

Biodiversity and Conservation

Insights into regional patterns of Amazonian forest structure, diversity, and dominance from three large terra-firme forest dynamics plots.

--Manuscript Draft--

Manuscript Number:	BIOC-D-16-00197R1	
Full Title:	Insights into regional patterns of Amazonian forest structure, diversity, and dominance from three large terra-firme forest dynamics plots.	
Article Type:	Original Research	
Keywords:	aboveground biomass, abundance, forest conservation, Fisher's alpha, rarity, species richness	
Corresponding Author:	Alvaro Duque, Ph.D. Universidad Nacional de Colombia Sede Medellin Medellín, COLOMBIA	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:	Universidad Nacional de Colombia Sede Medellin	
Corresponding Author's Secondary Institution:		
First Author:	Alvaro Duque, Ph.D.	
First Author Secondary Information:		
Order of Authors:	Alvaro Duque, Ph.D.	
	Helene Muller-Landau	
	Renato Valencia	
	Dairon Cárdenas	
	Stuart Davies	
	Alexandre de Oliveira	
	Alvaro Pérez	
	Hugo Romero-Saltos	
	Alberto Vicentini	
Order of Authors Secondary Information:		
Funding Information:	Center for Tropical Forest Science - Forest Global Earth Observatory (CTFS-ForestGEO) of the Smithsonian Tropical Research Institute	Mr. Alvaro Duque
	Pontifical Catholic University of Ecuador (PUCE, research grants of Donaciones del Impuesto a la Renta from the government of Ecuador)	Mr Renato Valencia
	National Science Foundation (DEB-1046113)	Mr Stuart Davies
Abstract:	<p>We analyze forest structure, diversity, and dominance in three large-scale Amazonian forest dynamics plots located in Northwestern (Yasuni and Amacayacu) and central (Manaus) Amazonia, to evaluate their consistency with prevailing wisdom regarding geographic variation and the shape of species abundance distributions, and to assess the robustness of among-site patterns to plot area, minimum tree size, and treatment of morphospecies. We utilized data for 441,088 trees (DBH\geq1 cm) in three 25-ha forest dynamics plots. Manaus had significantly higher biomass and mean wood density than Yasuni and Amacayacu. At the 1-ha scale, species richness averaged 649 for trees \geq 1 cm DBH, and was lower in Amacayacu than in Manaus or Yasuni; however, at the</p>	

	<p>25-ha scale the rankings shifted, with Yasuni<Amacayacu<Manaus. Within each site, Fisher's alpha initially increased with plot area to 1-10 ha, and then showed divergent patterns at larger areas depending on the site and minimum size. Abundance distributions were better fit by lognormal than by logseries distributions. Results were robust to the treatment of morphospecies. Overall, regional patterns in Amazonian tree species diversity vary with the spatial scale of analysis and the minimum tree size. The minimum area to capture local diversity is 2 ha for trees ≥ 1 cm DBH, or 10 ha for trees ≥ 10 cm DBH. The underlying species abundance distribution for Amazonian tree communities is lognormal, consistent with the idea that the rarest species have not yet been sampled. Enhanced sampling intensity is needed to fill the still large voids we have in plant diversity in Amazon forests.</p>
Response to Reviewers:	<p>Medellín, August 8th, 2016</p> <p>Editorial office Biodiversity and Conservation</p> <p>Reference: BIOC-D-16-00197</p> <p>Dear Dr. R. Nidhi:</p> <p>According to the decision received on July 4th to the referenced manuscript submitted to Biodiversity and Conservation, in which we were invited to answer the comments and suggestions made by Dr. Hans ter Steege and an anonymous reviewer, herewith we are submitting the new version of the manuscript entitled "Insights into regional patterns of Amazonian forest structure, diversity, and dominance from three large terra-firme forest dynamics plots". As you will see, all comments and suggestions made by you and reviewers were carefully addressed. We left in red all changes made to the document to facilitate yours and reviewers assessments of them.</p> <p>We sincerely thank you all for the wonderful and fructiferous comments, which surely helped us to improve the quality of the study. In the response to reviewers, you will find the detailed answer to each one of the comments and suggestions.</p> <p>The new version of the manuscript has 5589 words in the main text, 1 tables, and 6 figures. We are also including a Supporting Information file, which has 4 tables and 6 figures.</p> <p>I hope you will find this new version suitable for publication.</p> <p>Thank you very much for your help.</p> <p>Yours Sincerely,</p> <p>Alvaro J. Duque M. Associate Professor Departamento de Ciencias Forestales Universidad Nacional de Colombia Email: ajduque09@gmail.com/ajduque@unal.edu.co</p> <p>ANSWER TO REVIEWERS</p> <p>I will use R for Reviewer (in italics) and A for Answer throughout this response to the comments made by the reviewers.</p> <p>REVIEWER 1</p> <p>R: The authors aim to understand regional forest structure and diversity in the Amazon basin by examining three large (25 ha) plots, an improvement over the traditional 1 ha plots that have been used to infer these regional properties. The paper is very well written and organized. I believe it will make a solid contribution toward how future field data collections are performed and data is analyzed to better understand regional species patterns in Amazonia.</p> <p>A: Thank you very much for your comments.</p>

Major comments:

R: I have concerns about the validity of the comparisons in the discussion between the 25ha plots and the smaller, more distributed plots. First, the 25ha plots are much more representative of their surrounding landscape, but without knowing more about the small plot sampling that you compare them to, this is hard to assess. Knowing how these small plots are generally distributed, I am pretty confident that they do not actually sample the landscapes very well, but other readers may not know this.

Providing a map or a description of the spatial distribution of the small plots that you are comparing your 25ha plots to may be necessary, and could be added to Fig 1.

A: I find a bit difficult to attend the suggestion made by the reviewer. Overall, we are not comparing with a specific dataset that can be mapped, but with different sources of information employed in published papers that can completely differ or even partially overlap between them. In many of these cases (perhaps in all of them), the data employed by each one of the studies uses a different dataset that varies in plot size, sample size, or size cutoff. For example, we are referring our comparisons to the work published by Duivenvoorden et al. (2011) that employs 0.1 ha plots (DBH \geq 2.5 cm) as well as to the work published by Ter Steege et al (2003) and Baker et al. (2004), which primarily use 1-ha plots. Therefore, we are comparing with the data reported in the literature itself rather than with some specific set of data. Including this sort of information in our map doesn't seem appropriate. The readers can look at the papers to see what data they are referred to.

R: Additionally, the comparison between the regional patterns of structural variation from this study and the small plot studies is tenuous. While the 25ha plots are more intensive local samples, they are extremely limited in the regional context. Whereas the diffuse small plots may actually be more representative of regional patterns, at least in terms of structural variables. I believe some discussion needs to be added to this section concerning the tradeoffs between sample size, scale, and plot distributions when assessing forest structural variables. However, I don't think this is necessary for the species diversity comparisons, where you do a nice job of discussing these issues already.

A: I completely agree with you that the use of scattered plots is more appropriate to assess variations in forest structure at regional scales. To accomplish with your suggestion, a new paragraph was included in page 16, lines 423 to 427, in which we acknowledge the limitations we have by employing contiguous plots instead of scattered plots.

R: You might want to also work on synthesizing the species diversity results and the structural variation results since they are largely kept separate in the discussion.

A: We added a couple of sentences at the beginning of the Conclusions and future directions section that aim to include the synthesis of structural outputs along with those related to diversity.

R: In the abstract and conclusion, you suggest the need for increased sampling intensity to better understand Amazon plant diversity. How should this be done? Just installing more 25 ha plots in more locations? Providing some more detail around the type and frequency of sampling would strengthen your conclusion and provide readers with more than just the problem that you identified.

A: I think this information is already included in the Conclusions and future directions section, where we suggest new strategies of sampling for understanding tree diversity in the Amazon terra firme forests based on plots \geq 2 ha that include all individuals with DBH \geq 1 cm. A couple of words emphasizing this issue were added in this paragraph.

Minor comments:

R: Line 209-212: how often did this occur in each case?

A: Please, check Table 1. You can deduce from this table how many individuals were fully identified and how many were treated as morphospecies.

R: Line 219: "honest"

A: Fixed.

R: Line 289/665: Be consistent in the terms that you use in the main text and in the Tables/Figs. Currently, NI and AGB are used in text vs. Individuals and Total biomass in the tables.

A: The terms NI, BA, and AGB were all included in Table 1 to be in line with the wording employed in the text.

R: Line 330-331: Wording here is confusing. Please rephrase.

A: Fixed

R: Line 389-391: It is hard to tell in what way your results are partially consistent with the literature without providing the values from the literature. Please provide the values and their associated references. This way the reader can better assess the consistency or inconsistency of your results and others.

A: The text was modified in this part. We added new information about the number of species we want to highlight here. The main source of comparison is Ter Steege et al. (2003; and the ATDN website) and the readers can have a look to these maps for comparisons, which actually do not provide values of species richness. Since this manuscript is already quite long, in this topic, we aimed to highlight the new insights from our new data rather than focusing on previous works.

R: Line 426: "At"

A: Fixed

R: Line 678: What is the source of the vector file that divides the Amazon basin into three regions?

A: The correspondent reference (Ter Steege et al. 2013) was added in the figure legend. Thanks for the note.

R: Line 692: The legend does not describe what is actually plotted. There are three solid and three dashed lines in red, green, and blue, but the legend has three solid color, one solid grey, and one dashed grey. Unless you are indicating that the colors describe the plots and the line style the DBH class? In which case there should be a better, less confusing way create a legend to show that. Same for Fig 4.

A: In my opinion, the reviewer already saw what we wanted to show in this legend (underlined text in her/his comment). It is a good way to shorten the number of lines required for the legend.

R: Line 711: Missing letters for each subfigure in the caption. Check other captions for this as well.

A: The letters were included in the caption.

REVIEWER 3 (Hans ter Steege)

R: P(age) 4 L(ine) 70. Condit et al did not suggest that "previous studies in tropical forests showed that local samples of fewer than 3000 individuals provide poor estimates of expected diversity patterns at large spatial scales". The differences in Condit et al 1998 for c. 500 stems (1 ha) and 2000 stems are quite small and Condit et al suggest to use these samples stratified to equal individuals samples. Fisher's alpha of a single community is theoretically independent of sample size, but this does not hold for very small samples. However, there is also a demonstrable effect of area on Fisher's alpha (Hubbell 2013, cited in this MS). So if area increases, so does the variation in environment and Fisher's alpha.

