, requires improved sampling of tropical forests, both in terms of geographic scope and taxonomic identification of trees within plots. Expanding sampling to include Central America, New Guinea, Micronesia and other regions would improve the generality of our results. Better identifying trees in existing plots would increase the utility of available samples: in our Southeast Asia region we excluded 142 plots


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

Open Access. *This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.*

© The Author(s) 2024
Methods

Data compilation and pre-processing
We collated data from forest inventory plots ≥0.2 ha in size, situated in structurally intact (no detectable past logging or fire), closed canopy (not dry forest or savanna) tropical forest, with enumeration of all stems ≥10 cm diameter, in which ≥80% of stems are identified to the species level. Following Sullivan et al. 63, small (≤0.5 ha) plots within 1 km of each other were grouped for analysis to minimize the effect of stochastic tree fall events in smaller areas. 3 These criteria allow direct comparisons to be made with hyperdominance results from Amazonia 64. The data from each continent comprise the following:

Africa: 483 plots, covering a total of 504 ha (mean plot area 1.04 ha, median 1 ha, range 0.2–10 ha). These data are from four sources: 299 plots from the African Tropical Rainforest Observatory Network 3,47 (AfriTRON: www.afritron.org, accessed 1 March 2020), curated at http://www.ForestPlots.net 4; 127 plots from the Central African Plot Network (https://central-african-plot-network.netlify.app); 52 plots from the TEAM network 2; and 5 × 1 ha plots from 5 different soil types, extracted from one 50-ha plot in Korup, Cameroon from the SIGEO/CTFS network 3.

Amazonia: 1,417 plots, covering a total of 1,591 ha (mean plot area 1.12 ha, median 1 ha, range 0.1–78.8 ha) from the Amazon Tree Diversity Network (ATDN: http://atdn.myspecies.info/), includes plots from the RAINFOR network, accessed 8 January 2020.

Southeast Asia: 230 plots, covering a total of 202 ha (mean plot area 0.88 ha, median 0.49 ha, range 0.21–4.5 ha). These data are from two sources: 143 plots from Silk et al. 4,25 – a decrease from the published Indo-Pacific dataset in Silk et al. 4 due to our ≥80% species identification criterion and our Southeast Asia study region excluding Australia, India, and Papua New Guinea; and 87 plots from the T-Forces network 44 curated at http://www.ForestPlots.net, accessed 03/02/2021.

Species names were checked for orthography and standardized (synonyms identified from the reference databases corrected to their accepted names) using the African Flowering Plants Database (https://www.ville-ge.ch/musinfo/bd/cjb/africa), Taxonomic Name Resolution Service 6, and Asian Plant Synonym Lookup (F. Silk, personal communication), for Africa, Amazonia and Southeast Asia, respectively. Trees not identified to species level (7.3%, 6.3% and 8.4% of stems in the Africa, Southeast Asia datasets respectively) were classed as ‘indeterminate’ (Indet). Indet stems contributed to plot-level and dataset-wide stem abundance totals but are necessarily absent from species totals.

For the purposes of our study we delimited tropical forests according to the ‘tropical and subtropical moist broadleaf forests’ biome delineation from the World Wildlife Fund ecoregion map 6,7. The total number of tropical trees ≥10 cm trunk diameter in each of our regions was then estimated by summing tree abundances in countries in which we have at least one sampled plot from the ‘map of Global Tree Density’ 8 (derived from 429,775 ground-based estimates of tree density) and masking according to the ‘tropical and subtropical moist broadleaf forests’ borders using ArcGIS v3.10.10. Thus, we estimate that there are ~92 billion, ~331 billion trees, and ~217 billion trees in our Africa, Amazonia, and Southeast Asia regions, respectively, totalling 640 billion trees. Including abundance from countries in the ‘tropical and subtropical moist broadleaf forests’ biome in which we have no sampled plots, we estimate ~799 billion total trees across all of Earth’s moist tropical forests.

