WILEY ECOLOGY

Counting niches: Abundance-by-trait patterns reveal niche partitioning in a Neotropical forest

Cluster Definition

Journal:	*Ecology*
Manuscript ID	ECY19-0381.R1
Wiley - Manuscript type:	Reports
Date Submitted by the Author:	n/a
Complete List of Authors:	D'Andrea, Rafael; University of Michigan, Ecology and Evolutionary Biology; University of Illinois at Urbana-Champaign College of Liberal Arts and Sciences, Plant Biology Guittar, John; Michigan State University Figueroa, Hector; University of Michigan, Ecology and Evolutionary Biology Condit, Richard; Smithsonian Tropical Research Institute Wright, S. Joseph; Smithsonian Tropical Research Institute, ; Smithsonian Tropical Research Institute, O'Dwyer, James; University of Illinois, Ostling, Annette; University of Michigan, Ecology and Evolutionary Biology
Substantive Area:	Niche Relationships/Resource Partitioning < Species Interactions < Community Ecology < Substantive Area, Interspecific Competition < Species Interactions < Community Ecology < Substantive Area, Species Interactions < Community Ecology < Substantive Area, Community Analysis/Structure/Stability < Community Ecology < Substantive Area
Organism:	Plants, Angiosperms < Plants
Habitat:	Rain Forest < Tropical Zone < Terrestrial < Habitat
Geographic Area:	Mainland < Central America < Geographic Area
Additional Keywords:	Trait-based clustering, Self-organized similarity, Emergent neutrality
Abstract:	Tropical forests challenge us to understand biodiversity, as numerous seemingly similar species persist on only a handful of shared resources. Recent ecological theory posits that biodiversity is sustained by a combination of species differences reducing interspecific competition and species similarities increasing time to competitive exclusion. Together, these mechanisms counterintuitively predict that competing species should cluster by traits, in contrast with traditional expectations of trait overdispersion. Here, we show for the first time that trees in a tropical forest exhibit a clustering pattern. In a 50 ha. plot on Barro Colorado Island in Panama, community trait values exhibit clusters in two traits connected to light capture strategy, suggesting that competition for light structures community composition. Notably, we find four clusters by maximum height, quantitatively supporting the classical grouping of

	Neotropical woody plants into shrubs, understory, midstory, and canopy layers.

Response to reviewers

12-Jul-2019

Dear Dr. D'Andrea:

Thank you very much for submitting your manuscript "Counting niches: Abundance-by-trait patterns reveal niche partitioning in a Neotropical forest" ECY19-0381 to Ecology. The reviewers and I appreciate the work you have accomplished. We are willing to consider a revised version for publication in the journal, assuming that you are able to modify the manuscript according to the recommendations.

…

Sincerely,

Prof. María Uriarte
Subject Matter Editor, Ecology
mu2126@columbia.edu

Thank you to the editor and reviewers for their positive responses to our work, and for their constructive comments. We have addressed all of their major concerns, and believe it has resulted in an even stronger manuscript. We now provide a more nuanced discussion of our results, a more thorough discussion of their implications, and potential caveats to our interpretations.

Our responses follow below the respective comments, in bold.

_____

Reviewer: 1
Comments to the Author

The article by D'Andrea et al investigates potential for niche structuring of tropical forests, as revealed by clustering of abundance-weighted trait distributions across species. The paper combines measures of size and abundance, taken from long term plot surveys, with trait data for four prominent physiological traits, and new methods for detecting clustering of species along the trait spectrum. The results suggest clustering is present in two of the four traits investigated: maximum height and wood density, but not seed mass or leaf mass per area. For both of the clustered traits, the statistical methods suggest 4-5 clusters; a number which corresponds with traditional groupings of species by experienced botanists into four height types. The existence of clusters is significant also, as earlier theory suggested over dispersion of trait values would be the a signature of niche differentiation. By contrast, more recent theory

Automatic Cluster Choice.

argues in favour of clustering. The authors therefore claim these results provide evidence of niche partitioning in these forests.

I like many aspects of this paper. It is clearly and concisely written, and tests recent developments in theory of niche differentiation at BCI - a location that has been pre-eminent in the debates around forces structuring forests. Notably, BCI has provided some of the strongest evidence for both neutral and niche theories. Building from a theoretical framework that partly integrates these perspectives, the current paper tests an important prediction of the new theory.

Thank you for these positive comments.

I have some concerns around the interpretation of the results, particularly of the presence or not of clusters. The authors claim that existence of clusters suggests "that competition for light structures community composition" (ln 27 & later). My concerns are two fold. First, can we or should we take clustering as evidence for or against niche partitioning? Is the evidence here stronger than that in the previous theory, where trait over dispersion was the signature of niche differentiation? Other theoretical models (e.g. Leimar et al 2013 (http://dx.doi.org/10.1016/j.jtbi.2013.08.005) show how niche differentiation can lead to a continuous distribution of traits, given a competition kernel with the right shape.

This is an important question. As the reviewer points out, classical literature predicts that niche partitioning will lead to trait overdispersion, whereas here we interpret trait clusters rather than overdispersion as evidence of niche partitioning. Our work follows up on recent literature drawing attention to the fact that the same models that predict overdispersion in stable equilibrium will predict clusters as transients (e.g. Scheffer and van Nes 2006, D'Andrea et al 2019). Following the initial findings by Scheffer and van Nes, some of us and others have shown that several ecological forces can maintain the clusters indefinitely as a stationary state. Those forces include immigration (D'Andrea et al 2019), appearance of new species through mutation (D'Andrea & Ostling 2016), periodic environments (Sakavara et al 2018), and "hidden niches" (i.e. independent niche axes providing extra intraspecific regulation, Scheffer and van Nes 2006, Barabas et al 2013). The 50-ha plot on Barro Colorado Island is very likely under the influence of one or several of these forces, such as immigration and other niche axes. Based on this theoretical literature, we therefore interpret clusters on BCI as a more general signature of niche differentiation than the stricter circumstances under which one might predict overdispersion. We added text in the Introduction to make clearer the state of theory regarding clusters versus overdispersion (lines 65-75).

As the reviewer notes, other literature has indicated that kernels of the right shape may allow for coexistence of a trait continuum (Leimar et al 2013). Specifically, the kernel must have a strictly positive Fourier transform, which is not a general property of competition based on trait differences. (For example, in the family of exponential kernels, only those with exponent

Presence of greater variability than expected.

≤ 2 will satisfy the condition, and small perturbations of these kernels will break the property and lead to cluster formation.) However, in the absence of empirical evidence that the competition kernels on BCI possess those special properties, we believe a more general expectation is for cluster formation rather than a continuum of traits. We now bring this up on lines 218-226.

Also, clustering was only present for two of the traits only - does that mean the other two traits are not involved in niche partitioning? Previous work at this site has provided reasonable evidence that seed mass (less so LMA) influences population dynamics.