A: The wording associated to Condit et al. (1998) was changed. Effect of area on Fisher's alpha is treated in the text in relation to the 75 ha sampled in this study.

R: P4L81 and onwards. Yes, there is a general decrease in diversity from west to east but there is a band of high diversity from WA to CA. The decrease is mainly at the edges. Plots in CA (around Manaus and Urucu) can have diversities as high as those in the richest plots in WA. So while it is true that it has been shown that forests have higher AGB in CA than in WA, this is not the case for alpha diversity. So the expectation of a decrease in diversity is not supported by the papers cited.

A: All this text was changed according to the suggestions made by Hans. We included the idea of a continuous band of high tree diversity around 5° S instead of a systematic

change from west to east. The research questions were also modified accordingly.

R: P5L109 But do realize that many of the very rare species are likely/possibly common somewhere else. We discuss this in an upcoming paper but also see Magurran who sees a lognormal for the species belonging to an area and a change into a log-series when the rare species (“vagrants”, as named by here) are taken into account.

A: Although I understand the point made by Hans here, I apologize but I don’t see the need of changing anything.

R: P6L126 But Pos et al made this statement only for analyses that use differences in composition and state that for comparison of diversity all species have to be taken into account, as the species usually unidentified are the rare ones. This strongly affects Fisher’s alpha.

A: In the introduction and section methods we made clear that we are interested in analyzing the main patterns of change between morphospecies and named species instead of net values, acknowledging the findings of Pos et al. (2014).

R: P6L141

Hypothesis i) diversity does follow from the ‘prevailing wisdom’.

Hypothesis ii) strongly if diversity is analyzed, as shown by Pos et al!

P9L232 Diversity with and without the morpho-species will differ predictably as shown by Pos et al.

A: Changes were made to the research questions in order to accomplish with the suggestions and comments made by Hans on these points.

R: P14L358

While the data I know of seems to agree with the pattern of alpha increasing with sample size (below randomized samples of Amacayacu), in other places (e.g. BCI) alpha stabilizes much earlier. Image further below.

A: Our main goal is to highlight and compare our findings within terra firme forests of the Amazon basin. Many other tropical forests, such as the over-studied dry BCI forests, can have a different pattern. The comparison we made among these three sites speaks for itself.

R: However for the fit of the logseries, I differ in opinion. In the top graph left is the unbinned RAD of Amacayacu with calculated lognormal (blue) and logseries (red). The hypothetical log-series of N and S for Amacayacu is near exact, even for the rare species. I am not sure how to interpret the analyses made for this manuscript.

A:

1) To analyze the Species Abundance Distributions (SADs) in the three plots we used the sads library available in R ((Prado PI, Miranda MD (2013); Fitting species abundance model with maximum likelihood. Quick reference for sads package. 1–20). As noted in the methods section, we used the poisson lognormal, which accounts for singletons and doubletons, instead of the truncated log normal, which assumes a lack of information. Based on the AIC assessed for both models, there is not way to claim that the logseries fits better than the lognormal. I kindly ask Hans to use this library and compare the results with those he has and tell us (we and the editor) if he finds a mistake with the analyses presented.

2) I don’t know which dataset is Hans employing to illustrate the case of Amacayacu. Until I know, we gave him some preliminary data of the largest trees (DBH ≥ 10 cm) of Amacayacu in the past. Neither Dairon Cardenas nor I have given Hans the last version of the dataset. We all know this kind of datasets is dynamic in terms of taxonomy. We are using the dataset currently available in the CTFS database. Since we are just finishing the second census, all updates to the new version of the database are expected to be ready at the end of this year.

3) I am presenting here the RADs for the three sites assessed for all individuals (DBH ≥ 1 cm) including morphospecies (Figure 1), where the same pattern observed with the SADs (octaves method) emerged: logseries underestimates the middle abundance species and generates more rare species than observed. As you can see, in all cases the fit of the logseries underperforms the fit of the lognormal.

Figure 1. RADs for the three 25 ha plots surveyed in the Amazon

R: I would like to be able to test this for the Yasuni site as well with our scripts. Data with species named from sp1 to sp900 or so, would work.

A: I apologize, but I don't have the authorization to distribute this dataset. What I would suggest is the ATDN to release the scripts you are using to prepare these figures and models to enable us to compare between scripts and outputs (I don't expect any difference anyway). The biggest advantage of the sads library is the use of likelihood methods to estimate the parameters and the calculus of the AIC to compare between models.

R: P15L416. As said above, the east-west pattern does not hold for a band along the Amazon. For the latest maps see ATDN website. The discussion on soil fertility is valid, although the areas with the lowest fertility area north of Manaus (Upper Rio Negro) and Guyana.

A: In both, introduction and discussion the issues related to the wrongly claimed east-west patterns were all modified.

R: I would also agree that taxonomic detail is influenced by area. Likely, the number of species is to increase in Ecuador if more collections are made and more work can be done on its flora.

A: I think this issue is out of the scope of this study.

R: P17L440. The Mid-domain is unlikely to play an effect as the ranges of most species are small, so the edge effect of the mid-domain is only found in a really small fringe of the Amazon.

A: This idea was deleted following your comments.

R: P17L481. Based on the simple analysis above, I do not agree with this statement. It seems the Logseries give a much better prediction than the lognormal. I see that in most of our samples.

I also have great problems with the veil, the unsampled species, as all species in the sample have been identified. So there are basically no missing species.

A: I disagree with Hans: the lognormal in all cases gave a better fit than logseries. I think it is a new contribution to the debate about the underlying model of species abundance distributions in the Amazon forests, which is based on more intensively samples of local communities.

R: P17L491. Maybe, but in a publication to appear in a week or so we show that 12,000 species of tree have been collected in the Amazon already with very low sample effort (500,000 collections). We also show that singletons are often species from other areas (which we predicted in 2013). So they may not be threatened in the Amazon. However, as deforestation rates in the areas where these species originate (Cerrado, Atlantic forest, Andean slopes) have much higher deforestation rates than the Amazon, this does not make them less threatened.

A: Independent of the range of species distribution, if we have the lognormal instead of the logseries as the underlying model, our claims remain valid. Some of the topics nicely highlighted in the new paper of Ter Steege et al are out of the scope of this study.

R: L493 – Here I agree completely but see above.

A: See also above.

R: P19. Do note that the estimate of ter Steege et al is not just an extrapolation of plot data. We estimated Amazon-wide population sizes of trees from our plots and extrapolated the curve from this data. Interestingly, calculating with just the plot data gave the same c. 16,000 spp.

A: The world extrapolation was changed by estimation. We are not interested in this issue at all due to the data we have is not appropriate to address this question at a regional scale (too few points).

Insights into regional patterns of Amazonian forest structure, diversity, and dominance from three large terra-firme forest dynamics plots.

Alvaro Duque^{1*}, Helene C. Muller-Landau², Renato Valencia³, Dairon Cardenas⁴, Stuart Davies², Alexandre de Oliveira⁵, Álvaro J. Pérez³, Hugo Romero-Saltos^{3,6}, and Alberto Vicentini⁷.

¹: Departamento de Ciencias Forestales, Universidad Nacional de Colombia – Sede Medellín. Calle 59A #63-20. Medellín, Colombia. Email: ajduque09@gmail.com/ajduque@unal.edu.co

²: Center for Tropical Forest Science-Forest Global Earth Observatory, Smithsonian Tropical Research Institute, PO Box 37012, Washington, DC 20013-7012, USA. Email: mullerh@si.edu - DaviesS@si.edu

³: Laboratorio de Ecología de Plantas, Escuela de Ciencias Biológicas, Pontificia Universidad Católica del Ecuador. Av. 12 de Octubre 1076 y Roca, Ed. De Ciencias. Email: lvalencia@puce.edu.ec; ajperezc@puce.edu.ec

⁴: Herbario Amazónico Colombiano, Instituto Amazónico de Investigaciones Científicas Sinchi. Calle 20 # 5-44. Bogotá, Colombia. Email: dcardenas@sinchi.org.co

⁵: Instituto de Biociências, Universidade de São Paulo, Rua do Matão, 321, Travessa 14, São Paulo, SP 0550890, Brazil. Email: adalardo@usp.br

⁶: Universidad Yachay Tech, Ciudad del Conocimiento, 100119 Urcuquí, Ecuador. Email: hromero@yachaytech.ecu.ec

⁷: Instituto Nacional de Pesquisas da Amazônia, Av. André Araújo, 2936, CP 478. Manaus, AM 69060001, Brazil. Email: vicentini.beto@gmail.com

Running-title: Amazon forests diversity

28 ***Correspondent author:** Alvaro Duque

29

30 **Number of words in the abstract:** 249

31

32 **Number of words in main body of the paper:** 5589

33

ABSTRACT

We analyze forest structure, diversity, and dominance in three large-scale Amazonian forest dynamics plots located in Northwestern (Yasuni and Amacayacu) and central (Manaus) Amazonia, to evaluate their consistency with prevailing wisdom regarding geographic variation and the shape of species abundance distributions, and to assess the robustness of among-site patterns to plot area, minimum tree size, and treatment of morphospecies. We utilized data for 441,088 trees (DBH \geq 1 cm) in three 25-ha forest dynamics plots. Manaus had significantly higher biomass and mean wood density than Yasuni and Amacayacu. At the 1-ha scale, species richness averaged 649 for trees \geq 1 cm DBH, and was lower in Amacayacu than in Manaus or Yasuni; however, at the 25-ha scale the rankings shifted, with Yasuni<Amacayacu<Manaus. Within each site, Fisher's alpha initially increased with plot area to 1-10 ha, and then showed divergent patterns at larger areas depending on the site and minimum size. Abundance distributions were better fit by lognormal than by logseries distributions. Results were robust to the treatment of morphospecies. Overall, regional patterns in Amazonian tree species diversity vary with the spatial scale of analysis and the minimum tree size. The minimum area to capture local diversity is 2 ha for trees \geq 1cm DBH, or 10 ha for trees \geq 10cm DBH. The underlying species abundance distribution for Amazonian tree communities is lognormal, consistent with the idea that the rarest species have not yet been sampled. Enhanced sampling intensity is needed to fill the still large voids we have in plant diversity in Amazon forests.