Data format, commonness and diversity parameters
The species abundance distribution (SAD), defined as a vector of abundances (number of individuals observed) of all species encountered in a community 7, formed the basis for our analyses of the three tropical forest datasets. For each dataset, we tallied the number of trees of each species in each plot to give plot-level SADs and combined these SADs across all plots to get regional-level abundance matrices with rows representing plots, columns representing species, and entries representing the abundance of each species in each plot. To capture patterns of commonness and species composition we calculated the number of hyperdominants (H#), defined as the minimum number of species required to account for 50% of the population of an assemblage, hyperdominant species identities, total number of species (TS), hyperdominant percentage of total species (H% = H#/TS) and Fisher’s α (ref. 68). To investigate the sensitivity of results to the ‘hyperdominant’ definition of the most common species, we looked beyond the 50% threshold used for hyperdominance, at the minimum number of species required to account for 10%, 20%, 30%, ..., 90% of the population, here termed ‘dominants’.

Sampling standardization, subsampling and comparison of continental data
We identified variations in the number of plots, stems, and species, and the size and spatial clustering of plots as potential confounding factors liable to skew dominance and diversity results from our regional datasets and impede rigorous comparisons between them. We used sample-based rarefaction to quantify and account for the effect of differences in sample size (number of plots and stems) on our diversity measures of interest; namely species richness, number, ranking and identity of hyperdominants, hyperdominant percentage of total species, and Fisher’s α. To quantify the effect of plot size, which is smaller in Southeast Asia data (mean 0.88 ha, median 0.49 ha) than in Amazonia and Africa data (both mean ~1 ha, median 1 ha) we compared results from the full data to those from plots >0.9 ha. We found that small plots (<1 ha) inflate per-plot species totals relative to larger plots (because the rate of encountering new species is higher the smaller the plot size; Extended Data Fig. 1), so we limited our analyses to plots >0.9 ha to enable like-for-like comparison.

For Africa, we retained 368 plots covering 450 ha (mean plot area 1.22 ha, median 1 ha, range 0.92–10 ha; 2% of plots 0.9–0.99 ha, 88% of plots 1 ha, 8% of plots 1.01–5 ha, 1% of plots >5 ha) with mean temperature of 24.3 °C (range 16.2–27.6 °C), mean annual precipitation 1,802 mm yr⁻¹ (range 1,066–2,747 mm yr⁻¹), and mean elevation of 511 m above sea level (range 41–2,070 m) per WorldClim 5. For Amazonia we retained 1,097 plots covering 1,434 ha (mean plot area 1.31 ha, median 1 ha, range 0.9–78.8 ha; 2% of plots 0.9–0.99 ha, 90% of plots 1 ha, 7% of plots 1.01–5 ha, 1% of plots >5 ha) with mean temperature of 25.7 °C (range 20.1–27.5 °C), mean annual precipitation 2,937 mm yr⁻¹ (range 1,119–4,284 mm yr⁻¹), and mean elevation of 154 m (range 0–1,412 m). For Southeast Asia we retained 103 plots covering 164 ha (mean plot area 1.59 ha, median 1 ha, range 0.96–4.5 ha; 1% of plots 0.9–0.99 ha, 48% of plots 1 ha, 52% of plots 1.01–5 ha, 0% of plots >5 ha) with mean temperature of 24.3 °C (range 20.1–27.5 °C), mean precipitation 2,680 mm yr⁻¹ (range 1,466–3,941 mm yr⁻¹), and mean elevation of 288 m (range 10–934 m). We assessed if the remaining differences in plot size affected the results, using only the 1 ha plots from Africa (n = 323) and Amazonia (n = 988), rarefied to the size of the Asia dataset, again finding near-identical per cent hyperdominance on the two continents (Africa: 7.30%, 95% confidence interval: 6.56–8.04; Amazonia: 7.35%, 95% confidence interval: 6.61–8.10).

To quantify the effect of the spatial clustering of plots, we compared results from the full Amazonia data, as the largest dataset, to those from subsets of the Amazonia data in which 1,2,3,...,10 plots were sampled from each spatial cluster. We found that spatial clustering had a negligible and not statistically significant effect on hyperdominant percentage and fitted values of Fisher’s α (Extended Data Fig. 2). Therefore, we retain all plots for our analyses to maximize sample sizes. Computation of percentage hyperdominance and dominance accounts for the effects of variations in species richness on the number of hyperdominants and dominants.