Thank you for raising this valuable point, which we now emphasize in our manuscript (lines 275-287). The absence of clustering in LMA and seed mass does not preclude a role for these traits in niche partitioning on BCI. In D'Andrea et al 2018, we show that a loose connection between trait value and niche strategy can mask clustering along the relevant niche axis. Seed mass is linked to dispersal and recruitment ability (Muller-Landau et al 2008) and may reflect a species' position on a niche axis based on a competition-colonization tradeoff, and is therefore a candidate for clustering -- hence its inclusion in our study. However, seed mass alone may be a poor predictor of a species' strategy along this niche axis, due, for example, to the effects of other traits such as drought tolerance and dispersal mode. LMA is theoretically connected to niche strategies (Wright et al. 2010), but there is substantial ontogenetic variation in its predictive ability of growth, with the relationship being stronger in the seedling stage. Eventually, other aspects of physiology become limiting as the plant grows. Hence, LMA may be a noisy predictor of the niche strategy (Wright et al. 2010). We conclude that the lack of clustering in these traits may be due to their loose connection to niche strategies rather than no involvement in niche partitioning.

My other concern is about the method of cluster analysis. To detect significant clustering, the method uses abundance and compares to null communities where abundance's are shuffled across species. But the authors do not report if traits on their own show any evidence of clustering, without considering abundances. If so, how does this impact significance testing? If traits are significantly clustered on their own, I would consider this strong evidence favouring your general hypothesise. Yet, the effect may not appear significant using the methods applied, as the null communities also include this effect.

[Handwritten margin note: Physicality of traits.]

An analysis based on species trait values alone, independent from their local abundances, would require a different null model than the one we currently use in the paper. Whereas our null model randomly shuffles species abundances among the observed trait values at BCI, a null model that ignores species abundances would likely rely on random selections of species drawn from the regional pool. However, trait data for all species in the regional pool, most of which are rare, are not available, making this approach infeasible at this time. Another type of null model focused solely on species trait values could be to randomly choose trait values from the observed ranges.

However, non-random patterns compared to this null would more likely be the result of an evolutionary process (D'Andrea and Ostling 2016) rather than than niche differentiation during community assembly. Preliminary results using a small number of starting points for the k-means analysis suggest the trait values are not significantly clustered compared with this null. Due to the difficulties in generating these alternative null model communities and/or their associated issues of interpretation we did not pursue these analyses further.

It would also be good to explain further why wood density shows significant clustering but LMA and seed mass do not, as visually, I cannot see the patterns are that different.

One of the strengths of our statistically-grounded method is that it detects patterns even when they are not visually obvious, and it tells us when we are "seeing" patterns that actually do not exist.

While the discussion was admirably brief, I feel there was a missed opportunity, particularly given the background knowledge available for this study site. One of my favourite papers on the topic of clusters is Barabas et al 2013 "Emergent neutrality or hidden niches?", on which D'Andrea & Ostling are co-authors. That article points out that Scheffer & Nes 2006's model, which kicked off the current theory lineage on clustering of traits, involved a strong assumption required to induce the observed clustering. That is it requires differentiation along an additional niche axis. One possible source for such differentiation is in the interaction with pathogens and herbivores. BCI has a long history of research in this area, and it would be a shame not to consider this in relation to the assumptions of the model.

Thank you for this insightful question and suggestion. As the reviewer notes, Barabas et al 2013 pointed out that additional niche axes, such as investment in defense against herbivores and pathogens, are required to generate permanent clusters in a closed community subject to no temporal variation in the environment. As the reviewer points out, defense is a well-known important niche axis in tropical forests, and as such is likely to contribute to clusters along the trait axes we report in our study, along with immigration (in fact, we believe it likely that our species also cluster by traits related to defense, and are planning a follow-up study along those lines). We now emphasize this point in our revised Discussion (lines 241-253).

It is also essential to discuss the results further in relation to successional dynamics. Height appears as a clustered trait, but surely there is a big difference between a shade tolerant shrub and a short early successional pioneer?

Thank you for this insightful comment. The cluster of species at the low end of the max-height trait axis on BCI may indeed comprise a mix of shade-tolerant shrubs and short, early-successional pioneers. These species may have similar light acquisition strategies (the niche axis represented by maximum height) but be niche-differentiated in terms of

their successional strategy, which is a niche axis that would be better reflected on other traits such as wood density rather than maximum height.

To reiterate the point above, the four maximum height strategies reflect coexistence that could happen without disturbance and successional dynamics, perhaps due to a tradeoff between investment in light acquisition versus investment in recruitment (Kohyama 1993). On the other hand, the coexistence between species with similar height strategies (i.e. those within the same height cluster) requires differentiation along additional niche axes, such as successional dynamics -- which would be reflected in the distribution of wood density values and/or other traits related to successional strategy. We now make this point in the revised Discussion (lines 241-253).

Minor issues

Method of clustering identifies clusters that "maximizes the difference between within-cluster trait dispersion in the observed and null communities. " (ln 130). This is not straight forward to visualise, hence i recommend a figure to illustrate the concept.

What we mean by this language is that for a given number of clusters, we first (i) run the k-means algorithm to find the best assignment of species to clusters on BCI, which is the assignment that minimizes the average distance between species in each cluster; then (ii) repeat this procedure on each of a set of null communities, and take the difference between the BCI value and the mean value of the null communities. This difference we call the gap index for this number of clusters; (iii) we calculate the gap index for a range of numbers of clusters, and select the one that gives us the biggest value. That biggest value is the gap statistic, and the number of clusters that gives us that maximum is the estimate for the number of clusters on BCI.

This metric is explained in further detail in D'Andrea, Riolo & Ostling PLOS Comp Bio 2019. Fig 1 in S1 Box in the supporting information of that paper illustrates the concept.

Are the clusters necessarily equally spaced? In Fig2a, for height, it is clear there is a strong break around 27m, but below this height the distribution looks more even than clustered. In other words, two clusters: one for talon trees, then the rest.

There is no theoretical reason why the clusters should necessarily be equally spaced, nor is this a requirement for our detection of clustering using our statistic (we now note this on lines 154-155). Small between-species differences at one section of the trait axis may be ecologically similar to large between-species differences at another section of the trait axis. Thus, we do not necessarily expect even distribution of clusters.

As for the possibility of there being only two clusters in maximum height, our statistical test reveals that the biggest difference in "clusteredness" (formally, the total abundance-

weighted within-cluster dispersion) between the BCI community and a set of null communities occurs for four clusters. This means that upon testing for different numbers of clusters, four is the number that gives the strongest signal relative to null expectations. However, this finding does not rule out the possibility of layered height structure, with a larger split into tall trees and the rest subsuming substructure made up of three clusters in the bulk. That said, we think the best approach is to defer to the statistical results provided by our test, rather than visual intuition.

Ln 72 - for what taxa?

Our methodology for detecting clusters developed in D'Andrea, Riolo, & Ostling (2019) is not taxon-specific. The method is shown to detect clusters that arise in a variety of "stochastic niche models", i.e. models that include niche differentiation along with stochastic birth, death, and immigration. This includes the MacArthur-Levins (1967) niche axis model, which is not mechanistic, as well as mechanistic models where niche differentiation is specifically deriving from differences in use of resources (whose dynamics are included in the model), differences in habitat use, and differences along a competition-colonization tradeoff. These mechanisms could be at work for a wide variety of taxa and our models are not constructed in a taxon-specific manner. We considered them in order to demonstrate that niche differentiation produces clustering independently from the details of the dynamics of competitive interactions.