Keywords: aboveground biomass, abundance, forest conservation, Fisher's alpha, rarity, species richness.

INTRODUCTION

The Amazon basin harbors the largest and most species-rich tropical forest on earth (Myers et al. 2000; Slik et al. 2015). An understanding of how forest structure and diversity vary across the whole Amazon region is critical for the development of effective regional conservation strategies. So far, these questions have been addressed mostly using census data for plots of ~1-ha in area that included only trees with diameter at breast height (DBH) ≥ 10 cm, henceforth referred to as large trees (Gentry 1988b; Phillips et al. 1998; Ter Steege et al. 2003; Phillips et al. 2004; Ter Steege et al. 2006). However, previous studies in tropical forests **out of the Amazon region** showed that local samples of fewer than 3000 individuals **tend to underestimate differences among species-rich sites** (Condit et al. 1998), and these small plot data all fall within that sample size. Here, we take advantage of a new dataset of three intensively sampled large permanent plots (25-ha each; DBH ≥ 1 cm) located in central (CA) and northwestern Amazon (NWA) to test the validity of the prevailing wisdom (founded on small samples) about forest structure and diversity in Amazon forests.

Assessments of the variation in forest structure and diversity based on intensively sampled tree communities may inform us about how to implement sound programs of forest management and conservation in this important ecosystem.

In the Amazon basin, rainfall seasonality, soil fertility, and forest turnover rates **have been** associated with local variation in forest productivity and patterns of **tree density or number of individuals** (Ter Steege et al. 2003, Duivenvoorden et al. 2011) and aboveground biomass (AGB) (Malhi et al. 2006; Saatchi et al. 2007). **At a regional scale, tree density has shown to be positively associated with the dry season length** (Ter Steege et al. 2003); **on contrast, at a local scale, soil fertility seems to be inversely associated with it** (Duivenvoorden et al. 2011). Variation in forest AGB within the Amazon is also strongly positively associated with variation in wood density (WD),

which is higher in areas of more seasonal rainfall, lower soil fertility, and lower forest turnover, while basal area shows little regional variation (Baker et al. 2004). Forests with richer soils and more constant climates also show higher turnover (recruitment and mortality) and systematic differences in tree species functional composition, a pattern that reaches its apogee in the NWA region (Phillips et al. 1994; Phillips et al. 1998; Ter Steege et al. 2006). However, tree species richness and diversity from northwestern (NWA) to central (CA) Amazonian terra firme forests did not significantly vary along a geographical band around 5° S (Ter Steege et al. 2003). Tree diversity around the regions of Urucu and Manus in CA has found to be as high as that reported in the richest plots of NWA (<http://atdn.myspecies.info>). Therefore, considering that our three study sites are located in the NWA (Yasuni and Amacayacu) and CA (Manaus), and assuming that these sites reflect the same general patterns in structure and diversity as the previously censused 1-ha plots, we expect a systematic decrease in tree density, a systematic increase in AGB, but similar values of species richness and diversity from west to east (Yasuni to Amacayacu to Manaus).

The design of effective conservation strategies also depends on knowledge of patterns of rarity and dominance (Pitman et al. 1999; Pitman et al. 2001), and thus on species abundance distributions. Species abundance distributions are commonly modeled as either lognormal or logseries (McGill 2003). Both these distributions predict relatively few dominant species, but they differ in the expected number of rare species. Under the logseries distribution, most species are rare (Fisher et al. 1943); whereas under the lognormal distribution, most species have intermediate abundance, with few rare species (Preston 1948; McGill 2003; Connolly 2005; Connolly et al. 2014). If the logseries is the better model of species abundance distributions, this implies the existence of many rare species, which means that a much more radical conservation strategy is required to avoid considerable species extinction (Hubbell *et al.*, 2008). Past analyses of tree species abundance distributions in Amazonia have been based on analyses of 1-ha, large-tree censuses, either individually or pooled, and have generally supported the use of the logseries distribution (Hubbell et al. 2008; Ter Steege et al.

2013; Slik et al. 2015). However, we might expect that abundance distributions could take different shapes when smaller woody individuals are included, and at different spatial scales. Here, we evaluate the fit of the logseries relative to the lognormal for larger plots, and for all trees (≥ 1 cm dbh), thus testing the appropriateness of the logseries as the underlying species abundance distribution (SAD) model in tropical forests (Hubbell 2001; Hubbell et al. 2008; Hubbell 2013; Ter Steege et al. 2013; Slik et al. 2015).

One of the main difficulties of working **with** species-rich communities like the Amazon is plant identification, and this raises the question of the degree to which observed diversity patterns depend on taxonomic resolution. A previous analysis based on the three plots employed in this study showed that differences among research teams in morphotyping non-fully identified specimens could lead to biases in plant classification (Gomes et al. 2013). In contrast, Pos et al. (2014) argued that analysis that includes only fully botanically identified species (hereafter referred to as named species) should find patterns **of species similarity between sites** similar to those obtained when also including morpho-types (hereafter referred to as morphospecies), **but not for assessments of species diversity**. Here, we quantitatively assess whether **the main pattern and trend of variation in species relative abundance distributions and diversity patterns, rather than the net values**, are indeed robust to this difference in taxonomic resolution. **For the species abundance models, we expect an increase of rare species due to the inclusion of morphospecies, and thus, a better fit by the logseries than by the lognormal; on contrary, the exclusion of non-fully identified species would promote a better fit of the lognormal than the logseries (see Pos et al. 2014), which also have implications for the shape of the diversity curves produced by both datasets.**

In this study, we analyze patterns of variation in forest structure, diversity, and dominance across CA and NWA, to evaluate their consistency with prevailing wisdom regarding geographic variation, species abundance distributions, and robustness to

details of the census methods. We use a dataset composed of 441,088 individuals (DBH ≥ 1 cm) surveyed in the three most intensively sampled large permanent plots that currently exist in the Amazon basin, all associated with the Center for Tropical Forest Science - Forest Global Earth Observatory (CTFS-ForestGEO) global network. We address the following specific questions:

- i) Is there systematic variation from central to northwestern Amazonia in tree density **and** aboveground biomass **as reported from 1 ha plots?**
- ii) **Can we confirm the existence of a high tree species diversity band around the 5° S in the NWA and CA Amazon?**
- iii) How do the details of tree censuses, specifically differences in minimum tree size (DBH ≥ 1 cm *versus* DBH ≥ 10 cm), plot size (1-ha *versus* 25-ha), and taxonomic resolution (all morphospecies *versus* only named species), affect among-site patterns in forest structure?
- iv) Are species abundance distributions better fit by the logseries or the lognormal?
- v) What insights into metacommunity diversity and abundance patterns can we obtain by combining data from multiple large plots?

METHODS

Study sites

Data were collected in three permanent 25-ha plots established in terra-firme forests located in NWA and CA in Ecuador (Yasuni), Colombia (Amacayacu), and Brazil (Manaus), respectively. These plots are arrayed roughly on a straight line, with distances of 700 km between Yasuni and Amacayacu, 1100 km between Amacayacu and Manaus, and 1800 km between Yasuni and Manaus (Fig. 1). Yasuni National Park and Biosphere Reserve and the adjacent Huaorani Indian territory cover 1.6 million ha of forest and form the largest protected area in Amazonian Ecuador (Valencia et al.

2004). Amacayacu National Park covers around 220.000 ha of forest and is part of the protected system of national parks in the Colombian Amazon. The Manaus plot is located about 80 km north of the city of Manaus. The three plots are all located on terra-firme forests at elevations below 200 m asl. Precipitation at Yasuni and Amacayacu is aseasonal, with mean annual rainfall ~3000 mm and no months with less than 100 mm. Mean annual rainfall at Manaus is ~3500 mm, with a dry season of 1-2 months between June and October (Sombroek 2001).

Tree censuses

In each 25-ha plot (500×500 m), each individual free-standing woody plant with a DBH ≥ 1 cm was mapped, tagged and measured, including shrubs, trees, and palms (but not lianas). Multiple stems were separately recorded. Voucher collections were made for each unique species in each plot. We collected vouchers in all cases in which there was any doubt about a plant's similarity with another individual that was already collected within the same plot. The taxonomic identifications were made by comparing the specimens with herbarium material and with the help of specialists. All of the samples are kept at the COAH, QCA, and INPA Herbaria. We assumed that all specimens with the same botanical name represented the same species, even though we did not standardize the taxonomy between plots. The plants that could not be identified as named species were separated into morphospecies that were treated as distinct species. Variation between sun-exposed and shaded leaves and between young and old leaves was documented in vouchers deposited in reference collections, to avoid splitting species with high plasticity and/or ontogenetic variation. Identifications were done by separate teams at each site, and thus there may be differences in the species concept between sites. For instance, the morphospecies classification in the Amacayacu and Yasuni plots was conservative with morphospecies including a relatively wide range of variation, while in Manaus, the classification allowed less variation within morphospecies.

Structural variation

We analyzed the variation in the number of individuals (NI), basal area (BA in m²), aboveground biomass (AGB in Mg), and mean wood density (WD in g cm⁻³). The aboveground biomass (AGB) of each tree (in kg) was calculated using the general model without tree height developed for tropical forests by Chave et al. (2014), which employs DBH, wood specific gravity (referred to as wood density, WD), and a new site-specific environmental variable called E. The new parameter E is a coefficient derived from global databases on temperature seasonality (TS), the maximum climatological water deficit (CWD), and precipitation seasonality (PS) (Chave et al. 2014). The equation is

$$AGB = \exp(-1.802 - 0.976 * E + 0.976 * \log(WD) + 2.673 * \log(DBH) - 0.0299 * (\log(DBH))^2)$$

Wood density values of each species found in all plots were assigned following Chave et al. (2006), Zanne et al. (2009), and databases compiled by the CTFS-ForestGEO. In cases in which we could not assign a WD value at the species level, we used the average value at the genus or family level. For individuals without a botanical identification, we used the average WD value of all other individuals found in the same plot. E takes value -0.075 in Amacayacu, -0.111 in Manaus, and -0.023 in Yasuni. The total AGB in each quadrat, subplot, or plot was obtained by summing the AGB of all trees present including palms, but excluding lianas and tree ferns.