For sample-based rarefaction, 200 subsamples of 1, 2, ..., Np plots were drawn, without replacement, from the Np total number of plots in the pth dataset, the stems contained in each subsample were pooled,
and the mean total species, number of hyperdominants, hyperdominance percentage, and Fisher’s α were calculated across the subsamples. Similarly, we tallied the number of subsamples in which each species in the dataset qualified as hyperdominant at each level of subsampling and compared results between datasets at subsample sizes equating to a mean 10,000, 20,000, …, 10p, individual trees, where Ip is the total number of trees in the pth dataset. Confidence intervals were calculated as confidence interval = μ ± 1.96 × σ, where μ values are the means of the diversity metrics calculated across the 200 iterations of subsamples taken without replacement, and σ values are the s.d. of the mean of diversity metrics calculated across the 200 iterations of subsamples taken with replacement (to reduce the degree to which confidence intervals were conditional on the sample). For point estimates, all datasets were compared at the common sample size of the Southeast Asia dataset (77,587 stems equivalent to ISO, 116 and 103 plots in Africa, Amazonia and Southeast Asia, respectively).

 Extrapolation and bias correction of log series fits to the empirical data
We extrapolated our empirical SADs to SADs at the scale of the entire Amazonian, African, and Southeast Asian regional level via analytical expansion and bias correction of Fisher’s log series fits following the methodology of ter Steege et al.21 developed using the ATDN data that comprise our Amazonia dataset.

Ter Steege et al. found that simulations of sampling of plots with conspecific aggregation from log series-modelled SADs provide extremely good approximations of the processes that generate tropical forest inventory data—that is, non-random sampling of plots containing species with limited dispersal and/or ecological preferences. They further found that estimates of species richness derived from samples taken with conspecific aggregation from the simulated SADs substantially underestimated the true species richness of the simulated SADs, but that a linear relationship with low variance existed between the true and sample-derived values. Thus, although conspecific aggregation in the empirical data introduces bias in the log series-modelled SADs extrapolated therefrom, quantification and correction of the effects of this bias on regional estimates of species richness is possible. Therefore, to estimate species richness at the regional level, they fitted Fisher’s log series to empirical species abundance data, quantified the effect of conspecific aggregation on these estimates via simulation, and applied quantified corrections to give more accurate estimates of regional species richness taking into conspecific aggregation. Thus, this approach corrects for species-specific aggregation at the plot scale depending on species density.

To estimate regional numbers and proportions of dominants and hyperdominants as well as species richness, we extended the methodology of ter Steege et al.21 to log series-derived estimates of regional numbers and proportions of dominants and hyperdominants. Initially, values of Fisher’s α were fitted to the empirical species abundance vectors from each region using maximum likelihood and numerical optimization in the ‘sads’ R package22 and fits visualized with Preston plots23 and rank abundance distributions (RAD)24 (Extended Data Fig. 4).

Regional species totals S, not accounting for bias introduced by conspecific aggregation, were then estimated24 via S = α × ln (1 + 2N) with total number of trees ≥10 cm trunk diameter at the continental level (N) from the Global Tree Density map of Crowther et al.49 with each tropical region delineated within the ‘tropical and subtropical moist broadleaf forests’ biome of Olson et al.50. An inverse quantile function from the sads R package22 was then applied to generate (uncorrected) continental-scale SADs for each region using the above fitted α, estimated S and N.

For the quantification of bias and computation of corrections, we first simulated 250 log series SADs with known values of total species, Su, randomly drawn from the range of plausible regional species totals (10,000–25,000 in Amazonia and Southeast Asia; 2,000–10,000 in Africa) and N, the number of trees in each region ≥10 cm trunk diameter from Crowther et al.49. We calculated known values of numbers of hyperdominants, Hu, and percentage hyperdominance, Pu, from each of these simulated distributions. Using a negative binomial distribution to simulate conspecific aggregation per ter Steege et al.21, we then simulated j random samples of 1 ha plots from each of the 250 simulated SADs, with/without to the number of plots in the empirical data, and the expected abundance of each species in each plot equal to its mean regional density (total abundance/regional area). We then estimated (uncorrected) species richness, Su, from each of the samples by fitting Fisher’s α to the sampled data and applying the formula Su = α × ln (1 + N)/μ. From each of the samples we also derived continental-scale uncorrected SADs (see above), from which the number of hyperdominants, Hu, and percentage hyperdominance, Pu, could be directly calculated, via analytical expansion of the log series using the fitted values of α and corresponding values of Su. We then regressed the known values of Su, Hu, and Pu from the simulated SADs against the estimated (uncorrected) values Su, Hu, and Pu, from the samples drawn with conspecific aggregation across all 250 simulations—that is, fit linear models of the form Ai = m × A + c for A = S, H, P. This same procedure was also applied to the number and proportion of dominants.

