Hyperdominance in the Amazonian Tree Flora
Hans ter Steege et al.
Science 342, (2013);
DOI: 10.1126/science.1243092

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of October 17, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/342/6156/1243092.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2013/10/16/342.6156.1243092.DC1.html

This article cites 50 articles, 10 of which can be accessed free:
http://www.sciencemag.org/content/342/6156/1243092.full.html#ref-list-1
Hyperdominance in the Amazonian Tree Flora

Hans ter Steege* et al.

Introduction: Recent decades have seen a major international effort to inventory tree communities in the Amazon Basin and Guiana Shield (Amazonia), but the vast extent and record diversity of these forests have hampered an understanding of basinwide patterns. To overcome this obstacle, we compiled and standardized species-level data on more than half a million trees in 1170 plots sampling all major lowland forest types to explore patterns of commonness, rarity, and richness.

Methods: The ~6-million-km² Amazonian lowlands were divided into 1° cells, and mean tree density was estimated for each cell by using a loess regression model that included no environmental data but had its basis exclusively in the geographic location of tree plots. A similar model, allied with a bootstrapping exercise to quantify sampling error, was used to generate estimated Amazon-wide abundances of the 4962 valid species in the data set. We estimated the total number of tree species in the Amazon by fitting the mean rank-abundance data to Fisher’s log-series distribution.

Results: Our analyses suggest that lowland Amazonia harbors $3.9 \times 10^{11}$ trees and ~16,000 tree species. We found 227 “hyperdominant” species (1.4% of the total) to be so common that together they account for half of all trees in Amazonia, whereas the rarest 11,000 species account for just 0.12% of trees. Most hyperdominants are habitat specialists that have large geographic ranges but are only dominant in one or two regions of the basin, and a median of 41% of trees in individual plots belong to hyperdominants. A disproportionate number of hyperdominants are palms, Myristicaceae, and Lecythidaceae.

Discussion: The finding that Amazonia is dominated by just 227 tree species implies that most biogeochemical cycling in the world’s largest tropical forest is performed by a tiny sliver of its diversity. The causes underlying hyperdominance in these species remain unknown. Both competitive superiority and widespread pre-1492 cultivation by humans are compelling hypotheses that deserve testing. Although the data suggest that spatial models can effectively forecast tree community composition and structure of unstudied sites in Amazonia, incorporating environmental data may yield substantial improvements. An appreciation of how thoroughly common species dominate the basin has the potential to simplify research in Amazonian biogeochemistry, ecology, and vegetation mapping. Such advances are urgently needed in light of the >10,000 rare, poorly known, and potentially threatened tree species in the Amazon.
Hyperdominance in the Amazonian Tree Flora


The vast extent of the Amazon Basin has historically restricted the study of its tree communities to the local and regional scales. Here, we provide empirical data on the commonness, rarity, and richness of lowland tree species across the entire Amazon Basin and Guiana Shield (Amazonia), collected in 1170 tree plots in all major forest types. Extrapolations suggest that Amazonia harbors roughly 16,000 tree species, of which just 227 (1.4%) account for half of all trees. Most of these species are habitat specialists and only dominant in one or two regions of the basin. We discuss some implications of the finding that a small group of species—less diverse than the North American tree flora—accounts for half of the world’s most diverse tree community.

Much remains unknown about the Amazonian flora, the richest assemblage of plant species on Earth. Tree inventories carried out over the past two decades have helped improve our understanding of regional-scale patterns of distribution and abundance in Amazonian tree communities, but similar advances at the basin-wide scale remain scarce. Scientists still do not know how many tree species occur in the Amazon (1), how many tree species have been recorded to date in the Amazon, how those species are distributed across the basin, and in what regions or forest types they are rare or common. So uncertain are patterns at the largest scales that even the simplest question of all—what is the most common tree species in the Amazon?—has never been answered.

In practical terms, this lack of basic information means that the largest pool of tropical carbon on Earth remains a black box for ecologists. Also, given that abundance and frequency are the primary currencies of conservation status, there is essentially no information on which Amazonian tree species face the most severe threats of extinction and where to protect them. Although ecologists have an ever-larger toolbox of methods to extrapolate from local surveys to larger-scale patterns, well-known barriers to sampling, identifying, and putting valid names on tropical trees have long been assumed to make Amazon-wide extrapolations impossible.