We used One Way Anova (ANOVA) to test for significant differences in the stand-level mean NI, BA, AGB, and WD at the 1-ha scale. That is, we divided each plot into 25 square 1-ha subplots (100×100 m) and treated these as replicate samples. When significant differences were found, a Tukey's honest significant difference (TukeyHSD)

test was used to compare the main trend of variation between sites. ANOVAs were done separately for two size categories: all individuals with DBH ≥ 1 cm (hereafter referred to as *all individuals*) and only individuals with DBH ≥ 10 cm (hereafter referred to as *large individuals*). For each site, we also characterized the distributions of these structural variables at the scale of 20×20 m (0.04 ha) quadrats using probability density functions.

Species diversity patterns

As above, we used ANOVAs and a subsequent TukeyHSD test to evaluate differences in species richness (SR) and species diversity (SD; assessed by the Fisher's alpha index) among sites, based on 25 square 1-ha subplots (100×100 m) for each site. ANOVAs were performed for both size categories (all individuals and large individuals) and for both morphospecies and named species. The morphospecies dataset included all named and unnamed species that were compared with each other and classified as different within each site based on the morphology of vegetative characters, excluding individuals not collected and those for which no morphospecies assignment was possible. The named species dataset contained all individuals identified to species, and excluded non-fully identified species and uncollected individuals. We analyzed both morphospecies and named species in order to understand and compare results under these approaches, and thereby identify the uncertainty associated with the morphotyping of sterile specimens by different teams at different sites (Gomes et al. 2013). Overall, we are interested in evaluating whether the main pattern of variation within plots change with the use of either named species or morphospecies, rather than to compare the net values of diversity estimated by each one of them, which are expected to differ (Pos et al. 2014).

We used species-individual curves and graphs of Fisher's alpha vs. area (henceforth Fisher's alpha-area curves) to describe the overall patterns of species diversity at both

plot and meta-community scales. We chose to use Fisher's alpha over other commonly used diversity metrics both because of its conceptual roots (Fisher et al. 1943; Hubbell 2001) and because it is relatively less dependent on sample size than other metrics (e.g., Condit et al. 1996). At the plot scale, the development of the species-individual and Fisher's alpha-area curves followed the approach of Condit et al. (1996). To build the species-individual curves at the plot scale, we employed 100 randomly chosen points as centers of progressively larger plots. The size of the square plots employed to build the curves increased from 0.01-ha (10×10 m) to 25-ha (500×500 m) (0.01-ha, 0.04-ha, 0.25-ha, 1-ha, 2-ha, 3-ha, 4-ha, 5-ha, 10-ha, 15-ha, 20-ha, and 25-ha). At the plot scale, species-individual and Fisher's alpha-area curves were analyzed for both morphospecies and named species datasets, for both all individuals and large individuals. To perform these analyses we used the CTFS R package (<http://ctfs.arnarb.harvard.edu/Public/CTFSRPackage/>). While the plot-level analysis sampled individuals within contiguous areas, the meta-community analyses were based on 500 random draws from the full merged dataset, with species-individuals curves based on random draws of individuals, and Fisher's alpha-area curves based on random draws of complete 1-ha plots. The metacommunity analyses were performed only on named species because morphospecies could not be matched across plots. The metacommunity analyses of species-individuals and Fisher's alpha – area curves were done using the vegan library for R (Oksanen et al. 2013).

Species abundance distributions

We analyzed species abundance distributions (SAD) at the plot and metacommunity scales. We characterized and fit the SAD for each plot, for named species and morphospecies as well as for all individuals and large individuals. At the metacommunity scale we characterized and fit the SAD only for named species in both size classes (as in ter Steege *et al.*, 2013; Connolly *et al.*, 2014; Slik *et al.*, 2015). We used maximum likelihood methods to fit the lognormal (specifically the Poisson-lognormal) and logseries to each distribution (Prado and Miranda 2013), choosing

these models because they have been found to be the most suitable SAD models for species rich communities (Wilson 1991; Hubbell 2001). We ranked models using the Akaike Information Criterion (AIC).

All statistical analyses were performed using the Statistical Software R version 3.02 (R Development Core Team 2014).

RESULTS

Structural variation

A total of 441,088 individuals with DBH ≥ 1 cm and 46,456 individuals ≥ 10 cm were recorded in the three 25-ha plots. When each plot was divided into 25 1-ha subplots, there were significant differences among sites in NI, BA, AGB, and WD for all individuals and large individuals (Table 1). Amacayacu had significantly lower values of NI and BA than Yasuni for all individuals (DBH ≥ 1 cm) and large individuals (DBH ≥ 10 cm). Manaus was similar to Yasuni in NI and BA of all individuals, similar to Amacayacu in the NI of large individuals, and indistinguishable from the other two sites in the BA of large individuals. The central Amazonian site of Manaus had significantly higher AGB and mean wood density than the two northwestern Amazonian sites for both all individuals and large individuals. Amacayacu also had significantly higher mean wood density than Yasuni. The distribution of structural parameters across 20x20 m quadrats illustrated the patterns found with 1-ha subplots in greater detail (Fig. 2). Overall, for all individuals, the distribution of NI differed noticeably among all three plots (Fig. 2A), while BA distributions were remarkably similar except for the longer tail due to the presence of larger trees in Manaus (Fig. 2B). WD varied strongly across sites with the highest values in Manaus (Fig. 2D), which then translates into the AGB distributions, where Manaus again

stands out (Fig. 2C). For large individuals (DBH ≥ 10 cm), the basic patterns of variation remained almost the same for BA, GB, and WD, but NI was partially reversed, being higher in Yasuni than in Amacayacu and Manaus (Table 1; Fig. 2E).

Species diversity

A total of 2993 morphospecies, belonging to 419,576 individuals with DBH ≥ 1 cm (95% of total) were recorded in the three 25-ha plots, of which 70% were fully identified to species. The 2095 fully identified species (named species) accounted for 83% of the total number of individuals. When all individuals ≥ 1 cm were included, 1-ha subplots had average species richness of 649 ± 41 for morphospecies and 513 ± 32 for named species, with significantly lower richness in Amacayacu than in Yasuni and Manaus for both morphospecies and named species (Table 1). When only large individuals (≥ 10 cm) were included, species richness averaged 234 ± 19 for morphospecies and 204 ± 16 for named species, with Amacayacu again showing the lowest value and Manaus the highest (Table 1). The sites had a different ranking in species richness at the 25-ha scale, with Yasuni having the fewest morphospecies and named species for all individuals and large individuals, while Manaus had the most (Table 1). For all individuals the pattern of among-site variation in diversity, as measured by the mean Fisher's alpha in 1-ha subplots, very much resembled the pattern of species richness. However, diversity pattern for large individuals in 1-ha subplots differed, with Manaus showing markedly higher Fisher's alpha values than Yasuni and Amacayacu for both morphospecies and named species (Table 1). Among-site patterns in species richness and diversity in 1-ha subplots were qualitatively similar whether analyzing morphospecies or just named species.

Species-individuals patterns showed different patterns of variation between the size categories **among plots**. Overall, large individuals in Yasuni showed a higher number of species for a given number of individuals than Amacayacu and Manaus. In contrast,

for all individuals, Yasuni showed a lower number of species for a given number of individuals than Amacayacu and Manaus, which followed exactly the same pattern of species accumulation with increasing sample size (Figure 3). In Yasuni, a sample of a given number of large individuals had more species than an equivalently sized sample of all individuals, while Manaus showed the opposite pattern and Amacayacu had similar numbers of species in both size classes (Figure S1, Table 1). These patterns were qualitatively the same whether analyses were restricted to named species or not.

Fisher's alpha varied strongly with area in all analyses, with considerable variation in Fisher's alpha-area between size categories and among sites. In the 25-ha plots, large individuals in Yasuni ($\text{DBH} \geq 10$ cm) had the highest Fisher's alpha, but all individuals ($\text{DBH} \geq 1$ cm) the lowest (Figure 4). All the curves showed a strong increase to 1 ha. Above 1 or 2 ha, the curves for all individuals tended to plateau (Amacayacu and Manaus) or even decrease (Yasuni). In contrast, the curves for large individuals continued to increase with area to larger areas, at best plateauing above 4-10 ha. Fisher's alpha values for all individuals were larger than those for large individuals at areas < 1 ha in all sites, with divergent patterns at larger areas. At Manaus and Amacayacu, the differences between the curves declined above 1 ha, and at Amacayacu the curves actually crossed above 10 ha and values remained quite similar beyond that. In contrast at Yasuni, the curves crossed between 1 and 2 ha, with Fisher's alpha for large individuals becoming increasingly larger than that for all individuals at larger areas (Figure S2). The observed patterns were very similar for morphospecies compared with named species.

Species abundance distributions

Species abundance distributions in all three 25-ha plots were better fit by the lognormal than by the logseries, for both morphospecies and named species as well as

for all individuals and large individuals (Figure 5, Figure S3, Table S2). Although the lognormal model tended to systematically underestimate the number of the rarest species (those with just 1 individual in 25 hectares), it performed better at fitting the number of species with the most common intermediate abundances than the logseries. In contrast, the log series tended to systematically overestimate rare species and to underestimate those with intermediate abundances for both all individuals (Figure 5) and large individuals (Figure S3). The observed patterns were similar for morphospecies and for named species.

Metacommunity patterns

The metacommunity species-individual curves based on random draws of individuals of named species from across all three plots showed higher species richness in samples of all individuals than in equal-sized samples of just large individuals (Fig. 6a). These differences were statistically significant in samples of 2000 or more individuals (Table S3). The Fisher's alpha vs. area curves for all and large individuals crossed, with the all individuals curve showing higher diversity below 12 ha, and the large individuals higher diversity at larger areas (Fig. 6B, Table S4). The SADs for both all individuals and large individuals showed the same shape, but with a considerable increase in the number of rare species in the latter (Fig. 6C). For both SADs (all individuals and large individuals), the lognormal provided a better fit than the logseries (Figure S4).