Across all three regional datasets, the above procedure outlined a linear relationship with low variance between known values of species richness, number of dominants and hyperdominants, and percentage hyperdominance and dominance, and values thereof estimated from sampling with conspecific aggregation (Extended Data Fig. 5). Thus, constant terms with low variance were readily applicable to correct bias in the point estimates of species richness, number of dominants/hyperdominants, and percentage hyperdominance/dominance, derived from the empirical Africa, Amazonia, and Southeast Asia data. To capture uncertainty around each bias-corrected point estimate, prediction intervals (PI) were then developed as PI = μ + 1.96 × σ, where μ is the predicted mean value of the point estimate according to the linear regression, and σ is the PI standard error, calculated as σPI = √(σ2 + α2), where σ is the standard error of predicted means and α is the residual s.d. (and 1.96 is the 0.05 quantile of a t-distribution).

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

 Data availability
The species abundance data that support the findings of this study are available from https://doi.org/10.6084/m9.figshare.21670883 (formatting notes: a column for each species, rows for each plot, entries are the number of trees ≥10 cm diameter of each species in each plot). WorldClim bioclimatic data are available from https://www.worldclim.org/data/bioclim.html.

 Code availability
R code (version 4.3.1) to run the analyses and produce the figures and tables is available from https://github.com/declancoooper/Common-Species2022.git.

70. Prado, P. I., Miranda, M. D., Chalom, A., Prado, M. P. I. & Imports, M. & d’enseignement supérieur ARES (AFORCO project) and by the European Union through the Council VLIR-UOS (CD2018TEA459A103, FORMONCO II); by L’Académie de recherche et FED-tWIN2019-prf-075/CongoFORCE, EF/211/TREE4FLUX); by the Flemish Interuniversity Office BELSPO (SD/AR/01A/COBIMFO, BR/132/A1/AFRIFORD, BR/143/A3/HERBAXYLAREDD, Republic of the Congo (Yangambi and Yoko sites) was funded by the Belgian Science Policy Research Fellowship and a Leverhulme Trust Research Fellowship. Fieldwork in the Democratic Republic of the Congo was supported by the London Natural Environmental Research Council Doctoral Training Partnership grant (grant no. NE/L002485/1). This paper developed from analysing data from the African Tropical Rainforest Observatory Network (AfriTRON), curated at ForestPlots.net. AfriTRON has been supported by numerous people and grants since its inception. We sincerely thank the people of the many villages and local communities who welcomed our field teams and without whose support this work would not have been possible. Grants that have funded the AfriTRON network, including data in this paper, are a European Research Council Advanced Grant (T-FORCES, 291585), Tropical Forests in the Changing Earth System), a NERC standard grant (NER/A/3/2000/01002), a Royal Society University Research Fellowship to S.L.L., a NERC New Investigators Grant to S.L.L., a Philip Leverhulme Award to S.L.L., a European Union FP7 grant (GEOCARBON, 283080), Leverhulme Program grant (Valuing the Acer), a NERC Consortium Grant (TROBIT, NE/D00559/0/1), NERC Large Grant (CongoPeat, NE/R018680/1) the Gordon and Betty Moore Foundation the David and Lucile Packard Foundation, the Centre for International Forestry Research (CIFOR), and Gabon’s National Parks Agency (ANPN). This paper was supported by ForestPlots.net approved Assistant.