We challenged those assumptions with a wide-ranging assessment of the composition and biogeography of Amazonian tree communities (2). We compiled stem density and species abundance data from 1170 tree inventory plots across the Amazon (Fig. 1), well distributed among all regions and major forest types (table S1 and figs. S1 to S3), to generate basin-wide estimates of the abundance, frequency, and spatial distribution of thousands of Amazonian tree species.

Results

A Rank-Abundance Distribution for Amazonian Trees

The plots contained a total of 4962 valid species, 810 genera, and 131 families of trees [free-standing stems ≥ 10 cm in diameter at breast height (dbh)]. By using stem density and species abundance data collected in the individual plots, we constructed a spatial model that yielded estimated basin-wide population sizes for every valid species in the data set. The rank-abundance distribution (RAD) of these data (Fig. 2) offers four important insights regarding Amazonian tree communities.

First, it provides the most precise estimates yet of two numbers that have been debated for decades: How many trees and how many tree species occur in the ~6-million-km² landscape of Amazonia (1, 3–5). Our estimate of tree density yielded a total of 3.9 × 10¹¹ individual trees and a median tree density of 565 trees/ha (fig. S4). Assuming that our population size estimates for the common species are reasonable (fig. S5) and that Fisher’s log-series model fits our data (table S2 and figs. S6 and S7) (1), we estimated the total number of tree species in the Amazon to be about 16,000 (Fig. 2). A second estimate based on the Fisher’s alpha scores of all plots yielded a similar figure: 15,182 species (fig. S8).

Second, the RAD suggests that just 227 (1.4%) of the estimated 16,000 species account for half of all individual trees in Amazonia. We refer to these species, all of which have estimated populations of >3.7 × 10¹² trees, as hyperdominant species (see a list of the 20 most abundant species in Table 1 and a full list in appendix S1). These hyperdominant species form the basis of the tree communities in individual plots as well, accounting for a median of 41% of trees (range = 0 to 94%, fig. S9) and 32% of species (range = 0 to 78%) per plot (fig. S9).

Third, all species ranking in abundance from 5000 to 16,000 are very rare. These species in the tail of the RAD have total populations of <10⁶ individuals and together account for just 0.12% of all trees in Amazonia. Although some of these species may be treelets or climbers rarely reaching tree stature or “vagrants” spilling over from extra-Amazonian biomes such as the Cerrado and Andes, thousands must be Amazonian endemics that run a high risk of going extinct, even before they can be found and described by biologists. The rarest 5800 species have estimated population sizes of <1000, which is sufficient to classify those that are endemic as globally threatened (6).

Together, these taxa (the rarest 36% of species)
account for just 0.0003% of all trees in Amazonia. Given the extreme likelihood of locating a fertile individual of one of these species, we believe that discovering and describing the unknown portion of Amazonian biodiversity will be a long-term struggle with steeply diminishing returns and not an easy linear process (7). Indeed, the RAD suggests that floras of even well-collected areas may remain half-finished for decades. For example, our model predicts that ~4500 tree species occur in the Guianas (fig. S10), but centuries of collecting there have yielded just half that number (8). Some of these species may be present among the unidentified species of our plots or as undescribed specimens in herbaria (9), but the majority may yet have to be collected.

Fourth, there are strong similarities between theoretical models of tree species richness in the Amazon (7) and our distribution of species abundances based on empirical data. For example, Hubbell et al. (7) used a log-series distribution to predict that the most common species in the Amazon should account for 1.39% of all trees. This is very close to our estimate for the most common species in our data set, the palm *Euterpe precatoria* (1.32%). Our empirical estimate of Fisher’s alpha for the Amazon (fig. S8) is also extremely close to Hubbell et al.’s modeled prediction [754 versus 743 in (7)]. Although these strong correlations between predictions and our data set suggest that the log series may offer useful insights on the most poorly known tree species in the Amazon (e.g., the number of undescribed taxa), they should not necessarily be interpreted as evidence for any one theory of how these tree communities are structured (10, 11).