DISCUSSION

Structural variation

The density of large trees ($\text{DBH} \geq 10 \text{ cm}$) was partially consistent with literature findings of a decrease from west to east (see Figure 3 in Ter Steege et al. 2003), with Yasuni showing the highest values and Amacayacu and Manaus substantially lower values. We expect soil fertility to be highest in Yasuni (Lips and Duivenvoorden 2001) and lowest in Manaus (Sombroek 2000), and thus our findings partially agree with the hypothesis that soil fertility drives large individual density in the Amazon terra firme forests (Ter Steege et al. 2003). In contrast, the density of all individuals ($\text{DBH} \geq 1 \text{ cm}$) showed a different pattern, with Manaus having the highest values, Yasuni the next-highest, and Amacayacu a much lower value (Table 1). High densities of small individuals at Manaus can perhaps be explained by lower soil fertility, which is expected to promote increases in plant defenses and reduction in mortality of juveniles and shrubs (Duivenvoorden et al. 2005). In contrast, high densities at Yasuni might be explained by higher turnover and local disturbance rates (Phillips et al. 1994; Phillips et al. 1998). Higher rates of disturbances in the more fertile soils of Yasuni than in the other two sites may also in part explain why this site has the lowest mean wood density (Ter Steege et al. 2006). We must acknowledge that a regional sampling of spread out small plots can represent better the structural variation than contiguous samples as those employed here. However, in the long-term the large permanent plots will surely help to identify the mechanisms acting on a fine-grain resolution that determines the structural variation of tropical forests at local scales.

In accordance with expectations, aboveground biomass was similar in the two northwest Amazon plots, and higher in the eastern central Amazon plot of Manaus. However, forest basal area was similar in Yasuni and Manaus, and considerably lower in Amacayacu. Thus, differences in wood density among plots appear as the main driver of the observed variation in aboveground biomass. Amacayacu had somewhat higher wood density than Yasuni, thus compensating for its lower basal area (Figure 2). Likewise, Manaus's much higher wood density clearly explains its higher aboveground biomass relative to Yasuni, which had the same basal area. Therefore, our results agree with previous findings from 1-ha plots that identified wood density

as a major driver of regional variation in aboveground biomass in Amazonian terra firme forests (Baker et al. 2004). We obtained the same among-site pattern with the older moist forest biomass allometry equation of Chave *et al.* (2005), which yielded higher mean biomass values than the new model without height proposed by Chave *et al.* (2014: see Table 1): 298.5 Mg ha⁻¹ for Amacayacu, 297.7 Mg ha⁻¹ for Yasuni, and 380.6 Mg ha⁻¹ for Manaus. This demonstrates that the among-site pattern is not merely a consequence of the new environmental factor (*E*) introduced in Chave *et al.* (2014).

Species diversity

Our results confirm the existence of a high tree species diversity band around 5° S in the NWA and CA as proposed by Ter Steege et al. (2003). A mean value of 649 ± 50 species (DBH ≥ 1 cm) per hectare is an unprecedented value of tree species richness that exceeds any previous report made in tropical forests. However, within this geographic band, we found differences in both tree species richness and diversity between plots, which also varied according to size. At the 1-ha subplot scale and for large individuals (DBH ≥ 10 cm), species richness and diversity patterns followed the not systematic west-east trend Yasuní > Manaus > Amacayacu. For all individuals (DBH ≥ 1 cm) and at the 1-ha scale, Manaus was as rich and diverse as Yasuni, with Amacayacu again having the lowest diversity. Therefore, for all individuals, this result is inconsistent with the hypothesis that species richness and diversity increase with soil fertility (after Gentry, 1988). To some extent, it could be argued that our results are likely influenced by the different taxonomic treatment of species at each site. However, the relatively large differences found here, and their consistency in the named species dataset, suggest that such results reflect patterns that can be found even if we standardize the taxonomy across the three sites. Competing theories could explain the high species richness and diversity found in Manaus. First, the greater age of CA relative to the younger areas of NWA may have provided a longer time for species to arrive via dispersal. In contrast, the high species richness of Yasuni and

NWA in general, may in part reflect higher speciation rates triggered by the uplift of the Andean mountains (Hoorn et al. 2010), which could partially balance the lower time and opportunity to accumulate species.

Among-site patterns in the species-individuals and Fisher's alpha-area curves were dependent on both sampled area and size class. The species-individual curves assessed at 10,000 large individuals (DBH \geq 10 cm) or more showed the Yasuni region as the most diverse and Manaus the least. In contrast, if all individuals (DBH \geq 1 cm) are considered, the expected trend was basically reversed: Manaus and Amacayacu were more diverse than Yasuni at sample sizes larger than 20,000 individuals. At samples of less than 1000 individuals, it was difficult to differentiate the curves for all individuals among plots (Condit et al. 1996). For large individuals, at sample sizes of less than 1000 individuals, Yasuni appeared on top of the other two plots, thus confirming the high diversity of large trees reported for the Andean foothills (Gentry 1988a; Ter Steege et al. 2003).

For Fisher's alpha-area curves, the most striking pattern was the one found in Yasuni, where the accumulation trend in the Fisher's alpha of all individuals and large individuals took different directions at sample sizes larger than 1-ha. In Yasuní, the Fisher's alpha of all individuals showed a clear trend to systematically decrease with areas above 1 ha, whereas the value for large individuals continued to increase albeit at a progressively slower rate. The lack of an asymptote in the Fisher's alpha for all individuals in Yasuní does not support the logseries expectation of a linear species accumulation with sample size (Hubbell 2001; Hubbell 2013), which challenges the use of this function to extrapolate species richness to larger geographical areas (e.g., Hubbell *et al.*, 2008). In the other two sites, Fisher's alpha in samples of all individuals tended to level off around 1 ha or earlier, suggesting that samples incorporating all individuals should be considered more appropriate to extrapolate species richness at larger areas than samples based on only large individuals (DBH \geq 10 cm). At sample sizes larger than or equal to 10 ha, diversity patterns for different minimum individual

sizes in Amacayacu and Manaus tended to converge and asymptote, suggesting that 10 ha might be a minimum ideal sample size to assess Fisher's alpha in local surveys based only on larger trees, particularly in cases in which the aim is to **estimate** species richness **in large geographic regions** (e.g., Ter Steege *et al.*, 2013).

Species abundance distribution models of independent communities

The results of this study are inconsistent with the hypothesis that the logseries is the "universal" SAD model that best fits the relative abundance distributions of tree communities in tropical forests (Hubbell 2001; Hubbell *et al.* 2008; Hubbell 2013; Ter Steege *et al.* 2013; Slik *et al.* 2015). All three sites assessed here were better fit by the lognormal than the logseries. Therefore, our results support the "veil effect" hypothesis (Preston 1948; Connolly 2005) as the most likely explanation of the observed SADs of tree communities in the Amazon basin. The "veil effect" hypothesis simply emphasizes that the underlying shape of the SAD is lognormal because the rarest species have not been sampled yet (Preston 1948). The lognormal distribution has many fewer rare species than the logseries, which has practical implications for the development of effective conservation strategies. For example, the recently estimated number of globally threatened Amazonian tree species (Ter Steege *et al.* 2015), may be reduced. Overall, our results propose that in more intensive local samplings, such as those employed in this study, many rare species in 1-ha plots could be common elsewhere.

Metacommunity patterns

In recent years, a number of studies have sought insights into metacommunity diversity and abundance patterns by analyzing pooled datasets comprised of fully identified species (named species) censused in multiple spatially separate sampling

units (Ter Steege et al. 2013; Connolly et al. 2014; Slik et al. 2015). We take the same approach here, pooling data for our three large plots to investigate diversity and abundance patterns in the metacommunity, after first establishing that patterns observed within each site are qualitatively similar whether we use named species or morphospecies (see also Pos *et al.*, 2014). Our analyses of metacommunity species-individual and Fishers alpha-area curves found that samples of large individuals show different patterns than samples of all individuals. In general, large individuals are a highly nonrandom subset of all individuals, demonstrating that the inclusion of all individuals will bring additional information in terms of diversity and species composition. Finally, our metacommunity species abundance distributions were better fit by the lognormal than by the logseries for both all individuals and just large individuals. This has consequences for the quantification of species rarity and dominance (Pitman et al. 1999; Pitman et al. 2001), including estimates of the number of hyperdominant species (*sensu* Ter Steege *et al.*, 2013). The inclusion of all individuals and larger local samples should reduce the proportion of dominant species (Figure S5).

Conclusions and future directions

The use of plots larger than 1 ha that includes smaller sizes than the usually 10 cm DBH employed will surely shed new insight on forest structure and diversity of Amazon forests. The use of large permanent plots, although limited to describe structural patterns at the landscape and regional scales, will surely help to unravel the main mechanisms that maintain and regulate forests structural dynamics and the capability of these ecosystems to respond to climate change. However, based on our findings in these three large plots in Amazonia, we recommend that the minimum census area to adequately capture local tree diversity in the Amazon is 2 ha for the ≥ 1 cm size class, or 10 ha for the ≥ 10 cm size class. Below these areas, Fisher's alpha continues to increase with increasing area. We emphasize that censuses of all individuals ≥ 1 cm capture more species and additional kinds of species relative to

those of only individuals ≥ 10 cm, and that Fisher's alpha values tend to be lower when only larger individuals are sampled. The sampling efficiency of large individuals tallied in 1-ha plots was approximately 40% relative to that observed for all individuals in the same plot, and roughly 30% relative to all species included in a 25-ha plot (Figure S6). It is clear that we still have much to learn about patterns of forest structure and tree species diversity in the Amazon. Enhanced sampling intensity, including more large plots, ≥ 2 ha each sampled to smaller size classes, is needed if we are to fill the still large voids in our knowledge of plant diversity in Amazon *terra firme* forests and tropical ecosystems more generally (Feeley 2015).