Acknowledgements D.L.M.C. was supported by the London Natural Environmental Research Council Doctoral Training Partnership grant (grant no. NE/L002485/1). This paper developed from analysing data from the African Tropical Rainforest Observatory Network (AfriTRON), curated at ForestPlots.net. AfriTRON has been supported by numerous people and grants since its inception. We sincerely thank the people of the many villages and local communities who welcomed our field teams and without whose support this work would not have been possible. Grants that have funded the AfriTRON network, including data in this paper, are a European Research Council Advanced Grant (T-FORCES, 291585), Tropical Forests in the Changing Earth System), a NERC standard grant (NER/A/3/2000/01002), a Royal Society University Research Fellowship to S.L.L., a NERC New Investigators Grant to S.L.L., a Philip Leverhulme Award to S.L.L., a European Union FP7 grant (GEOCARBON, 283080), Leverhulme Program grant (Valuing the Acer), a NERC Consortium Grant (TROBIT, NE/D00559/0/1), NERC Large Grant (CongoPeat, NE/R018680/1) the Gordon and Betty Moore Foundation the David and Lucile Packard Foundation, the Centre for International Forestry Research (CIFOR), and Gabon’s National Parks Agency (ANPN). This paper was supported by ForestPlots.net approved Assistant.


Competing interests The authors declare no competing interests.

Additional information Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-06820-z.

Correspondence and requests for materials should be addressed to Declan L. M. Cooper or Simon L. Lewis.

Peer review information Nature thanks Miguel Martinez-Ramos, Thorsten Wiegand and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permissions information is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Impact of plot size on rarefaction curves of total species (a) and number of hyperdominants (b) in the Asia data. Red points represent the full Southeast Asia data (mean values across iterations of subsamples), including all plots (mean plot size: 0.877 ha, median plot size: 0.5 ha); purple points represent the Southeast Asia data restricted to plots ≥0.9 ha (mean plot size: 1.59 ha, median plot size: 1 ha).
Extended Data Fig. 2 | Impact of spatial clustering of plots on rarefaction curves of hyperdominant percentage (first row) and Fisher’s Alpha (second row) in the Amazonia data. Purple points and confidence intervals represent the full data; black points and confidence intervals represent a subset of the data in which one plot is sampled from each spatial cluster of plots; other coloured points represent subsets of the data in which 2, 3, 4, ..., 10 plots (or the total number of plots in the cluster) are sampled from each spatial cluster of plots. Points give the mean values across iterations of subsamples. Confidence intervals are derived via the standard deviation across iterations of subsamples taken with replacement at each sampling point. Note that although resampling for rarefaction was done by subsampling tree inventory plots, the curves are re-plotted with an x-axis of number of stems.
Extended Data Fig. 3 | Complete rarefaction curves showing the effect of increasing sampling on the number of hyperdominants (a), total species (b), hyperdominant percentage (c), and fitted values of Fisher’s α (d). In tropical Africa (magenta), Amazonia (cyan), Southeast Asia (blue). Markers represent rarefied points (mean values across iterations of subsamples); shaded areas represent confidence intervals (CIs). Confidence intervals are derived via the standard deviation across iterations of subsamples taken with replacement at each sampling point. Note that although resampling for rarefaction was done by subsampling tree inventory plots, the curves are re-plotted with an x-axis of number of stems.
Extended Data Fig. 4 | Preston plots (top row) and rank abundance distributions (bottom row) showing the empirical species abundance distributions for Africa (left) Amazonia (middle) and Southeast Asia (right) with log series fits overlaid. Histogram bars display the empirical species abundance distributions as Preston plots (top row); black markers show the empirical species abundance distributions as rank abundance distributions (bottom row); overlaid points and lines show log series fits to empirical species abundance distributions in Africa (magenta), Amazonia (cyan), and Southeast Asia (blue).
Extended Data Fig. 5 | Bias correction of estimates of species richness (first column), number of hyperdominants (second column), percentage hyperdominance (third column) for the Amazonia (first row), Africa (second row) and Southeast Asia (third row) datasets. X-axes show estimated values derived from samples of the simulated communities taken with conspecific aggregation. Y-axes show true values of the simulated communities. Points show estimated true values for each of the 250 simulated communities. 1:1 equivalence shown by straight line in each plot. For number of hyperdominants and total species plots, simulated communities containing 100 to 25,000 species in Amazonia and Southeast Asia, 100 to 10,000 species in Africa are shown. For percentage hyperdominance, simulated communities containing 10,000 to 25,000 species in Amazonia and Southeast Asia, 2,000 to 10,000 species in Africa are shown.
Extended Data Table 1 | Empirical summary statistics and hyperdominance results for tree species data in Africa, Amazonia, and Southeast Asia