Hyperdominant Patterns in Regions, Forest Types, and Taxonomic Groups

We examined species’ geographic ranges and abundances by plots, regions, and forest types to explore how hyperdominant species differ from other taxa, as a first step toward understanding what makes them so successful. Hyperdominant species have larger ranges than other taxa (Fig. 3A) and reach greater maximum relative abundances in plots (Fig. 3B). Most hyperdominant species (121 out of 227) are habitat specialists (Fig. 3C) [i.e., they show a strong preference for one of the five major Amazonian forest types: terra firme (53 spp.), várzea (26), white-sand forest (16), swamps (14), and igapó (12)]. Most are only dominant within one or two forest types. When the study area was divided into six regions (Guiana Shield and northwest, southwest, south, east, and central Amazonia), most hyperdominant species (73%) were found to be dominant within only one or two regions (Table 2).

It is thus important to emphasize that, although the Amazonian RAD is dominated by a small suite of species, most of those species are only dominant in certain forest types and in certain regions of the basin. Just one species qualified as dominant in all six regions (Espchevleria coriacea), no species were dominant in all five forest types, and only four species were dominant in four forest types (*E. precatoria*, *Oenocarpus batua*, *Licania apetala*, and *Euterpe oleracea*). Much more representative of the 227 hyperdominant species are taxa like *Siparuana decipiens* (112th largest population size overall), only dominant in terra firme forests in southwest Amazonia, and *Eperua falcata* (13th), only dominant in the Guiana Shield. Indeed, 58% of hyperdominant species qualify as both dominant in one or two regions and dominant in one or two forest types.

Within each region, an even smaller number of species (75 to 163) typically accounts for
50% of all individual trees, and most of these regional dominants are also hyperdominant species (Fig. 4A). For example, the data suggest that half of all individual trees in southwest Amazonia belong to just 76 species, 50 of which are also hyperdominant species. The same pattern holds for forest types, which are individually dominated by 25 to 195 species (Fig. 4B). Half of all individual trees in white-sand forests belong to just 25 species, 15 of which are also hyperdominant species. Because most hyperdominant species are only dominant in one or two regions or forest types, in any single region or forest type the majority of the 227 hyperdominant species are not locally dominant.

Given these results, it seems likely that the basinwide patterns of dominance we describe here arise in part from regional-scale patterns of dominance described previously at various sites in upper Amazonia (12, 13). There is substantial compositional overlap between Pitman et al.’s (12) “oligarchies” in Peru and Ecuador and our hyperdominant species, even though those authors’ plots represent just 2.1% of the full Amazon Tree Diversity Network (ATDN) data set and only include terra firme forests. Sixty-eight “oligarchs” of (12) are on the list of 227 hyperdominant species, including 8 of the top 10 most common hyperdominants. The 250 oligarchic species in (12) account for 26.9% of all trees in Amazonia, according to the RAD in Fig. 2. These results suggest that the regional-scale and Amazon-wide patterns derive from similar processes.

Hyperdominants are more frequent in some families (table S3). Arecaceae, Myristicaceae, and Lecythidaceae have many (~four to five times) more hyperdominant species than expected by chance, whereas Myrtaceae, Melastomataceae, Lauraceae, Annonaceae, and Rubiaceae have fewer, probably because many of their species are shrubs or treelets that do not reach our 10-cm-diameter cutoff. In Fabaceae, the most abundant and most diverse family in the data set, the observed number of hyperdominant species is not significantly different from the expected.

We observed a negative relationship between the number of species in a genus and the frequency of hyperdominant species (fig. S11). This pattern has been observed in several plant communities worldwide, and scientists have yet to determine whether it is ecologically informative or an artifact of rank-based taxonomy (14, 15). The 227 hyperdominant species belong to 121 genera, and 68 of these contain more hyperdominants than expected by chance (appendix S3). The highest number of hyperdominant species is found in moderately diverse Eschweilera (52 species overall; 2.4 hyperdominant species expected versus 14 observed), also the most abundant genus in the ATDN data set (5.2% of all stems). Given that the families and genera mentioned here dominate Amazonian forests, it remains a key goal to determine why some achieve dominance with a large number of mostly rare species (e.g., Inga, Sapotaceae) whereas others do so with a small number of common species (palms), differences that may result from variation in speciation and extinction rates (14–17). Although genetics data may reveal some hyperdominant species to be species complexes, there is not yet enough knowledge on how widespread such complexes are, where they are located along our RAD,
and to what degree they could alter the patterns described here ([18] and references therein).