ACKNOWLEDGEMENTS

We gratefully acknowledge the contributions of the many people who assisted in collecting the tree census data. In Colombia, this work was made possible by the Parques Nacionales de Colombia, and in particular to Eliana Martínez and staff members of the Amacayacu Natural National Park. The census of Yasuni plot was financed by Pontifical Catholic University of Ecuador (PUCE, research grants of Donaciones del Impuesto a la Renta from the government of Ecuador). The Yasuni plot census was endorsed by the Ministerio de Ambiente del Ecuador through several research permits. We also thank the Center for Tropical Forest Science - Forest Global Earth Observatory (CTFS-ForestGEO) of the Smithsonian Tropical Research Institute for partial support of plot census work. This manuscript was advanced at a working group meeting funded by a grant from the US National Science Foundation (DEB-1046113). Comments from Hans ter Steege and an anonymous reviewer helped to improve the contents of this manuscript.

SUPPORTING INFORMATION

Figure S1. Species – individual curves by site.

Figure S2. Fisher's alpha – area curves by site.

584 Figure S3. Species abundance distributions of large individuals (DBH \geq 10 cm).
 585 Figure S4. Best-fit lognormal and logseries distributions.
 586 Figure S5. Rank abundance distribution curves (RADs) by size category.
 587 Figure S6. Sampling efficiency of 1-ha plots.
 588 Table S1. Species – individuals mean and 95% confidence intervals at the plot scale.
 589 Table S2. Fit of the Species abundance models (SADs) evaluated at the plot scale.
 590 Table S3. Species – individuals mean and 95% confidence intervals at the
 591 metacommunity scale.
 592 Table S4. Mean Fisher’s alpha values.

593

594 REFERENCES

595

- 596 Baker T, Phillips O, Malhi Y, et al (2004) Variation in wood density determines spatial
 597 patterns in Amazonian forest biomass. *Glob Chang Biol* 10:545–562. doi:
 598 10.1111/j.1529-8817.2003.00751.x
- 599 Chave J, Andalo C, Brown S, et al (2005) Tree allometry and improved estimation of
 600 carbon stocks and balance in tropical forests. *Oecologia* 145:87–99. doi:
 601 10.1007/s00442-005-0100-x
- 602 Chave J, Réjou-Méchain M, Búrquez A, et al (2014) Improved allometric models to
 603 estimate the aboveground biomass of tropical trees. *Glob Chang Biol* 20:3177–
 604 3190. doi: 10.1111/gcb.12629
- 605 Condit R, Foster RB, Hubbell SP, et al (1998) Condit et al MABSeries1998.pdf. In:
 606 Dallmeier F, Comiskey JA (eds) *Forest biodiversity research, monitoring and*
 607 *modeling*. UNESCO, Paris, and The Parthenon Publishing Group, New York, pp
 608 247–268
- 609 Condit R, Hubbell SP, Lafrankie J V, et al (1996) Species-Area and Species-Individual
 610 Relationships for Tropical Trees: A Comparison of Three 50-ha Plots. *J Ecol*
 611 84:549–562. doi: 10.2307/2261477
- 612 Connolly SR (2005) *Community Structure of Corals and Reef Fishes at Multiple Scales*.

613 Science (80-) 309:1363–1365. doi: 10.1126/science.1113281

614 Connolly SR, MacNeil MA, Caley MJ, et al (2014) Commonness and rarity in the marine
615 biosphere. *Proc Natl Acad Sci U S A* 111:8524–9. doi: 10.1073/pnas.1406664111

616 Duivenvoorden JF, Duque A, Cavelier J, et al (2005) Density and diversity of plants in
617 relation to soil nutrient reserves in well-drained upland forests in the north-
618 western Amazon basin. *K Danske Vidensk Selsk Biol Skr* 55:25–35.

619 Feeley K (2015) Are We Filling the Data Void? An Assessment of the Amount and
620 Extent of Plant Collection Records and Census Data Available for Tropical South
621 America. *PLoS One* 10:e0125629. doi: 10.1371/journal.pone.0125629

622 Fisher RA, Steven Corbet A, Williams CB (1943) Fisher et al_Intl.J.Pl.Sci._1997.pdf. *J*
623 *Anim Ecol* 12:42–58.

624 Gentry a H (1988a) Tree species richness of upper Amazonian forests. *Proc Natl Acad*
625 *Sci U S A* 85:156–159. doi: 10.1073/pnas.85.1.156

626 Gentry AH (1988b) Changes in Plant Community Diversity and Floristic Composition
627 on Environmental and Geographical Gradients. *Ann Missouri Bot Gard* 75:1–34.

628 Gomes ACS, Andrade A, Barreto-Silva JS, et al (2013) Local plant species delimitation
629 in a highly diverse Amazonian forest: do we all see the same species? *J Veg Sci*
630 24:70–79. doi: 10.1111/j.1654-1103.2012.01441.x

631 Hoorn C, Wesselingh FP, Steege H, et al (2010) Amazonia Through Time : Andean.
632 *Science* 330:927–931. doi: 10.1126/science.1194585

633 Hubbell SP (2001) The Unified Neutral Theory of Biodiversity and Biogeography.

634 Hubbell SP (2013) Tropical rain forest conservation and the twin challenges of
635 diversity and rarity. *Ecol Evol* n/a–n/a. doi: 10.1002/ece3.705

636 Hubbell SP, He F, Condit R, et al (2008) Colloquium paper: how many tree species are
637 there in the Amazon and how many of them will go extinct? *Proc Natl Acad Sci U S*
638 *A* 105 Suppl :11498–11504. doi: 10.1073/pnas.0801915105

639 Lips J, Duivenvoorden JF (2001) Caracterización ambiental. In: Duivenvoorden JF,
640 Balslev H, Cavelier J, et al. (eds) Evaluación de recursos naturales no maderables
641 en la Amzonía noroccidental. Universiteit van Amsterdam, Amsterdam, pp 19–46

642 Malhi Y, Wood D, Baker TR, et al (2006) The regional variation of aboveground live
643 biomass in old-growth Amazonian forests. *Glob Chang Biol* 12:1107–1138. doi:
644 10.1111/j.1365-2486.2006.01120.x

645 McGill B (2003) A test of the unified neutral theory of biodiversity. *Nature* 422:881–
646 885. doi: 10.1038/nature01569.1.

647 Myers N, Mittermeier RA, Mittermeier CG, et al (2000) Biodiversity hotspots for
648 conservation priorities. *Nature* 403:853 – 858.

649 Oksanen J, Blanchet F, Kindt R, et al (2013) vegan: Community Ecology Package. R
650 package version 2.0-10. R Packag. version 1:10.4135/9781412971874.n145.

651 Phillips O, Hall P, Gentry A, et al (1994) Dynamics and species richness of tropical rain
652 forests. *Proc. Natl. Acad. Sci.* 91:2805–2809.

653 Phillips OL, Baker TR, Arroyo L, et al (2004) Pattern and process in Amazon tree
654 turnover, 1976-2001. *Philos Trans R Soc B Biol Sci* 359:381–407. doi:
655 10.1098/rstb.2003.1438

656 Phillips OL, Malhi Y, Higuchi N, et al (1998) Changes in the Carbon Balance of Tropical
657 Forests: Evidence from Long-Term Plots. *Science* (80-) 282:439–442. doi:
658 10.1126/science.282.5388.439

659 Pitman NCA, Terborgh J, Silman MR, Nuez P (1999) Tree species distributions in an
660 upper Amazonian forest. *Ecology* 80:2651–2661.

661 Pitman NCA, Terborgh JW, Silman MR, et al (2001) Dominance and distribution of tree
662 species in upper Amazonian terra firme forests. *Ecology* 82:2101–2117.

663 Pos E, Guevara Andino JE, Sabatier D, et al (2014) Are all species necessary to reveal
664 ecologically important patterns? *Ecol Evol* 4:4626–36. doi: 10.1002/ece3.1246

665 Prado PI, Miranda MD (2013) Fitting species abundance model with maximum
666 likelihood Quick reference for sads package. 1–20.

667 Preston FW (1948) The commonness, and rarity, of species. *Ecology* 29:254–283. doi:
668 10.2307/1930989

669 R Development Core Team (2014) R: A Language and Environment for Statistical
670 Computing. R Found. Stat. Comput. Vienna Austria 0:{ISBN} 3–900051–07–0.

671 Saatchi SS, Houghton RA, Dos Santos Alvalá RC, et al (2007) Distribution of
672 aboveground live biomass in the Amazon basin. *Glob Chang Biol* 13:816–837. doi:
673 10.1111/j.1365-2486.2007.01323.x

674 Slik JWF, Arroyo-Rodríguez V, Aiba S-I, et al (2015) An estimate of the number of
675 tropical tree species. *Proc Natl Acad Sci* 112:7472–7477. doi:
676 10.1073/pnas.1423147112

677 Sombroek WG (2000) Amazon landforms and soils in relation to biological diversity.
678 Acta Amaz. 30:81–100.

679 Ter Steege H, Pitman N, Sabatier D, et al (2003) A spatial model of tree a -diversity and
680 tree density for the Amazon. Biodivers Conserv 12:2255–2277.