We did not make changes to the text of our manuscript regarding this point as it takes too much space to give a detailed sense of the models used to validate the metric in our other paper, and the reader can refer to it for them.

———————————————————————————————————————

Reviewer: 2
Comments to the Author

Dear editor and authors,

D'Andrea et al. assess to what extent the distribution of traits show multiple clusters (i.e. different optimum values) across tropical tree species in BCI. They show that maximum tree height clusters around four optima, wood density around five optima, and leaf mass area and seed mass show no clusters. At the regional scale, only maximum height shows clusters, although much less clearly than for BCI alone.

This is an interesting and well-written manuscript. Nevertheless, some aspects are unclear, and other aspects are rather poorly explained. This limits the ability to see these results in a more general context, rather than a specific site study. Below I provide some suggestions for improvement.

Thank you for the positive comments about our manuscript. We agree that our manuscript could benefit from additional discussion of the broader implications of our results, and we include several points in the new version of our manuscript (see below).

- The authors mention that it is unknown whether clustering by size occurs in tropical forests (L65-66), but this seems not true: we know that there are different strata in tropical forests, and the authors themselves provide some of the relevant references for this (e.g. L160-161).

It is true that conventional descriptions of tropical forest structure have long included references to different strata (sub-canopy, canopy, etc.), but to our knowledge there has never been any quantitative empirical support for their existence. Indeed, we were surprised and excited to see that the number of clusters based on tree maximum height conformed to these conventional descriptions, and we think this actually makes our study more compelling. We now highlight this on lines 208-210.

- It is unclear to me why only reproductive individuals were included. Their hypotheses underlying the formation of trait clusters should hold for all established individuals, not only for the reproductive individuals. Perhaps the formation of clusters becomes stronger during ontogeny (species with non-optimal traits eventually drop out), which would by itself be interesting to test, and may strengthen the authors' conclusions.

Current models predicting clusters do not include age structure and essentially only track reproductive individuals. Therefore, their predictions pertain to adults. If it could be assumed that the number of non-reproductive individuals is proportional to the number of adults, then our census restriction is indeed unnecessary. However, this is probably too strong an assumption for trees on BCI. Further work should confirm that age-structured models predict clusters for all individuals.

We did not add further analysis with non-reproductive individuals to our manuscript because these analyses are computationally demanding and would have taken us longer than the allowed time frame for revision. We see this as an important question to explore in future work.

- At regional scales, only one peak for maximum height is found, and the authors say that this may "arise through physiological constraints against extreme height". This, however, does not explain why there would be no clusters at smaller maximum height.

The reviewer is correct that constraints against extreme height do not explain why there would be no clusters at smaller maximum height. In fact, while our statistic identifies a single cluster, it also reveals four clusters subsumed within this larger cluster (see lines 190-193 and caption to Table 1).

The single regional peak reflects a unimodal abundance trend along the maximum height axis, with trees in the second height cluster being consistently the most abundant in

their communities. This abundance trend may result from an interplay of higher light incidence at larger heights and the potential for a larger number of smaller-sized trees in a given area. We now point this out on lines 193-198.

-       The results for height are intuitive and agree with the canopy strata that we are familiar with. However, the results for wood density are not clear and not well explained. Why would there be clusters around .35, .5, .58, .65 and .8? What is so special with these values that species tend to cluster around them? Also, why is this not visible at regional scale? Does that mean that immigration might be weaker (and therefore clusters more apparent) in BCI compared to all other plots? And if so, then perhaps we would not expect strong clusters in larger continuous stretches of forest? It would be good if the authors could elaborate on this.

Thanks for bringing this up. The positions of the clusters generally have many causes: the available pool of traits, external filtering from the environment, resource availability, and the shape of the competition kernel (the function specifying how species compete given their trait values). In the case of maximum height, it is possible that the optimal heights are determined by the distribution of light below the canopy (the sunfleck model, lines 228-232). In the case of wood density, given its connection to the successional niche, the distribution of light gaps in the forest may possibly determine the positions of the niches. The number of clusters may also be influenced by the rate at which competition drops between species of disparate wood density: the faster the drop, the more niches are allowed. We now mention this in the Discussion (lines 260-264).

As for the regional pool, we hypothesize that differences in environmental or ecological conditions across the region may have led to different trait optima (cluster centroid values) in different communities, which led to the erosion of cluster structure when these communities are combined (lines 199-202).

-       No explanation is given as to why no clustering is found for leaf mass area and seed mass. It is important to understand not only the 'significant' results, but especially also the non-significant ones. What does this mean?

We hope our response to a similar question raised by the other reviewer (page 3 of this letter, lines 275-287on the manuscript) satisfactorily answers this question.

A few minor comments:
-       L44-45: this sentence is unclear. It seems to me that you have just specified testable predictions?

By this phrase we meant the lack of model-validated quantitative predictions for trait-based impacts of niche-based competition beyond intuitive ideas of greater-than-chance trait differences. We rephrased for clarity (lines 44-45).

-       L52-53: it is not clear why a potential invader will have higher fitness when it is similar to one of the residents.

This is borne out of the calculation of the total competition faced by the invader from both residents. Essentially, the increase in competition with the closer resident is more than compensated by an even stronger decrease in competition with the other resident. We now mention this on lines 54-55.

-       L107: usually height (also maximum height) is a highly skewed variable that would be good to log-transform.

In our study we use maximum potential species height, not individual height. We tested all traits for a skewed distribution and found that only LMA and seed mass were significantly skewed, hence our choice to log-transform these two traits but not maximum height or wood density.

-       L118: what happened with the other two plots?

In the remaining two plots, all individuals with maximum DBH > 10cm were censused in the entire plot. We now mention this on lines 135-136.

-       L173: may K=1 not just refer to environmental filtering?

In the case of height, our intuition is that the environment would not filter for one single intermediate height. One would expect filtering for the maximum biologically possible height, or according to the sunfleck model, for different height layers where the incidence of light has local maxima. Our most likely explanation for the presence of a single large cluster in the regional pool is an interplay of higher light incidence at larger heights and the potential for a larger number of smaller-sized trees in a given area (lines 196-198).