Discussion

Exploring Potential Causes for Hyperdominance

We found no evidence that two key functional traits for trees, seed mass and wood density, vary consistently with hyperdominance. The 227 hyperdominant species include both shade-tolerant, typically large-seeded climax species with dense wood (e.g., Chlorocardium rodiei, Clathrotropis spp., and Eperua spp.) and shade-intolerant, small-seeded pioneers with light wood (e.g., Cecropia spp., Jacaranda copaia, and Laetia procera). Given that most hyperdominant species attain very high local densities (>60 trees/ha) somewhere in the plot network, we predict that they will be found to be disproportionately resistant to pathogens, specialist herbivores, and other sources of frequency-dependent mortality (19, 20).

Table 1. Population characteristics of the 20 most abundant tree species of the Amazon.

<table>
<thead>
<tr>
<th>Species</th>
<th>Mean estimated population in the Amazon</th>
<th>SD estimated population (%)</th>
<th>No. trees in data set</th>
<th>% of all plots where present</th>
<th>Maximum abundance recorded (trees/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euterpe precatoria</td>
<td>5.21 × 10^9</td>
<td>9.9</td>
<td>5903</td>
<td>32.7</td>
<td>168</td>
</tr>
<tr>
<td>Protium altissimum</td>
<td>5.21 × 10^9</td>
<td>18.0</td>
<td>5889</td>
<td>15.6</td>
<td>128</td>
</tr>
<tr>
<td>Eschweilera coriacea</td>
<td>5.00 × 10^9</td>
<td>5.6</td>
<td>9047</td>
<td>47.9</td>
<td>28</td>
</tr>
<tr>
<td>Pseudolmedia laevis</td>
<td>4.30 × 10^9</td>
<td>8.9</td>
<td>5285</td>
<td>36.1</td>
<td>121</td>
</tr>
<tr>
<td>Iriartea detoidea</td>
<td>4.07 × 10^9</td>
<td>13.1</td>
<td>8405</td>
<td>18.5</td>
<td>169</td>
</tr>
<tr>
<td>Euterpe oleracea</td>
<td>3.78 × 10^9</td>
<td>17.5</td>
<td>8572</td>
<td>7.4</td>
<td>397</td>
</tr>
<tr>
<td>Oenocarpus bataua</td>
<td>3.71 × 10^9</td>
<td>10.7</td>
<td>4767</td>
<td>29.9</td>
<td>108</td>
</tr>
<tr>
<td>Trattinnickia burserifolia</td>
<td>2.78 × 10^9</td>
<td>29.4</td>
<td>3023</td>
<td>10</td>
<td>125</td>
</tr>
<tr>
<td>Socratea exorrhiza</td>
<td>2.68 × 10^9</td>
<td>10.8</td>
<td>863</td>
<td>28.6</td>
<td>82</td>
</tr>
<tr>
<td>Astracarum murumuru</td>
<td>2.41 × 10^9</td>
<td>11.2</td>
<td>5748</td>
<td>16.7</td>
<td>325</td>
</tr>
<tr>
<td>Brosimum lactescens</td>
<td>2.28 × 10^9</td>
<td>10.0</td>
<td>2234</td>
<td>28.2</td>
<td>106</td>
</tr>
<tr>
<td>Protium heptaphyllum</td>
<td>2.13 × 10^9</td>
<td>32.2</td>
<td>1365</td>
<td>11.3</td>
<td>169</td>
</tr>
<tr>
<td>Eperua falcata</td>
<td>1.95 × 10^9</td>
<td>15.8</td>
<td>1898</td>
<td>10.9</td>
<td>266</td>
</tr>
<tr>
<td>Hevea brasiliensis</td>
<td>1.91 × 10^9</td>
<td>15.5</td>
<td>6031</td>
<td>14.8</td>
<td>179</td>
</tr>
<tr>
<td>Eperua leucantha</td>
<td>1.84 × 10^9</td>
<td>32.3</td>
<td>1453</td>
<td>1.4</td>
<td>282</td>
</tr>
<tr>
<td>Helicostylis tomentosa</td>
<td>1.79 × 10^9</td>
<td>25.6</td>
<td>1948</td>
<td>36.5</td>
<td>89</td>
</tr>
<tr>
<td>Attalea butyacea</td>
<td>1.78 × 10^9</td>
<td>16.2</td>
<td>2561</td>
<td>5.8</td>
<td>73</td>
</tr>
<tr>
<td>Rinorea guianensis</td>
<td>1.69 × 10^9</td>
<td>18.6</td>
<td>1243</td>
<td>13.7</td>
<td>182</td>
</tr>
<tr>
<td>Licania heteromorpha</td>
<td>1.57 × 10^9</td>
<td>14.4</td>
<td>2483</td>
<td>35</td>
<td>173</td>
</tr>
<tr>
<td>Metrodorea flavida</td>
<td>1.55 × 10^9</td>
<td>14.7</td>
<td>1326</td>
<td>7.7</td>
<td>128</td>
</tr>
<tr>
<td>Median of other hyperdominant species</td>
<td>5.79 × 10^8</td>
<td>808</td>
<td>11.4</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Median of non-hyperdominant species</td>
<td>1.11 × 10^7</td>
<td>15</td>
<td>0.5</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
Widespread pre-1492 cultivation by humans is a compelling hypothesis to explain hyperdominance (21). Numerous hyperdominant species are widely used by modern indigenous groups (Hevea brasiliensis, Theobroma cacao, and many palms), and some are associated with pre-Columbian settlements (Attalea butyracea, A. phalerata, Mauritia flexuosa) (22–26). On the other hand, most hyperdominant species are not commonly cultivated; many of the most commonly used hyperdominants (palms) belong to a family that appears to have been dominant in tropical South America since the Paleocene (27), and large portions of the Amazon Basin do not appear to have been heavily cultivated before 1492 (28).