681 Ter Steege H, Pitman NC a, Sabatier D, et al (2013) Hyperdominance in the Amazonian
682 tree flora. Science 342:1243092. doi: 10.1126/science.1243092

683 Ter Steege H, Pitman NCA, Killeen TJ, et al (2015) Estimating the global conservation
684 status of more than 15 , 000 Amazonian tree species. Science (80-) 9–11. doi:
685 10.1126/sciadv.1500936

686 Ter Steege H, Pitman NCA, Phillips OL, et al (2006) Continental-scale patterns of
687 canopy tree composition and function across Amazonia. Nature 443:444–447.
688 doi: 10.1038/nature05134

689 Wilson JB (1991) Methods for fitting dominance / diversity curves. J Veg Sci 2:35–46.
690 doi: 10.2307/3235896

691

692

Table 1. Forest structural variables, species richness, and species diversity for the three 25-ha plots located in the Amazon basin, by size class (DBH ≥ 1 cm or DBH ≥ 10 cm). Except for the total individuals and total species values, reported statistics are the means \pm standard deviations over 25 square 1-ha (100 \times 100 m) subplots, and F statistics test for differences among sites in these 1-ha statistics. ^{ns} = non significant; * = $P \leq 0.05$; ** = $P \leq 0.01$; *** = $P \leq 0.001$. Different letters indicate significant differences according to the Tukey Honestly Significant test (Tukey's HSD test). **NI: number of individuals (ha⁻¹). BA: Basal area (m² ha⁻¹). AGB: aboveground biomass (Mg ha⁻¹). WD: wood density (g cm⁻³). Total values are indicated for the whole 25 ha of each plot or the joined 75 ha for the three plots.**

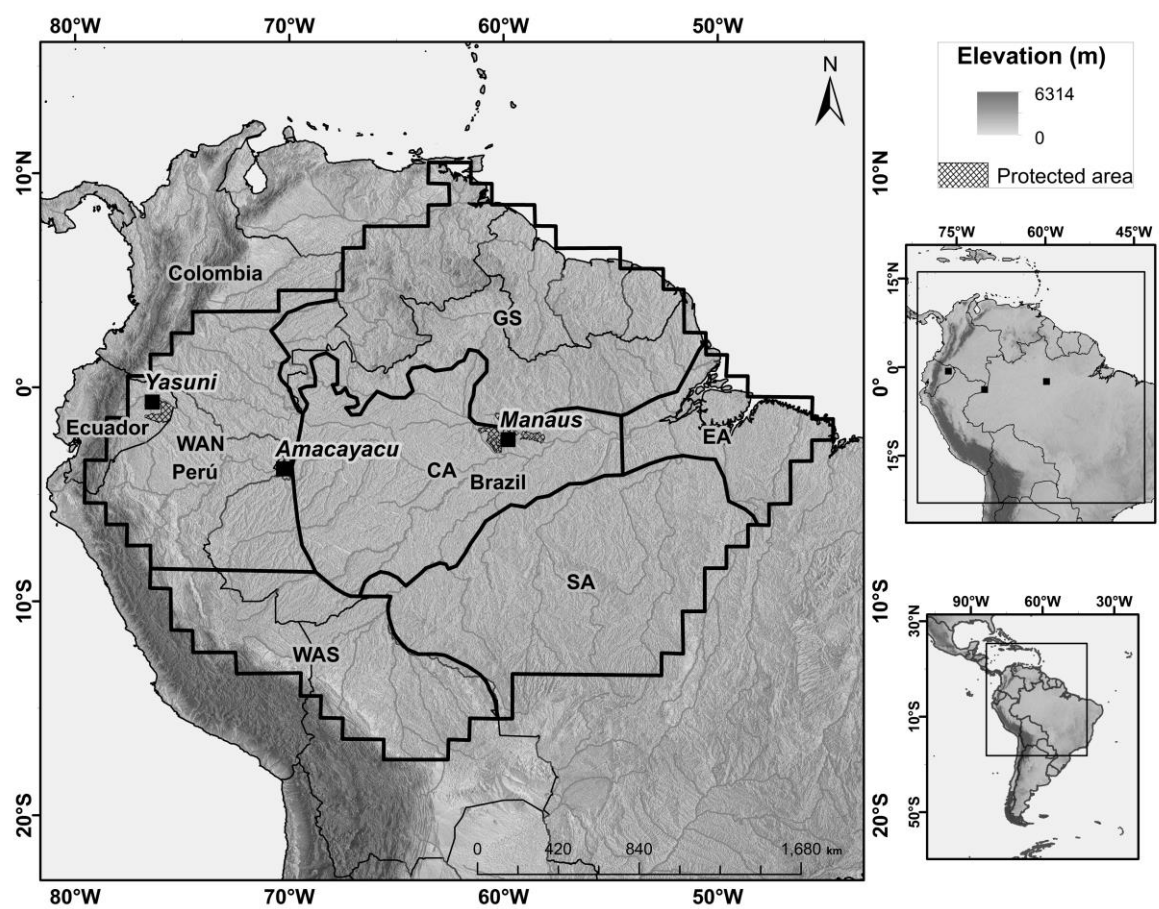
	Yasuni	Amacayacu	Manaus	All	F
All individuals					
Total NI (≥ 1 cm)	154,328	123,714	163,046	441,088	
Total NI (≥ 10 cm)	17,518	14,464	14,474	46,456	
NI (≥ 1 cm)	6173 \pm 743 a	4949 \pm 518 b	6522 \pm 579 a	5882 \pm 572	49.9***
NI (≥ 10 cm)	701 \pm 42 a	579 \pm 38 b	579 \pm 42 b	619 \pm 41	75.3***
BA (≥ 1 cm)	32.9 \pm 3.9 a	29.1 \pm 2.7 b	31.6 \pm 3.1 a	31.2 \pm 3.2	8.6***
BA (≥ 10 cm)	27.1 \pm 3.9 a	24.6 \pm 2.5 b	26.3 \pm 3.1 a,b	26.0 \pm 3.2	4.0***
AGB (≥ 1 cm)	256.2 \pm 49.2 a	262.8 \pm 33.0 a	345.3 \pm 54.1 b	288.1 \pm 45.4	28.7***
AGB (≥ 10 cm)	234.1 \pm 49.0 a	244.3 \pm 32.9 a	322.0 \pm 54.4 b	266.8 \pm 45.4	26.9***
Mean WD (≥ 1 cm)	0.60 \pm 0.16 a	0.61 \pm 0.13 b	0.66 \pm 0.13 c	0.62 \pm 0.14	245.7***
Mean WD (≥ 10 cm)	0.57 \pm 0.20 a	0.59 \pm 0.16 b	0.67 \pm 0.14 c	0.61 \pm 0.18	159.8***
Morphospecies					

Total NI (≥ 1 cm)	154,135	115,319	150,122	419,576	
Total NI (≥ 10 cm)	17,507	14,272	13,345	45,124	
Total morphospecies (≥ 1 cm)	1084	1230	1302	2993	
Total morphospecies (≥ 10 cm)	795	830	930	2095	
Morphospecies (ha^{-1}) (≥ 1 cm)	679.6 ± 33.8 a	597.5 ± 50.6 b	671.0 ± 38.0 a	649.4 ± 40.8	29.6***
Morphospecies (ha^{-1}) (≥ 10 cm)	257.2 ± 16.9 a	218.8 ± 23.3 b	241.3 ± 15.6 c	234.1 ± 18.6	25.9***
Fisher's alpha (≥ 1 cm)	195.9 ± 13.6 a	183.4 ± 17.4 b	194.1 ± 15.7 a	191.2 ± 16.4	4.7*
Fisher's alpha (≥ 10 cm)	147.4 ± 16.2 a	130.6 ± 20.2 b	171.8 ± 24.0 c	149.9 ± 26.4	25.8***
Named species					
Total NI (≥ 1 cm)	127,879	104,494	134,473	366,846	
Total NI (≥ 10 cm)	15,743	13,294	12,172	41,209	
Total named species (≥ 1 cm)	824	908	959	2068	
Total named species (≥ 10 cm)	615	642	730	1527	
Named species (ha^{-1}) (≥ 1 cm)	527.5 ± 26.3 a	484.4 ± 40.2 b	527.8 ± 28.3 a	513.2 ± 31.6	15.0***
Named species (ha^{-1}) (≥ 10 cm)	213.3 ± 13.3 a	191.2 ± 20.8 b	208.5 ± 13.1 a	204.4 ± 15.7	13.0***
Fisher's alpha (≥ 1 cm)	148.2 ± 9.5	142.2 ± 13.6	145.4 ± 12.1	145.3 ± 12.0	2.8ns
Fisher's alpha (≥ 10 cm)	114.2 ± 12.1 a	107.8 ± 16.8 a	139.8 ± 18.5 b	120.6 ± 21.1	28.1***

704

705

706

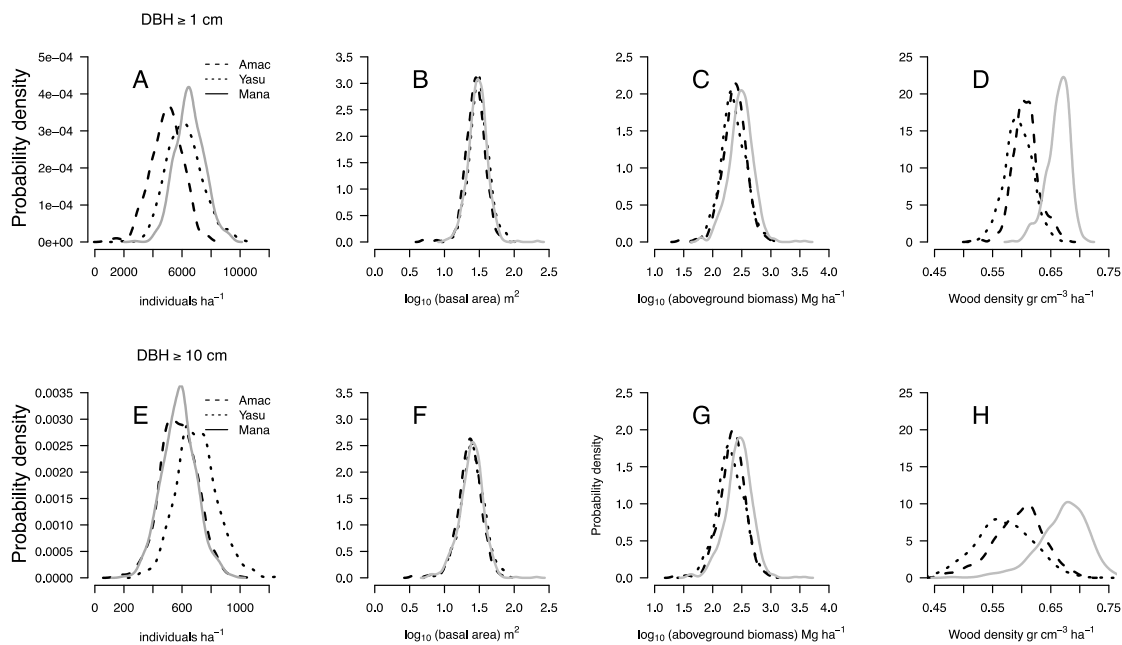


707

708

709 Figure 1. Geographical location of the three 25 ha plots employed in this study
710 (adapted from Ter Steege et al. 2013). CA: central Amazon. EA: eastern Amazon. GS:
711 Guyana shield. SA: southern Amazon. WAN: northwestern Amazon. WAS:
712 southwestern Amazon.