1   **Counting niches: Abundance-by-trait patterns reveal niche partitioning in a Neotropical**

2   **forest**

3   Rafael D'Andrea* [1,2,†], John Guittar[1,3,‡], Hector Figueroa[1,#], Richard Condit[4,&], S. J. Wright[4,$],

4   James P. O'Dwyer[2,§] & Annette Ostling[1,¶]

5

6   [1]Dept. Ecology & Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

7   [2]Dept. Plant Biology, University of Illinois, Urbana-Champaign, IL, USA

8   [3]Dept. Plant Biology, Michigan State University, Hickory Corners, MI, USA

9   [4]Smithsonian Tropical Research Institute, Apartado 0843–03092, Balboa, Republic of Panama

10

11 [†]rdandrea@illinois.edu ORCID 0000-0001-9687-6178 (corresponding author),

12 [‡]guittarj@msu.edu, [#]hecfox@umich.edu, [&]conditr@gmai.com, [$]wrightj@si.edu ORCID 0000-

13 0003-4260-5676, [§]jodwyer@illinois.edu, [¶]aostling@umich.edu

14

19   **Abstract**

20   Tropical forests challenge us to understand biodiversity, as numerous seemingly similar species

21   persist on only a handful of shared resources. Recent ecological theory posits that biodiversity is

22   sustained by a combination of species differences reducing interspecific competition and species

23   similarities increasing time to competitive exclusion. Together, these mechanisms

24   counterintuitively predict that competing species should cluster by traits, in contrast with

25   traditional expectations of trait overdispersion. Here, we show for the first time that trees in a

26   tropical forest exhibit a clustering pattern. In a 50 ha. plot on Barro Colorado Island in Panama,

27   community trait values exhibit clusters in two traits connected to light capture strategy,

28   suggesting that competition for light structures community composition. Notably, we find four

29   clusters by maximum height, quantitatively supporting the classical grouping of Neotropical

30   woody plants into shrubs, understory, midstory, and canopy layers.

31

32   **Keywords.** niche differentiation, trait-based clustering, self-organized similarity, emergent

33   neutrality, competition, tropical forests, Barro Colorado Island, community structure.

34

35   **Introduction**

36   A basic principle of community ecology is that coexistence requires niche differentiation, i.e.

37   species differences that stabilize communities by reducing interspecific competition. However,

38   the high diversity and seemingly continuous phenotypic variation of trees in tropical forests

39   raises questions about the importance of niche differentiation as a contributor to species co-

40   occurrence and patterns of diversity, relative to other community assembly processes such as

41   dispersal and ecological drift (Hubbell 2001, Levine and Murrell 2003, Mouquet and Loreau

42   2003, Chase and Myers 2011). Indeed, neutral models reflecting chance events provide a good fit

43   to the distribution of species abundances in tropical forests (Hubbell 2001, Volkov et al. 2003),

44   although other studies have found evidence of niche differentiation (Kraft and Ackerly 2010,

45   Götzenberger et al. 2012). One likely reason for these mixed results is a lack of model-validated

46   quantitative predictions for how niche-based competition shapes community trait structure.

47   Classical theory posits that only species whose trait differences are large enough to

48   sufficiently lower interspecific competition will stably coexist. Hence, attempts to find evidence

49   of niche differentiation in nature typically look for greater-than-chance differences in trait values

50   among species (Chase and Leibold 2003, Kraft et al. 2008, Lake and Ostling 2009, Kraft and

51   Ackerly 2010, Götzenberger et al. 2012, D'Andrea and Ostling 2016). However, competition can

52   produce counterintuitive patterns. If two resident species are not distant enough from each other

53   in trait space to allow a third species with an intermediate trait value to invade, the potential

54   invader has higher fitness if it is similar to one of the residents than if it is maximally different

55   from both (MacArthur and Levins 1967). This is because the increase in competition with the

56   close resident is more than compensated by the decrease in competition with the other resident.

57   When this phenomenon is extrapolated to multispecies communities, species are competitively

58   excluded at rates inversely related to trait distance from the nearest trait optimum (D'Andrea and

59   Ostling 2017). As competition among species proceeds, a transient pattern arises characterized

60   by distinct clusters of species with similar traits, centered around locally optimal ecological

61   strategies (Scheffer and van Nes 2006). While competition among species in each cluster

62   eventually leads to exclusion of all but the most competitive species, the similar strategies among

63   the species renders their dynamics near-neutral; indeed, clusters can persist indefinitely under

64   restorative forces such as immigration (D'Andrea et al. 2019), environmental fluctuations

65   (Sakavara et al. 2018), and specialist enemies (Scheffer and van Nes 2006, Barabás et al. 2013).

66   In other words, while classical literature predicts that niche partitioning will lead to trait

67   overdispersion, more recent literature suggests trait clusters rather than overdispersion as

68   evidence of niche partitioning. We note that the same models that predict overdispersion in

69   stable equilibrium will predict clusters as transients (Scheffer and van Nes 2006, D'Andrea et al.

70   2019). Furthermore, it has been shown that several ecological forces can maintain the clusters

71   indefinitely as a stationary state. Those forces include immigration (D'Andrea et al. 2019),

72   periodic environments (Sakavara et al. 2018), and "hidden niches", i.e. independent niche axes

73   providing extra intraspecific regulation such as specialist enemies (Scheffer and van Nes 2006,

74   Barabás et al. 2013). Given the widespread influence of these forces in nature, we expect clusters

75   to be a more general signature of niche differentiation than the stricter circumstances under

76   which overdispersion might arise.

77   Clustering by organismal size has been reported in animal (Scheffer and van Nes 2006)

78   and phytoplankton communities (Sakavara et al. 2018), but is not known to occur in tropical

79   forests – a focal point for the development and tests of coexistence theory. Furthermore, previous

80   efforts to detect species clusters have typically ignored species abundances and/or relied on

81   arbitrary binning of trait space (Scheffer and van Nes 2006, Yan et al. 2012) or temporal data to

82   distinguish between occasional and permanent species (Vergnon et al. 2009). However, recent

83   work has proposed a parameter-free method which uses trait and abundance data to identify and

84   count the clusters, which was validated using population dynamic models of competition with

85   immigration (D'Andrea et al. 2019).

86   Here we use this approach to test for trait clustering in the 50-ha Forest Dynamics Plot on

87   Barro Colorado Island, Panama (Condit 1998, Hubbell et al. 1999, Condit et al. 2012a) (BCI).

88   We quantify species strategy using four morphological traits: seed mass, maximum tree height,

89   wood density, and leaf mass per area (LMA). Seed mass is thought to reflect strategies along a

90   tolerance-fecundity tradeoff axis (Muller-Landau 2010), whereby large-seeded species specialize

91   in surviving stressful environmental conditions at the cost of low fecundity and vice-versa.

92   Maximum tree height, wood density, and LMA are associated with competition for light

93   (Kunstler et al. 2016), and as such may reflect plant strategies related to partitioning light

94   availability gradients both vertically (i.e. among forest canopy layers) and horizontally (i.e.

95   across forest patches at different stages of succession). Such partitioning of light access is central

96   to the forest architecture hypothesis (Kohyama 1993) and the related successional niche

97   hypothesis (Grime 1979, Tilman 1988).