<table>
<thead>
<tr>
<th></th>
<th>Plots</th>
<th>Stems</th>
<th>%ID</th>
<th>H#</th>
<th>TS</th>
<th>H%</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>368</td>
<td>189,948</td>
<td>92.7</td>
<td>82</td>
<td>1,416</td>
<td>5.79</td>
<td>210</td>
</tr>
<tr>
<td>Amazonia</td>
<td>1,097</td>
<td>736,270</td>
<td>93.7</td>
<td>224</td>
<td>4,492</td>
<td>4.99</td>
<td>644</td>
</tr>
<tr>
<td>Southeast Asia</td>
<td>103</td>
<td>77,587</td>
<td>91.6</td>
<td>172</td>
<td>2,585</td>
<td>6.65</td>
<td>526</td>
</tr>
</tbody>
</table>

*H = Number of hyperdominants, TS = Total Species, H% = Percentage hyperdominance, α = Fisher’s α, Stems = Total number of stems, Plots = Total number of plots, %ID = Percentage of stems identified to the species level.*
Extended Data Table 2 | Rarefied minimum number of species required to account for 10%, 20%, ..., 90% of trees in the Africa, Amazonia, and Southeast Asia data, resampled to the common sample size of the Asia dataset (77,587 stems)

<table>
<thead>
<tr>
<th></th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>4</td>
<td>13</td>
<td>26</td>
<td>46</td>
<td>77</td>
<td>121</td>
<td>186</td>
<td>296</td>
<td>592</td>
</tr>
<tr>
<td></td>
<td>[2,7 ]</td>
<td>[8,17]</td>
<td>[19,33]</td>
<td>[36,57]</td>
<td>[62,92]</td>
<td>[102,139]</td>
<td>[163,209]</td>
<td>[264,328]</td>
<td>[494,689]</td>
</tr>
<tr>
<td>Amazonia</td>
<td>8</td>
<td>28</td>
<td>60</td>
<td>106</td>
<td>174</td>
<td>276</td>
<td>434</td>
<td>709</td>
<td>1413</td>
</tr>
<tr>
<td></td>
<td>[4,13]</td>
<td>[17,38]</td>
<td>[42,77]</td>
<td>[79,133]</td>
<td>[134,215]</td>
<td>[214,337]</td>
<td>[339,528]</td>
<td>[553,864]</td>
<td>[1029,1797]</td>
</tr>
<tr>
<td>Southeast Asia</td>
<td>9</td>
<td>26</td>
<td>54</td>
<td>98</td>
<td>172</td>
<td>285</td>
<td>468</td>
<td>790</td>
<td>1778</td>
</tr>
<tr>
<td></td>
<td>[5,13]</td>
<td>[17,35]</td>
<td>[36,72]</td>
<td>[66,130]</td>
<td>[125,219]</td>
<td>[220,350]</td>
<td>[380,556]</td>
<td>[670,910]</td>
<td>[1427,2129]</td>
</tr>
</tbody>
</table>

Percentage headings represent the different dominance thresholds. Confidence intervals are derived from the standard deviation across iterations of subsamples taken with replacement at the sample size of the Asia dataset.
Extended Data Table 3 | Rarefied proportion of total species required to account for 10%, 20%, ..., 90% of trees in the Africa, Amazonia, and Southeast Asia data, resampled to the common sample size of the Asia dataset (77,587 stems).