713

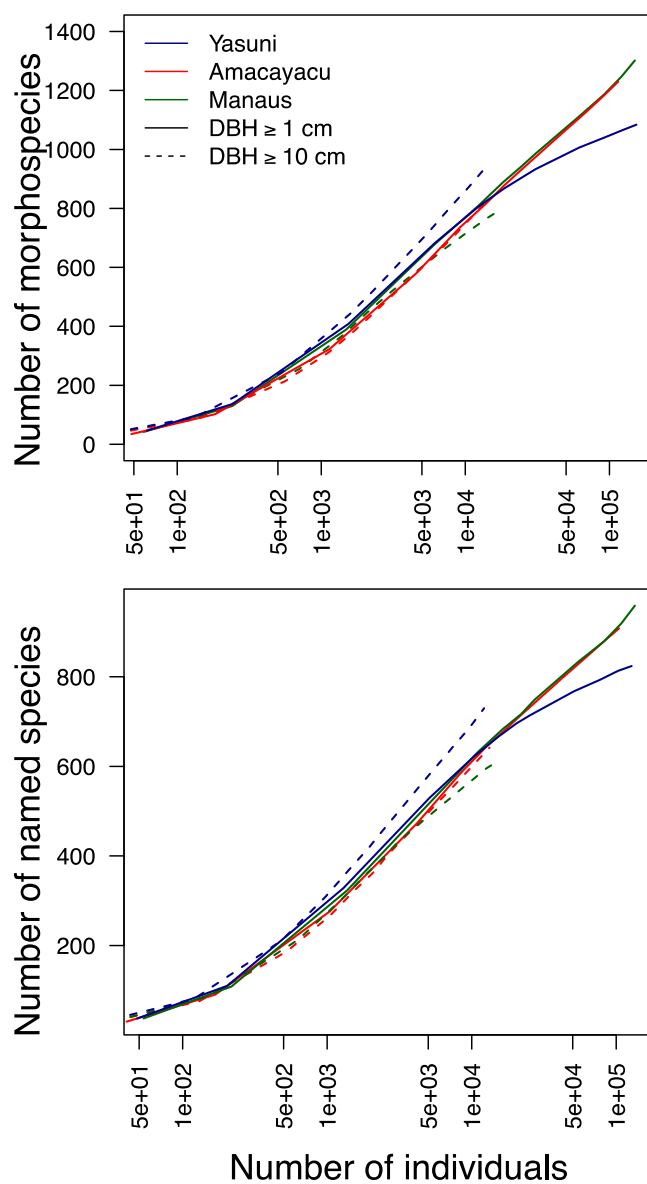


715

716

717 Figure 2. Variation in forest structure within and among sites. Lines show empirical
 718 probability density functions for each site calculated from data for all individuals ≥ 1
 719 cm DBH (panels A to D) and for large individuals ≥ 10 cm DBH (panels E to H).

720



722

723

724 Figure 3. Species-individual curves for all sites and size categories. Comparisons for
 725 morphospecies are in the upper panel. Comparisons for named species are in the
 726 lower panel. Separate analyses per site are shown in Figure S1.

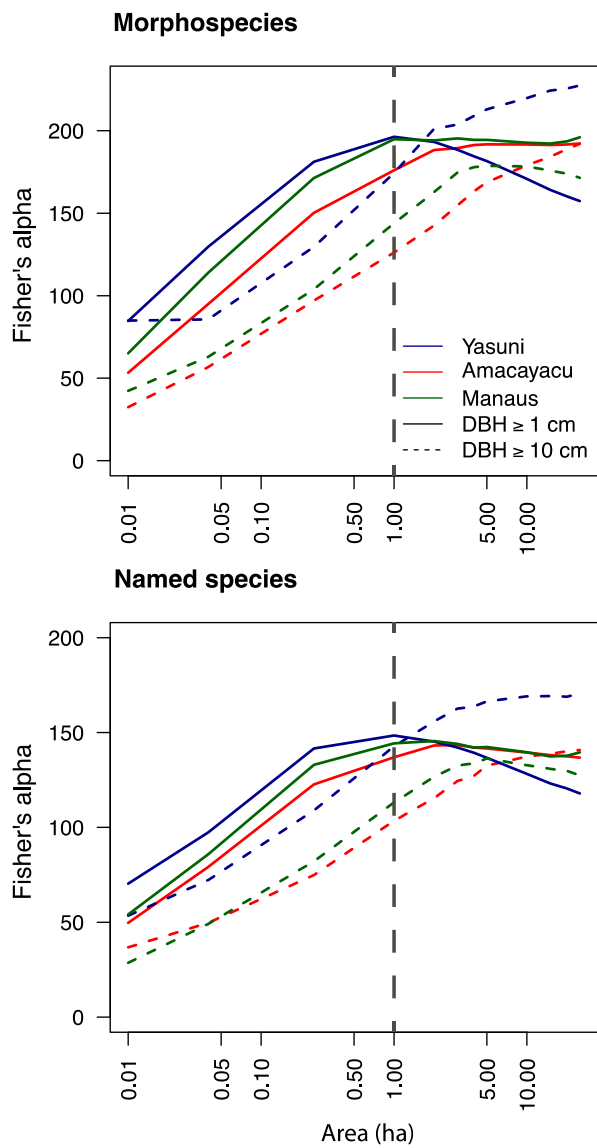


Figure 4. Fisher's alpha-area curves for all sites and size categories. Comparisons for morphospecies are in the upper panel. Comparisons for named species are in the lower panel. Separate analyses per site are shown in Figure S2.

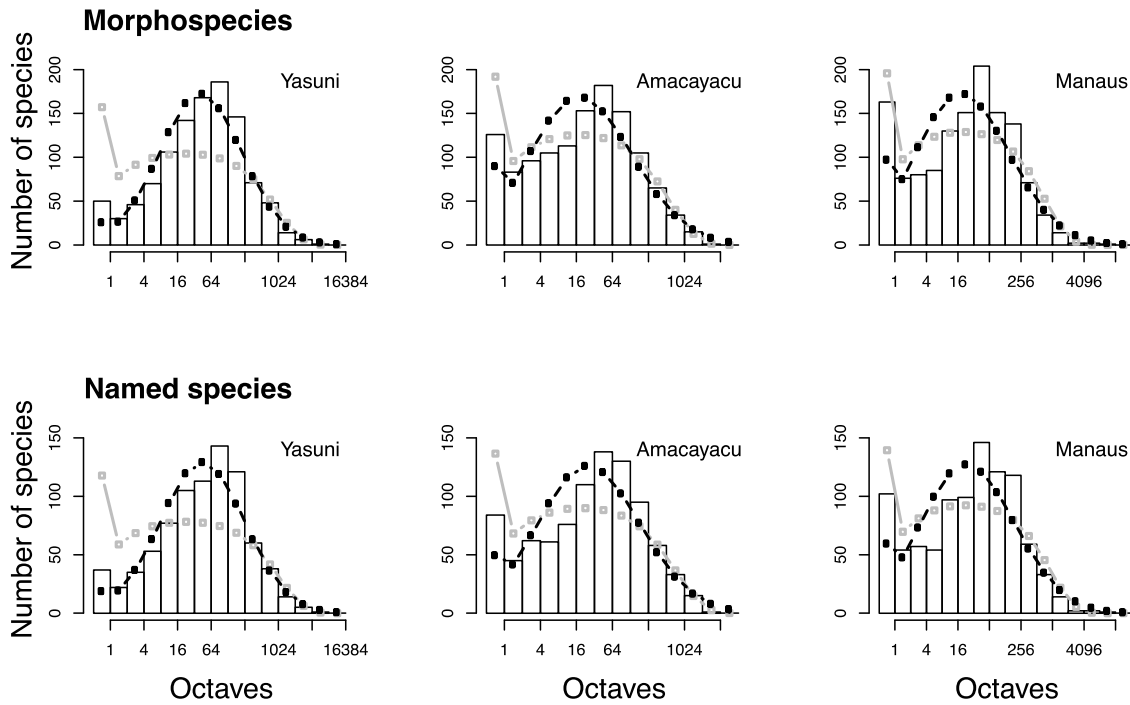
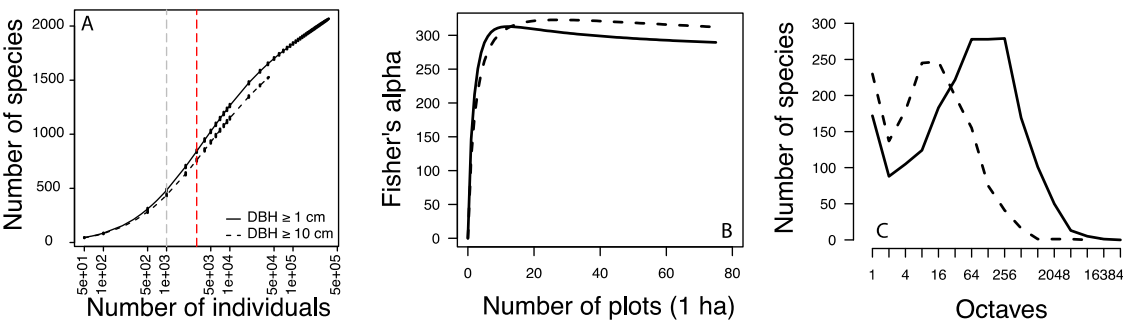


Figure 5. Species abundance distributions (bars) of all individuals ($\text{DBH} \geq 1 \text{ cm}$) for each 25-ha plot for all morphospecies (top) and just named species (bottom), along with best-fit lognormal (black) and logseries (grey) distributions. The parallel figure for large individuals only is shown in Figure S3.

740



741

742

743 Figure 6. Metacommunity patterns for species-individuals curves (A), Fisher's alpha
744 vs. area (B), and relative abundance distributions (C) based on pooling data for all
745 three 25-ha plots. Small vertical lines in the species-individual curves (panel A)
746 represent standard errors. The grey line indicates 1000 individuals and the red line
747 indicates 3000 individuals.

748

[Click here to view linked References](#)



Click here to access/download
Attachment to Manuscript
Duque et al SI.docx