98   Trait pattern indicative of a niche mechanism will likely only become evident at a spatial

99   scale above that of any spatial heterogeneity involved in the mechanisms (Kraft and Ackerly

100   2010, D'Andrea and Ostling 2016). Therefore, we look for signatures of niche partitioning at the

101   whole plot (50 ha) scale, large enough to encompass the spatial heterogeneity in patch-age since

102   disturbance involved in these niche mechanisms (Hubbell et al. 1999). Because immigration can

103   reinforce local clustering if surrounding forests have species clustered around the same trait-

104   based niche optima, or mask local clustering if surrounding forests have species clustered around

105   different trait-based niche optima, we compare results at the plot-scale with results from a

106   combined pool of 40 census plots within 30km of BCI (Condit et al. 2013, 2016) (Fig. 1).

107

Methods

109 The Barro Colorado Island Forest Dynamics Plot is a 1,000-by-500 meter tract of tropical forest

110 on Barro Colorado Island in Panama (Hubbell et al. 1999, Condit et al. 2012b) (Fig. 1). The plot

111 has been censused for stems above 1 cm diameter at breast height (DBH) in 1982, 1985 and

112 every five years through 2015. The plot contains approximately 210,000 stems with DBH ≥ 1

113 cm, and about 300 tree species. Results shown in the main text pertain to the 2000 census;

114 however, results are consistent across all available censuses (Appendix S1: Table S1, Figs. S3-

115 S6).

116   We used four traits related to plant architecture, leaf structure, and seed size. Trait data,

117 drawn from Wright et al. (2010), were based on the following protocols: 1) Maximum height

118 was estimated as the mean height of the six largest (by DBH) individuals, measured using a

119 rangefinder. 2) Wood density (g/cm^3) was measured after drying at 60°C. 3) Leaf mass per area

120 (g/m^2) was measured on leaf laminae, excluding petioles and petiolules, averaged across six

121 individuals. For species that reach the canopy we used leaves exposed to direct sunlight --- i.e.

122 sun leaves. For all others we used shade leaves collected from the crowns of the tallest

123 individuals in the 50 ha plot. 4) Seed mass (g) refers to endosperm plus embryo only and was

124 measured after drying at 60°C. Traits with skewed distributions (LMA and seed mass) were log-

125 transformed to reduce the skew. Palms were excluded from wood and leaf trait analyses, due to

126 their substantial differences in physiology from dicots.

127   We looked for clustering among reproductive individuals (adults), since population

128 dynamic models producing emergent clusters do not consider population structure. We consider

129 individuals to be reproductive when their DBH exceeds half the maximum DBH observed for its

130   species, a relationship that explains over 80% of interspecific variation in reproductive size

131   thresholds (see Visser et al. 2016: Fig. S9). Maximum DBH was estimated as the average of the

132   six largest individuals in the BCI plot in 2005, and an additional 150 ha of mapped tree plots

133   located within 30 km and mostly within 10 km of BCI (Visser et al. 2016).

134   The regional pool consists of 40 1-ha sampling sites within 30km of BCI (Condit et al.

135   2013, 2016). In 38 of these sites, individuals with maximum DBH > 1cm were censused in 40m-

136   by-40m internal subplots, while individuals with DBH > 10cm were censused in the entire plot.

137   In the remaining two plots, all individuals with maximum DBH > 10cm were censused in the

138   entire plot. Hence, we standardized the counts by the respective sampling area to estimate

139   densities in these sites. Of the circa 850 species contained in these sites, trait data were available

140   for those also found on BCI (see Table 1, caption to Fig. 1). While lack of pattern among this

141   subset of species in the regional pool does not rule out clustering among all species, it shows that

142   the pattern on BCI is not simply mirroring regional-scale pattern inherited via immigration.

143   Our clustering metric has two foundations: the k-means clustering algorithm (MacQueen

144   1967), an optimization procedure which assigns species to clusters by minimizing the average

145   trait distance between individuals in the same cluster; and the gap method (Tibshirani et al.

146   2001), which compares the observed data against null distributions, and selects the number of

147   clusters that maximizes the difference between within-cluster trait dispersion in the observed and

148   null communities. Specifically, given a candidate number of clusters, k, the k-means algorithm

149   finds the arrangement of species into k clusters that minimizes within-cluster trait dispersion: D_k

150    $= \frac{1}{k}\Sigma_C\Sigma_{i,j \in C}n_in_jd_{ij}^2$, where C refers to a cluster ($1 \le C \le k$), n_i is the abundance of species i,

151   and d_{ij} is the trait distance between species i and j. The algorithm starts with randomly chosen

152   trait values in the local community as possible cluster centers, then puts species into the cluster

153   whose center is the closest to them, then recalculates cluster centers, and so on until the

154   algorithm converges or changes in D_k fall below a specified threshold. Because the result can

155   depend on the starting set of cluster centers, the procedure is repeated with different starting sets,

156   and the final cluster arrangement with the lowest D_k across all sets is selected (note that the

157   clusters need not be equally spaced). We use enough starting sets that larger numbers do not

158   lower within-cluster dispersion further (typically between 1,000 and 10,000). We then apply this

159   k-means algorithm to 1,000 null communities, and the difference in within-cluster trait

160   dispersion D_k between the observed community and the mean null value is the gap index for k

161   clusters, $G_k = D_k - \overline{D}_{k,\text{null}}$. Finally, we find the number of clusters that maximizes the gap index.

162   This maximal value -- corresponding to the peak in the gap curves in Figs. 2, 3 -- is the gap

163   statistic, $G = \max(G_k)$, and the value of k at which it occurs is the estimated number of clusters

164   K.

165   We obtain a null distribution of gap statistics by performing the above routine on each of

166   the null communities, and from this distribution we extract significance (p-value) and

167   standardized effect size (z-score). The z-score is $Z = (G - \mu)/\sigma$, where $\mu = \frac{1}{1000}\Sigma_{\text{nulls}} G_{\text{null}}$ is the

168   mean of the null gap statistics and $\sigma^2 = \frac{1}{1000}\Sigma_{\text{nulls}}(G_{\text{null}} - \mu)^2$ is the variance. The p-value is the

169   proportion of null communities with a higher gap statistic than the observed community, $P =$

170    $\frac{1}{1000}\Sigma_{\text{nulls}} I(G_{\text{null}} > G)$, where the indicator function I is 1 if its argument is true, and zero

171   otherwise. Our null communities contain the same set of observed trait values and abundances,

172   with abundances randomized across the traits. We therefore test for a nonrandom association

173   between traits and abundances, as opposed to a nonrandom set of traits or abundances.

174

175   **Results**

176   We found significant clustering of species on BCI based on species maximum height (Fig. 2A;

177   Fig. 2E). There were four height-based clusters, falling at approximately eight-meter intervals.

178   Intriguingly, the pattern aligns well with the historical division of humid Neotropical forests into

179   four strata of about ten meters each: shrubs, understory, midstory, and canopy (Paul W Richards

180   1952, Terborgh 1985, Condit et al. 1995) (Fig. 3, Appendix S1: Fig S1). We also found a

181   significant pattern of five clusters based on wood density on BCI (Fig. 2F). Notably, species on

182   BCI were also significantly clustered around a single wood density optimum. This occurred

183   because species with intermediate wood densities were more common than those with extremely

184   low or high wood densities (Fig. 2B), potentially indicating physiological costs or environmental

185   filters against extreme trait values. We found no evidence for clustering based on LMA or seed

186   mass.