Testing the Validity of the Model Predictions

A fundamental assumption of our analyses is that the population-size estimates generated by the loess model were reasonably accurate for the most abundant species. This assumption is disputable for a few reasons: (i) The data set is very small compared with the community to which it was extrapolated; (ii) tree plots were not distributed randomly across the study area; (iii) trees were identified by many different research teams; and (iv) no environmental data were used by the model, even though many species in the ATDN data set are known to respond to environmental heterogeneity in the study area. A fifth problem makes the assumption especially difficult to test: (v) the fact that a basinwide population size has not been empirically determined for any Amazonian tree species, which precludes a comparison between projected and observed values. Here, we address these shortcomings by attempting to quantify the error that each could introduce into our results.

![Graph A](image)
![Graph B](image)
![Graph C](image)

Fig. 3. Characteristics of hyperdominant tree species of the Amazon. (A) Hyperdominant species (red) have larger geographic ranges than other species (gray), (B) reach higher maximum relative abundances in individual plots, and (C) are more likely to be habitat specialists.

Table 2. Hyperdominance by region and forest type. The number of hyperdominant species that are also dominant in individual forest types and regions. Most hyperdominants only dominate a single forest type, and most are dominant in one or two regions.

<table>
<thead>
<tr>
<th>No. forest types where dominant</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>47</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>65</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>17</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>9</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>total</td>
<td>35</td>
<td>148</td>
<td>28</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>227</td>
</tr>
</tbody>
</table>

To test how sampling intensity and the geographic distribution of plots (problems i and ii) affected the estimated population sizes of hyperdominant species, we recorded the frequency with which the 227 hyperdominants qualified as hyperdominant in the 500 runs of the bootstrap exercise described in the methods section. Most species (137, 60% of the total) qualified as hyperdominants in 90 to 100% of runs, whereas 207 species (91.2%) qualified as hyperdominants in more than half of runs (fig. S12A). Median (fig. S12B) and mean (fig. S12C) ranks for the 500 runs showed high stability.

In bootstrap runs for which a given hyperdominant species did not qualify among the top 227 species, it rarely qualified as rare. The lowest median rank observed for a hyperdominant species in the 500 bootstrap runs was 275, and hyperdominant species never ranked lower than 1000th (i.e., ranks 1000 to 4790). These analyses provide strong evidence that the identities and estimated population sizes of the hyperdominant species remain stable and predictable with varying levels of sampling intensity and geographic bias.

Taxonomic and identification problems (problem iii) are widespread in Amazonian tree inventories. However, two independent lines of evidence suggest that resolving these problems will not fundamentally alter the patterns described for hyperdominant species.