<table>
<thead>
<tr>
<th>Region</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>0.36</td>
<td>1.11</td>
<td>2.28</td>
<td>4.09</td>
<td>6.79</td>
<td>10.67</td>
<td>16.47</td>
<td>26.18</td>
<td>52.28</td>
</tr>
<tr>
<td></td>
<td>[0.12,0.61]</td>
<td>[0.71,1.51]</td>
<td>[1.64,2.93]</td>
<td>[3.1,5.07]</td>
<td>[5.39,8.20]</td>
<td>[8.89,12.44]</td>
<td>[14.28,18.65]</td>
<td>[23.22,29.15]</td>
<td>[43.00,61.56]</td>
</tr>
<tr>
<td>Amazonia</td>
<td>0.33</td>
<td>1.09</td>
<td>2.32</td>
<td>4.14</td>
<td>6.80</td>
<td>10.74</td>
<td>16.90</td>
<td>27.63</td>
<td>55.06</td>
</tr>
<tr>
<td></td>
<td>[0.15,0.51]</td>
<td>[0.67,1.50]</td>
<td>[1.63,3.01]</td>
<td>[3.09,5.19]</td>
<td>[5.24,8.36]</td>
<td>[8.40,13.09]</td>
<td>[13.30,20.50]</td>
<td>[21.68,33.58]</td>
<td>[40.37,69.76]</td>
</tr>
<tr>
<td>Southeast</td>
<td>0.35</td>
<td>1.01</td>
<td>2.09</td>
<td>3.79</td>
<td>6.65</td>
<td>11.03</td>
<td>18.10</td>
<td>30.56</td>
<td>68.78</td>
</tr>
<tr>
<td>Asia</td>
<td>[0.19,0.51]</td>
<td>[0.61,1.41]</td>
<td>[1.27,2.91]</td>
<td>[2.38,5.20]</td>
<td>[4.59,8.71]</td>
<td>[8.24,13.82]</td>
<td>[14.47,21.73]</td>
<td>[25.78,35.34]</td>
<td>[52.50,84.97]</td>
</tr>
</tbody>
</table>

Percentage headings represent the different dominance thresholds. Confidence intervals are derived from the standard deviation across iterations of subsamples taken with replacement at the sample size of the Asia dataset.
## Extended Data Table 4 | Extrapolated minimum number of species required to account for 10%, 20%, ..., 90% of trees in Africa, Amazonia, Southeast Asia, and the cross-regional total at the regional scale

<table>
<thead>
<tr>
<th>Percentage</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>7</td>
<td>22</td>
<td>41</td>
<td>69</td>
<td>104</td>
<td>154</td>
<td>228</td>
<td>354</td>
<td>713</td>
</tr>
<tr>
<td></td>
<td>[4,10]</td>
<td>[19,25]</td>
<td>[38,44]</td>
<td>[66,71]</td>
<td>[101,107]</td>
<td>[151,157]</td>
<td>[225,231]</td>
<td>[351,357]</td>
<td>[709,717]</td>
</tr>
<tr>
<td>Amazonia</td>
<td>20</td>
<td>62</td>
<td>120</td>
<td>196</td>
<td>299</td>
<td>443</td>
<td>651</td>
<td>1000</td>
<td>1892</td>
</tr>
<tr>
<td></td>
<td>[16,25]</td>
<td>[58,66]</td>
<td>[116,124]</td>
<td>[192,201]</td>
<td>[295,304]</td>
<td>[438,447]</td>
<td>[647,656]</td>
<td>[995,1005]</td>
<td>[1886,1899]</td>
</tr>
<tr>
<td>Southeast Asia</td>
<td>12</td>
<td>51</td>
<td>106</td>
<td>179</td>
<td>278</td>
<td>417</td>
<td>625</td>
<td>988</td>
<td>2243</td>
</tr>
<tr>
<td></td>
<td>[1,23]</td>
<td>[40,62]</td>
<td>[95,116]</td>
<td>[168,189]</td>
<td>[268,289]</td>
<td>[406,427]</td>
<td>[614,636]</td>
<td>[977,1000]</td>
<td>[2225,2262]</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>135</td>
<td>267</td>
<td>444</td>
<td>681</td>
<td>1014</td>
<td>1504</td>
<td>2342</td>
<td>4848</td>
</tr>
<tr>
<td></td>
<td>[21,58]</td>
<td>[117,153]</td>
<td>[249,284]</td>
<td>[426,461]</td>
<td>[664,700]</td>
<td>[995,1031]</td>
<td>[1486,1523]</td>
<td>[2323,2362]</td>
<td>[4820,4878]</td>
</tr>
</tbody>
</table>