187   The similar number of clusters by maximum height and wood density led us to ask if the

188   two trait-based patterns reflected the same underlying niche axis. Even though maximum height

189   and wood density did not correlate significantly (Pearson's $\rho = -0.10$, p-value $= 0.06$),

190   maximum height clusters correlated with wood density clusters (χ^2-test p-value $= 0.02$,

191   Cramér's V $= 0.20$), supporting this tentative hypothesis.

192   We found significant clustering by maximum height at the regional scale, with the

193   highest gap index occurring at $K = 1$ cluster, and a secondary but still significant peak at $K = 4$

194   clusters (Appendix S1: Fig. S2A, B). The existence of four height-based clusters at both local

195   and regional scales suggests a consistent organization of species into the same four niches across

196   many tropical forest communities. However, a key difference between local and regional height-

197   based patterns is the presence of a significant single cluster at the regional scale. This single

198   regional peak reflects a unimodal abundance trend along the maximum height axis, with trees in

199   the second height cluster being consistently the most abundant in their communities. This

200   abundance trend may result from an interplay of higher light incidence at larger heights and the

201   potential for a larger number of smaller-sized trees in a given area.

202   As for lack of clustering by wood density in the regional pool, we hypothesize that

203   differences in environmental or ecological conditions across the region may have created

204   different trait optima (cluster centroid values) in different communities, leading to the erosion of

205   cluster structure when the communities are combined.

206

Discussion

208   Our results underscore competitive partitioning of light as a major driver of tropical

209   forest community structure. Specifically, we found cluster-based evidence for niche partitioning

210   by maximum height, a trait associated with light capture, and wood density, a trait associated

211   with shade tolerance (Wright et al. 2010). Conventional descriptions of tropical forest structure

212   have long included references to different strata (sub-canopy, canopy, etc.), but to our knowledge

213   there has never been any quantitative empirical support for their existence.

214   Testing for trait-based clustering is an appealing approach for the detection of niche

215   partitioning because it conceptually unites the notions of stable species coexistence via

216   differences and transient species coexistence via similarity (Holt 2006). Perhaps

217   counterintuitively, it is precisely because species with similar strategies compete more strongly,

218   and experience similar competitive effects from the rest of the community, that clusters emerge

219   (D'Andrea et al. 2018). In other words, clusters reflect the simultaneous advantages of differing

220   from others while being similar to those favored strategies.

221   We note there are special cases where competition could allow for coexistence of species

222   with any trait value (Leimar et al. 2013). Specifically, the function specifying the relationship

223   between competition and species traits (often referred to as the competition kernel) must have a

224   strictly positive Fourier transform, which is not a general property of competition based on trait

225   differences. For example, in the family of exponential kernels, only those with exponent ≤ 2 will

226   satisfy the condition, and small perturbations of these kernels will break the property and lead to

227   cluster formation. However, in the absence of empirical evidence that the competition kernels on

228   BCI possess those special properties, a more general expectation is for cluster formation rather

229   than a continuum of traits.

230   The mechanistic origin of clustering may involve environmental selection favoring

231   certain traits over others. For example, according to Terborgh's sunfleck model (Terborgh 1985),

232   light availability below the canopy may peak at various vertical strata due to intersection of light

233   cones originating from gaps in the canopy. Latitudinal parameters which affect the angle of

234   incidence, combined with the shape of tree canopies in tropical forests, predict three such strata

235   below the canopy, for a total of four optimal tree heights based on light availability.

236   Alternatively, the height-based clusters we identified may arise through tradeoffs between the

237   light-related advantages of being tall and shading others, and the budgetary constraints that

238   investment in height impinges on reproductive growth, such as through losses in recruitment or

239   shade tolerance (Kohyama 1993, Kohyama and Takada 2009). These tradeoffs may result in an

240   effective relationship between interspecific competition and height similarity. In such a case, the

241   number of clusters would relate to how quickly competition drops with increasing separation in

242   maximum height. Regardless of the mechanism, the height-based clusters reflect the existence of

243   multiple light-related niches, as well as stabilization via light-related niche partitioning.

244   The existence of four clusters by height suggests that only four species can coexist on

245   BCI based on differences in height alone, but it does not mean that only four species can stably

246   coexist on BCI. Additional niche axes may support higher coexistence, and in fact may be

247   responsible for the maintenance of the observed clusters. Barabás et al. (2013) pointed out that

248   additional niche axes, such as susceptibility to different herbivores and pathogens, are required

249   for the permanence of clusters in a closed community subject to no temporal variation in the

250   environment. Though recent empirical studies report a good deal of overlap in enemy host use,

251   models suggest that small differences in susceptibility to enemies between tree species may still

252   contribute significantly to diversity maintenance (Sedio and Ostling 2013). Hence, they likely

253   also contribute to the permanence of clusters along the trait axes we report in our study. Another

254   example additional niche axis is the successional niche (Tilman 1988): trees sharing the same

255   height niche may niche-differentiate by shade tolerance. This could be reflected in the

256   distribution of wood density values and/or other traits related to successional strategy.

257   Wood density in Barro Colorado Island tree species has been found to be a good

258   predictor of species niche strategies along a tradeoff axis between survival under stressful

259   conditions and rapid growth under optimal conditions (Wright et al. 2010). Therefore, the

260   clustering we found in wood density suggests niche partitioning along a growth-mortality

261   tradeoff axis, i.e. the successional niche (Tilman 1988). The precise positions of the clusters

262   could have many interacting causes: the available pool of traits, external filtering from the

263   environment, resource availability, and the shape of the competition kernel. In the case of wood

264   density, given its connection to the successional niche, the distribution of light gaps in the forest

265   may possibly determine the positions of the niches. The number of clusters may also be

266   influenced by the rate at which competition drops between species of disparate wood density: the

267   faster the drop, the more niches are allowed.

268   The wood density results are not fully independent from our maximum height results, as

269   indicated by the correlation between height and wood density clusters. To fully delineate the

270   niche space driving pattern on BCI, a multivariate trait treatment would be required. However,

271   such an approach raises questions of how to best convert trait values along different trait axis

272   into a good predictor of competition. Moreover, traits could combine in different ways to form

273   multidimensional niche space (D'Andrea et al. 2018); for example, Euclidean distance may be a

274   poor predictor of competition, and therefore not be useful for detecting clustering. Determining

275   how competition maps to multivariate trait space requires mechanistic niche models going

276   beyond the existing theoretical literature, which typically focuses on contributions of individual

277   traits.

278   The absence of clustering in LMA and seed mass does not preclude a role for these traits

279   in niche partitioning on BCI. D'Andrea et al. (2019) showed that a loose connection between

280   trait value and niche strategy can mask clustering along the relevant niche axis. Seed mass is

281   linked to dispersal and recruitment ability (Muller-Landau et al. 2008), and is therefore a

282   candidate for clustering along a niche axis characterizing a competition-colonization tradeoff –

283   hence its inclusion in our study. However, seed mass alone may be a poor predictor of a species'

284   strategy along this niche axis, due for example to the effects of other traits such as drought

285   tolerance and dispersal mode. LMA is theoretically connected to niche strategies (Wright et al.