First, we observed a consistent relationship in the ATDN data set between the abundance of a species and the likelihood that it had been identified with a valid name. The percentage of identified species in individual plots was significantly higher than that of unidentified species-level taxa (87 versus 13% stems/ha, analysis of variance, F_s = 22.774, P << 0.001). Furthermore, very common morphospecies are very infrequent in the ATDN data set. Only 48 of the 1170 ATDN plots contained a morphospecies that accounted for >10% of all individuals, and only 10 plots contained a morphospecies that reached >20%. Given that all 227 hyperdominants reach high local relative abundances (Fig. 3B), these numbers suggest that very few currently unidentified species will eventually qualify as hyperdominant species.

Second, we see strong evidence that taxonomic and identification problems are less severe in hyperdominant species than in other species, in the form of a strong positive correlation between the abundance of a species in the field, the number of specimens in herbaria, and the number of fertile specimens (i.e., specimens with flowers or fruits) collected during field work. Common species are better represented in herbaria than rare species, because individual collectors are more likely to encounter them (29). Common species are also more likely than rare species to be collected fertile during the establishment of tree plots. For example, in 25 ATDN plots established in eastern Ecuador (30), we found that hyperdominant species were more likely than other species to be collected fertile (27.8 versus 17.7%). Botanists trying to identify a hyperdominant species thus...
have both a higher likelihood of matching their field specimens with museum specimens and a broader range of morphological features to facilitate identification.

The model we used to estimate population sizes was a loess function, parameterized exclusively with plot location and observed species abundances in plots and no environmental data (problem iv). This is a very different approach from the most commonly used class of species distribution modeling: maximum entropy modeling or Maxent (31, 32). Maxent uses presence-only data fitted to environmental variables of confirmed locations to produce a map of habitat suitability. In a Maxent model, a species known to occur under a given set of environmental conditions is predicted to occur in all environmentally similar areas, even when those areas are outside of the species’ known range. Because Amazonian tree species are known to respond strongly to environmental variation, an earlier version of our model included climatic data. That version, however, routinely predicted significant populations of species in regions of the Amazon where a large number of ATDN plots and other plant collection efforts had consistently failed to record those species (i.e., type I errors were common). Modeling with only latitude and longitude as predictive variables is a more conservative option, because it ensures that such errors will be made at a much lower frequency and that species will never be predicted far from confirmed records (Fig. 5). For the same reason, we used a span of 0.2; at higher span values, species ranges extended too far into areas with no known occurrence. Varying span values from 0.2 to 0.5 did not strongly affect population size estimates.

It is not possible to compare estimated population sizes with measured population sizes (problem v), because the latter do not exist for any Amazonian tree species. However, it is possible to compare the population sizes estimated by the loess model with population sizes estimated by using a different method based on the measured extent of Amazonian forest types. The estimated population of *Maurita flexuosa* is 1.5 billion stems. If we assume that one hectare of monodominant *M. flexuosa* swamp contains 565 *M. flexuosa* trees, then our 1.5-billion-stem estimate suggests that there are <3 million ha of monodominant *M. flexuosa* swamps in the entire basin. This appears reasonable, because the largest block of largely monodominant *M. flexuosa* stands in the basin (the Pastaza Fan) measures ~2.2 million ha. A similar test for white sands and podzol using *E. falcata* and *E. leucantha* (lumped together) was carried out. Together the model estimates that 3.9 billion trees in the greater Amazon belong to these species. If we assume that one hectare of white-sand [podzols and albic arenosols (33)] forest contains on average 150 stems that belong to these species, then the model suggests that there are roughly 26 million ha of white-sand and podzol forest in the greater Amazon. The extent of podzols in the greater Amazon has been estimated as 17 million ha (34). The estimate of podzols and arenosols (fig. S2) is 34 million ha (33).

We know of one study that attempted to estimate populations of trees over a large area in the Amazon Basin based on forest inventories of trees >30 cm dbh (35). The most abundant species in central western Amazonia (blocks: Roraima–Boa Vista, Manaus, and Rio Purus; total forest area 623,139 km$^2$) was *E. coriacea*, with an estimated population of 193 million individuals (this compares to roughly 800 million trees with >10 cm dbh, followed by *Goupia glabra* (93 million individuals, or 370 million trees with >10 cm dbh). Rollet concluded that *E. coriacea* is likely the most common tree species in the Brazilian Amazon. Although our data suggest that two other species have higher total population sizes (*E. precatoria* and *Protium altissimum*), a difference caused by our much larger study area (~10$^5$) and lower diameter cutoff (four times as many trees ha$^{-1}$), our estimate of *E. coriacea* (~5000 million) is of a similar order of magnitude (193 million × 10 × 4 = 7000 million). It is also worth noting that, in the forest inventories used by Rollet, other *Eschweilera* species were pooled more often with *E. coriacea* than in our inventories [see (36, 37) for a discussion on this].