Percentage headings represent the different dominance thresholds. Prediction intervals combine uncertainty from the standard error of predicted means and the residual standard deviation of the regression of the bias correction fit. ‘Total’ minimum number of species required to account for 10%-90% of trees across all of the regions are calculated as the sum of the number of hyperdominants across the three major tropical forest regions.
<table>
<thead>
<tr>
<th></th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Africa</strong></td>
<td>0.16</td>
<td>0.48</td>
<td>0.90</td>
<td>1.48</td>
<td>2.23</td>
<td>3.31</td>
<td>4.90</td>
<td>7.60</td>
<td>15.29</td>
</tr>
<tr>
<td><strong>Amazonia</strong></td>
<td>0.14</td>
<td>0.44</td>
<td>0.86</td>
<td>1.41</td>
<td>2.16</td>
<td>3.21</td>
<td>4.72</td>
<td>7.25</td>
<td>13.73</td>
</tr>
<tr>
<td><strong>Southeast Asia</strong></td>
<td>0.14</td>
<td>0.46</td>
<td>0.91</td>
<td>1.5</td>
<td>2.32</td>
<td>3.45</td>
<td>5.14</td>
<td>8.10</td>
<td>18.34</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>0.13</td>
<td>0.44</td>
<td>0.88</td>
<td>1.46</td>
<td>2.24</td>
<td>3.33</td>
<td>4.94</td>
<td>7.70</td>
<td>15.93</td>
</tr>
</tbody>
</table>

Percentage headings represent the different dominance thresholds. ‘Total’ minimum proportion of total species required to account for 10%-90% of trees across all of the regions are calculated as the proportion between the sum of the number of hyperdominants and the sum of total species across the three major tropical forest regions.
Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection: No software was used for data collection.

Data analysis: The custom R code (version 4.3.1) used to run the analyses and produce the figures is available on GitHub with the identifier https://github.com/declancouper/CommonSpecies2022.git

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

Data

Policy information about availability of data

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

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The species abundance data that support the findings of this study are available from https://doi.org/10.6084/m9.figshare.21670883. WorldClim bioclimatic data are available from https://www.worldclim.org/data/bioclim.html.
Human research participants

Policy information about: studies involving human research participants and Sex and Gender in Research.

| Reporting on sex and gender | NA |
| Population characteristics | NA |
| Recruitment | NA |
| Ethics oversight | NA |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

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

- [x] Life sciences
- [ ] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

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

Life sciences study design

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

| Sample size | No sample size calculation was performed. We selected all available plots meeting the following criteria: forest inventory plots $\geq 0.2$ ha in size, situated in structurally intact (no logging or fire), closed canopy (not dry forest or savanna) tropical forest, with enumeration of all stems $\geq 10$ cm diameter, in which $\geq 80\%$ of stems are identified to the species level. These criteria allow direct comparisons to be made with results from previous studies investigating common species abundances in Amazonia. The major tropical forest regions of Amazonia, Africa, and Southeast Asia are adequately represented. This is the largest dataset of repeatedly measured plots ever used to calculate long-term trends in African forest carbon dynamics. |
| Data exclusions | Plots not meeting the following criteria were excluded from analysis: forest inventory plots $\geq 0.2$ ha in size, situated in structurally intact (no logging or fire), closed canopy (not dry forest or savanna) tropical forest, with enumeration of all stems $\geq 10$ cm diameter, in which $\geq 80\%$ of stems are identified to the species level. These conditions were pre-established in line with previous studies. |
| Replication | Replication of repeated random subsampling 200 times ensured that derived results were stable and reproducible. |
| Randomization | Forest inventory data were partitioned into the the Amazonia, Africa, and Southeast Asia study regions by location. In all of the analyses, sampling and sub-sampling by plots and by trees within plots was done completely randomly within regions. |
| Blinding | Blinding is not relevant to this study. |

Reporting for specific materials, systems and methods

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

Materials & experimental systems

- [x] Antibodies
- [x] Eukaryotic cell lines
- [x] Palaeontology and archaeology
- [ ] Clinical data
- [ ] Dual use research of concern

Methods

- [ ] Involved in the study
- [x] ChIP-seq
- [ ] Flow cytometry
- [x] MRI-based neuroimaging