286   2010), but there is substantial ontogenetic variation in its predictive ability of growth, with the

287   relationship being stronger in the seedling stage. Eventually, other aspects of physiology become

288   limiting as the plant grows. Hence, LMA may be a noisy predictor of the niche strategy (Wright

289   et al. 2010). We conclude that the lack of clustering in these traits may be due to their loose

290   connection to niche strategies rather than no involvement in niche partitioning.

291   Our study provides a new line of evidence that niche differentiation shapes tropical

292   forests, notwithstanding neutral theory's success at fitting species abundance distributions

293   (Rosindell et al. 2012). Yet, our results go further by identifying specific numbers of niches and

294   their associated trait optima. Future extensions of clustering analysis could look at other traits

295   related to defense chemistry or abiotic preferences, further delineating the niches contributing to

296   coexistence in tropical forests.

297

306

307   **Literature cited**

308   Barabás, G., R. D'Andrea, R. Rael, G. Meszéna, and A. M. Ostling. 2013. Emergent Neutrality

309   or Hidden Niches? Oikos 122:1565–1572.

310   Chase, J. M., and M. Leibold. 2003. Ecological Niches: Linking Classical and Contemporary

311   Approaches. The University of Chicago Press.

312   Chase, J. M., and J. a. Myers. 2011. Disentangling the importance of ecological niches from

313   stochastic processes across scales. Philosophical Transactions of the Royal Society B:

314   Biological Sciences 366:2351–2363.

315   Condit, R. 1998. Tropical forest census plots: methods and results from Barro Colorado Island,

316   Panama and a comparison with other plots. Springer Science & Business Media.

317   Condit, R., R. A. Chisholm, and M. O. Hunter. 2012a. Thirty Years of Forest Census at Barro

318   Colorado and the Importance of Immigration in Maintaining Diversity. PLoS ONE 7:1–6.

319   Condit, R., B. M. J. Engelbrecht, D. Pino, R. Pérez, and B. L. Turner. 2013. Species distributions

320   in response to individual soil nutrients and seasonal drought across a community of tropical

321   trees. Proceedings of the National Academy of Sciences of the United States of America

322   110:5064–8.

323   Condit, R., S. Lao, R. Perez, S. B. Dolins, R. Foster, and S. Hubbell. 2012b. [Dataset:] Barro

324   Colorado Forest Census Plot Data (Version 2012).

325   Condit, R. S., S. P. Hubbell, and R. B. Foster. 1995. Mortality rates of 205 Neotropical tree and

326   shrub species and the impact of a severe drought. Ecological Monographs 65:419–439.

327   Condit, R. S., R. Pérez, S. Aguilar, and S. Lao. 2016. Data from Tree Censuses and Inventories

328   in Panama. repository.si.edu/handle/10088/28866.

329   D'Andrea, R., and A. Ostling. 2017. Biodiversity maintenance may be lower under partial niche

330   differentiation than under neutrality. Ecology 98:3211–3218.

331   D'Andrea, R., and A. M. Ostling. 2016. Challenges in linking trait patterns to niche

332   differentiation. Oikos 125:1369–1385.

333   D'Andrea, R., A. Ostling, and J. P. O'Dwyer. 2018. Translucent windows: How uncertainty in

334   competitive interactions impacts detection of community pattern. Ecology Letters:826–835.

335 D'Andrea, R., M. Riolo, and A. M. Ostling. 2019. Generalizing clusters of similar species as a

336   signature of coexistence under competition. PLoS Computational Biology 15:1–19.

337 Götzenberger, L., F. de Bello, K. A. Bråthen, J. Davison, A. Dubuis, A. Guisan, J. Lepš, R.

338   Lindborg, M. Moora, M. Pärtel, L. Pellissier, J. Pottier, P. Vittoz, K. Zobel, and M. Zobel.

339   2012. Ecological assembly rules in plant communities-approaches, patterns and prospects.

340   Biological Reviews 87:111–127.

341 Grime, J. P. 1979. Plant strategies and vegetation processes. Page Plant strategies and vegetation

342   processes. John Wiley and Sons.

343 Holt, R. D. 2006. Emergent neutrality. Trends in Ecology & Evolution 21:531–3.

344 Hubbell, S. P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Page

345   Monographs in Population Biology. Princeton University Press, Princeton, NJ.

346 Hubbell, S. P., R. B. Foster, S. T. O'Brien, K. E. Harms, R. Condit, B. Wechsler, S. J. Wright,

347   and S. Loo de Lao. 1999. Light-gap disturbances, recruitment limitation, and tree diversity

348   in a Neotropical forest. Science 283:554–557.

349 Kohyama, T. 1993. Size-Structured Tree Populations in Gap-Dynamic Forest--The Forest

350   Architecture Hypothesis for the Stable Coexistence of Species. Journal of Ecology 81:131–

351   143.

352 Kohyama, T., and T. Takada. 2009. The stratification theory for plant coexistence promoted by

353   one-sided competition. Journal of Ecology 97:463–471.

354 Kraft, N. J. B., and D. D. Ackerly. 2010. Functional trait and phylogenetic tests of community

355   assembly across spatial scales in an Amazonian forest. Ecological Monographs 80:401–422.

356 Kraft, N. J. B., R. Valencia, and D. D. Ackerly. 2008. Functional traits and niche-based tree

357   community assembly in an Amazonian forest. Science (New York, N.Y.) 322:580–2.

358   Kunstler, G., D. Falster, D. A. Coomes, F. Hui, M. Kooyman, Robert, D. C. Laughlin, L. Poorter,

359   M. Vanderwel, G. Vieilledent, S. J. Wright, M. Aiba, C. Baraloto, J. Caspersen, J. H. C.

360   Cornelissen, S. Gourlet-Fleury, M. Hanewinkel, B. Herault, J. Kattge, H. Kurokawa, Y.

361   Onoda, J. Penuelas, H. Poorter, M. Uriarte, S. Richardson, and P. Ruiz-Benito. 2016. Plant

362   functional traits have globally consistent effects on competition. Nature 529:1–15.

363   Lake, J. K., and A. M. Ostling. 2009. Comment on "Functional traits and niche-based tree

364   community assembly in an Amazonian forest". Science (New York, N.Y.) 324:1015; author

365   reply 1015.

366   Leimar, O., A. Sasaki, M. Doebeli, and U. Dieckmann. 2013. Limiting similarity, species

367   packing, and the shape of competition kernels. Journal of Theoretical Biology 339:3–13.

368   Levine, J. M., and D. J. Murrell. 2003. The Community-Level Consequences of Seed Dispersal

369   Patterns. Annual Review of Ecology, Evolution, and Systematics 34:549–574.

370   MacArthur, R. H., and R. Levins. 1967. The Limiting Similarity, Convergence, and Divergence

371   of Coexisting Species. The American Naturalist 101:377–385.