**Practical Implications**

The finding that Amazonia is dominated by just 227 tree species has important practical implications. It suggests that roughly half of all fruits, flowers, pollen, leaves, and biomass in the world’s most diverse forest belong to a very small suite of species, which must therefore account for a large proportion of Amazonian ecosystem services, including water, carbon and nutrient cycling. Our data also suggest that it may be possible to forecast a substantial proportion of the tree community composition and structure of unstudied sites in Amazonia with a purely spatial model. Although no one should underestimate the importance of the >10,000 rare and poorly known tree species in the Amazon (38), an appreciation of how thoroughly common species dominate the basin has the potential to greatly simplify
research in Amazonian biogeochemistry, plant and animal ecology, and vegetation mapping.

Materials and Methods
The ATDN network (39) comprises 1430 tree inventory plots distributed across the Amazon Basin and Guiana Shield, hereafter Amazonia (Fig. 1). Plots were established between 1934 and 2011 by hundreds of different botanists, some working in basinwide or global networks (39–42). Analyses of tree density were performed by using the 1346 plots with trees with ≥10 cm dbh that remained after plots with outlying density values (<100 or >1000 individuals/ha), poorly defined areas, or a different diameter cutoff level were removed.

Analyses of composition were performed with a subset of 1170 plots in which all 639,639 free-standing trees with ≥10 cm dbh had been identified with a valid name at the species (86.6%), genus (96.9%), or family (98.9%) level before our study. Most plots (852) measured 1 ha; 253 were smaller, 61 were larger, and 4 were plotless samples (point-centered quarter) for which the sampled area was unknown but the number of trees was equivalent to that typically found in 0.5 to 1 ha. We did not compare specimens or reidentify trees from these plots but resolved major nomenclatural issues (i.e., synonyms and misspellings) in the existing data sets by cross-checking all names with the TROPICOS database (43), via the Taxonomic Name Resolution Service [TNRS (44)(version October 2011)]. We made two adjustments to the names given in TROPICOS (supplementary text). Rollinia was merged with Annona, because phylogenetic analysis has revealed it to be nested inside that genus (45). Similarly, Crepidospermum and Tetragastris are nested in Protium (46) and were merged into that genus. For the small proportion of names whose validity could not be determined with those tools, we used The Plant List (47). Lianas, bamboos, tree ferns, and tree-sized herbs were excluded from all analyses. Varieties and subspecies were ignored (i.e., all individuals were assigned to the species level). Although some individuals may be misidentified, we assume that this error is within acceptable limits, especially for common species (see discussion above).

The total number of trees ≥10 cm dbh in Amazonia was estimated as follows. First, the study area was divided into 567 1°-grid cells (DGCs; Fig. 1). We constructed a loess regression model for tree density (stems ha⁻¹) on the basis of observed tree density in 1195 plots, with latitude, longitude, and their interaction as independent variables. The span was set at 0.5 to yield a relatively

---

**Fig. 5. Distribution maps of three hyperdominant Amazon tree species.**
Distribution maps estimated by the spatial loess model for three Amazonian hyperdominant species: (A) *E. falcata*, ranked 13th in abundance overall and with an eastern distribution; (B) *Iriartea deltoidea*, ranked fifth overall and with a western distribution; and (C) *E. coriacea*, ranked third overall and with a pan-Amazonian distribution. Black dots are tree plots where the species has been recorded, and dot size indicates the relative abundance of the species in the plot. Red dots are plots where the species has not been recorded. Shading in DGCs indicates the loess spatial average. For *E. falcata*, the relative abundance in individual plots ranged from 0 to 73.28%, and the loess spatial average in individual grid cells ranged from 0 to 11.27%. Comparable numbers for *I. deltoidea* are 0 to 38.47% and 0 to 12.17% and, for *E. coriacea*, 0 to 21.52% and 0 to 15.01%.
stems had). The total number of trees in each DGC ($D_{DGC}$) was then calculated by multiplying $D_{DGC}$ by 1,232,100 ha (the area of a DGC close to the equator—the deviation from this area is just 2.8% at 14°S and 1.1% at 8°N, our latitudinal range). Both empirical (plot data) and interpolated tree densities are illustrated in fig. S4.

The total number of trees belonging to each species in Amazonia was estimated as follows. Abundances of all valid species were converted to relative abundances for each plot: $R_{i} = n_i/N$, where $n_i$ = the number of individuals of species $i$ and $N$ = the total number of trees in the plot (including unidentified trees).

For each of the 4962 species with a valid name, we constructed a loess model for $R_A$, with latitude, longitude, and their interaction as independent variables and a span of 0.2. We used only spatially independent variables, because test runs including environmental variables commonly led to predictions of species occurrences in under-sampled areas where they had never been recorded in plots. For a similar reason (i.e., to keep predictions spatially conservative), a smaller span was used in the tree density analysis. Negative predicted abundances were set to 0. The loess model of a species predicted relative abundance in each DGC, yielding a map of its predicted variation in relative abundances across Amazonia. The total population size of each species was calculated by multiplying its relative abundance in each DGC by the total number of trees in that DGC and then summing these products for all DGCS.

To reduce the impact of individual plots and quantify uncertainty in the above procedure, we carried out a bootstrap exercise. This involved randomly drawing 1000 plots from the 1170-plot data set (with replacement) and calculating the population sizes of all species as described above. This was repeated 500 times, and the 500 population estimates per species were used to calculate mean estimated population size and 95% confidence intervals (i.e., mean ± 1.96 SD).

To estimate range size, we used the same data and methods as (48), standardized with TNRS and updated with specimen records from SpeciesLink (49). Species not found in this database were left out of the range size analysis ($n = 842$). Worldwide species diversity of genera was estimated by counting accepted species in (47). Seed mass and wood density data were obtained from sources described in (50).

Habitat preference was analyzed by means of Indicator Species Analysis, a permutation test that calculates indicator values for each species based on their frequency and relative abundance (50) in the five forest types (igapó, terra firme, swamp, várzea, and white-sand forest). To analyze regional-level dominance, we divided Amazonia into six regions and created a RAD for each region by summing population sizes in the DGCS they contained. RADS were also constructed for each forest type by summing the individuals of each species in all plots of a given forest type and calculating the average density of each species in that forest type. The forest-type RADS thus have their basis not in population estimates in DGCS but in the raw abundance data in our plots. A species was considered dominant in a given region or forest type if it appeared in the list of species comprising the upper 50% percentile of the respective RAD.

All analyses were carried out with the R software platform (57). For Indicator Species Analysis, we used the package labdsv. All other permutation tests were custom written.
Acknowledgments: This paper is the result of the work of hundreds of different scientists and research institutions in the Amazon over the past 80 years. Without their hard work this analysis would have been impossible. This work was supported by Alberta Mennega Stichting; ALCGA Suriname; Banco de la República; Center for Agricultural Research in Suriname; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Plano Nacional de Pós-Graduação); Conselho Nacional de Desenvolvimento Científico e Tecnológico of Brazil (CNPq) projects Programa de Pesquisas Ecológicas de Longa Duração (PELD) (S58069/2009-6), Programa de Apoio a Núcleos de Excelência da Fundação de Amparo à Pesquisa do Estado do Amazonas (PRONEX-FAPEAM) (1600/2006), Áreas Úmidas, and MAUA; PELD (403792/2012-6), PPBio, CENBAM, Universal (479599/2008-4), and Universal 307087-2009-6; Fundação de Amparo à Pesquisa do Estado do Amazonas (APEAM) projects: DCR/2006, Hidrovég with FAESP, and PRONEX with CNPq; FAPESP; Colciencias; Duke University; Ecopetroil; FEPAM 044/2003; the Field Museum; Conservation International/DC (TEAM/INPA Manuas), Gordon and Betty Moore Foundation; and U.S. National Geographic Society. O.L.P. is supported by a European Research Council Advanced Grant and a Royal Society Wolfson Research Merit Award. A summary of the data is given in appendix 1. Plot metadata are given in appendix 4.

Supplementary Materials
www.sciencemag.org/content/342/6156/1243092/suppl/DC1
Supplementary Text
Figs. S1 to S12
Tables S1 to S3
Appendices S1 to S4
References (S3–67)

11 July 2013; accepted 19 August 2013
10.1126/science.1243092