372   MacQueen, J. B. 1967. Some Methods for classification and Analysis of Multivariate

373   Observations. 5th Berkeley Symposium on Mathematical Statistics and Probability 1967

374   1:281–297.

375   Mouquet, N., and M. Loreau. 2003. Community Patterns in Source-Sink Metacommunities. The

376   American Naturalist 162:544–557.

377   Muller-Landau, H. C. 2010. The tolerance-fecundity trade-off and the maintenance of diversity

378   in seed size. Proceedings of the National Academy of Sciences of the United States of

379   America 107:4242–7.

380   Muller-Landau, H. C., S. J. Wright, O. Calderón, R. Condit, and M. O. Hunter. 2008.

381   Interspecific variation in primary seed dispersal in a tropical forest. Journal of Ecology

382   96:653–667.

383   Paul W Richards. 1952. The Tropical Rain Forest. The Univeristy Press, Cambridge.

384   Rosindell, J., S. P. Hubbell, F. He, L. J. Harmon, and R. S. Etienne. 2012. The case for

385   ecological neutral theory. Trends in Ecology & Evolution 27:204–209.

386   Sakavara, A., G. Tsirtsis, D. L. Roelke, R. Mancy, and S. Spatharis. 2018. Lumpy species

387   coexistence arises robustly in fluctuating resource environments. Proceedings of the

388   National Academy of Sciences.

389   Scheffer, M., and E. H. van Nes. 2006. Self-organized similarity, the evolutionary emergence of

390   groups of similar species. Proceedings of the National Academy of Sciences of the United

391   States of America 103:6230–6235.

392   Sedio, B. E., and A. M. Ostling. 2013. How specialised must natural enemies be to facilitate

393   coexistence among plants? Ecology Letters 16:995–1003.

394   Terborgh, J. 1985. The Vertical Component of Plant Species Diversity in Temperate and

395   Tropical Forests. The American Naturalist 126:760–776.

396   Tibshirani, R., G. Walther, and T. Hastie. 2001. Estimating the number of clusters in a data set

397   via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical

398   Methodology) 63:411–423.

399   Tilman, D. 1988. Plant strategies and the dynamics and structure of plant communities. Princeton

400   University Press.

401   Vergnon, R., N. K. Dulvy, and R. P. Freckleton. 2009. Niches versus neutrality: uncovering the

402   drivers of diversity in a species-rich community. Ecology letters 12:1079–90.

403   Violle, C., M.-L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel, and E. Garnier. 2007. Let

404   the concept of trait be functional! Oikos 116:882–892.

405   Visser, M. D., M. Bruijning, S. J. Wright, H. C. Muller-Landau, E. Jongejans, L. S. Comita, and

406   H. de Kroon. 2016. Functional traits as predictors of vital rates across the life cycle of

407   tropical trees. Functional Ecology 30:168–180.

408   Volkov, I., J. R. Banavar, M. O. Hunter, and A. Maritan. 2003. Neutral theory and relative

409   species abundance in ecology. Nature 424:1035–7.

410   Wright, S. J., K. Kitajima, N. J. B. Kraft, P. B. Reich, I. J. Wright, D. E. Bunker, R. Condit, J. W.

411   Dalling, S. J. Davies, S. Díaz, B. M. J. Engelbrecht, K. E. Harms, M. O. Hunter, C. O.

412   Marks, M. C. Ruiz-Jaen, C. M. Salvador, and A. E. Zanne. 2010. Functional traits and the

413   growth-mortality trade-off in tropical trees. Ecology 91:3664–74.

414   Yan, B., J. Zhang, Y. Liu, Z. Li, X. Huang, W. Yang, and A. Prinzing. 2012. Trait assembly of

415   woody plants in communities across sub-alpine gradients: Identifying the role of limiting

416   similarity. Journal of Vegetation Science 23:698–708.

417

418 **Tables and Figures**

419 **Table 1: Clustering results for the 50-ha plot and regional pool**. For each header, the left

420 column refers to the 50-ha BCI plot ("50 ha."), and the right column refers to the regional pool

421 ("Region"). The total number of species for which trait data were available is listed under "Species", and

422 the total number of individuals represented is listed under "Individuals." The number of clusters K is the

423 value at which the gap index peaks. z-scores and p-values were obtained by comparing the gap statistic

424 against 1,000 null communities. In the case of maximum height, in addition to the peak at $K = 1$ listed

425 below, the regional pool also had a significant subpeak at $K = 4$ ($Z = 2.0, P = 0.03$).

	Species		Individuals		Clusters (K)		z-score (Z)		p-value (P)	
	50 ha.	Region	50 ha.	Region	50 ha.	Region	50 ha.	Region	50 ha.	Region
Maximum height	259	242	20,620	13,743	4	1	2.3	3.4	0.01	0.02
Wood density	229	219	17,139	9,273	5	20	2.1	1.6	0.02	0.06
Leaf mass/area	250	234	18,368	9,528	3	3	0.2	1.5	0.40	0.07
Dry seed mass	185	177	18,259	10,667	15	3	0.2	1.0	0.38	0.16

426

427   **Figure 1. Field site**. Intact tropical forest covers 20% of Panama. Our regional pool is an aggregate of 40

428   sampling sites within 30km of BCI. Combined, these sites comprise approximately 850 species, of which

429   trait data were available for 242 species (max height), 219 (wood density), 234 (leaf mass per area), 177

430   (seed mass), comprising between 9,528 and 13,743 adult trees. Barro Colorado Island, spanning roughly

431   16 km^2, sits in an artificial lake in the Panama Canal. The 50-ha Forest Dynamics Plot, located near its

432   center, is a 1,000m by 500m patch of forest containing 300 tree species and 21,000 adult trees.

433

434   **Figure 2. Results.** Stem plots show trait distribution in the 50-ha plot on BCI for maximum height (**A**),

435   wood density (**B**), leaf mass per area (**C**), and seed mass (**D**). Each stem represents a species, with its

436   trait value plotted on the abscissa and abundance on the ordinate. Data shown for the 2000 census. The

437   corresponding gap curves (**E**-**H**) plot the gap index against the potential number of clusters. The gap

438   statistic is the highest point on the curve, with the red line indicating the 95th quantile of the expected

439   gap statistic under no pattern, obtained from 1,000 null sets. The peaks above the red line in (**E**) and (**F**)

440   reveal clustering by maximum height, with four clusters, and wood density, with five. The alternating

441   colors in the respective stem plots show cluster membership of each species.

442

443   **Figure 3**. **Height layers**. Visual representation of the forest on Barro Colorado Island, highlighting the

444   four height-based groups identified by our metric: shrubs (61 $\pm$ 2 species; 8,071 $\pm$ 712 adult

445   individuals), understory treelets (71 $\pm$ 3; 7,308 $\pm$ 165), midstory trees (61 $\pm$ 1; 2,957 $\pm$ 114), and canopy

446   trees (64 $\pm$ 1; 2,836 $\pm$ 82), where the numbers represent mean and standard deviation across the 7

447   censuses. Horizontal lines show maximum height of the dominant (most abundant) species in each
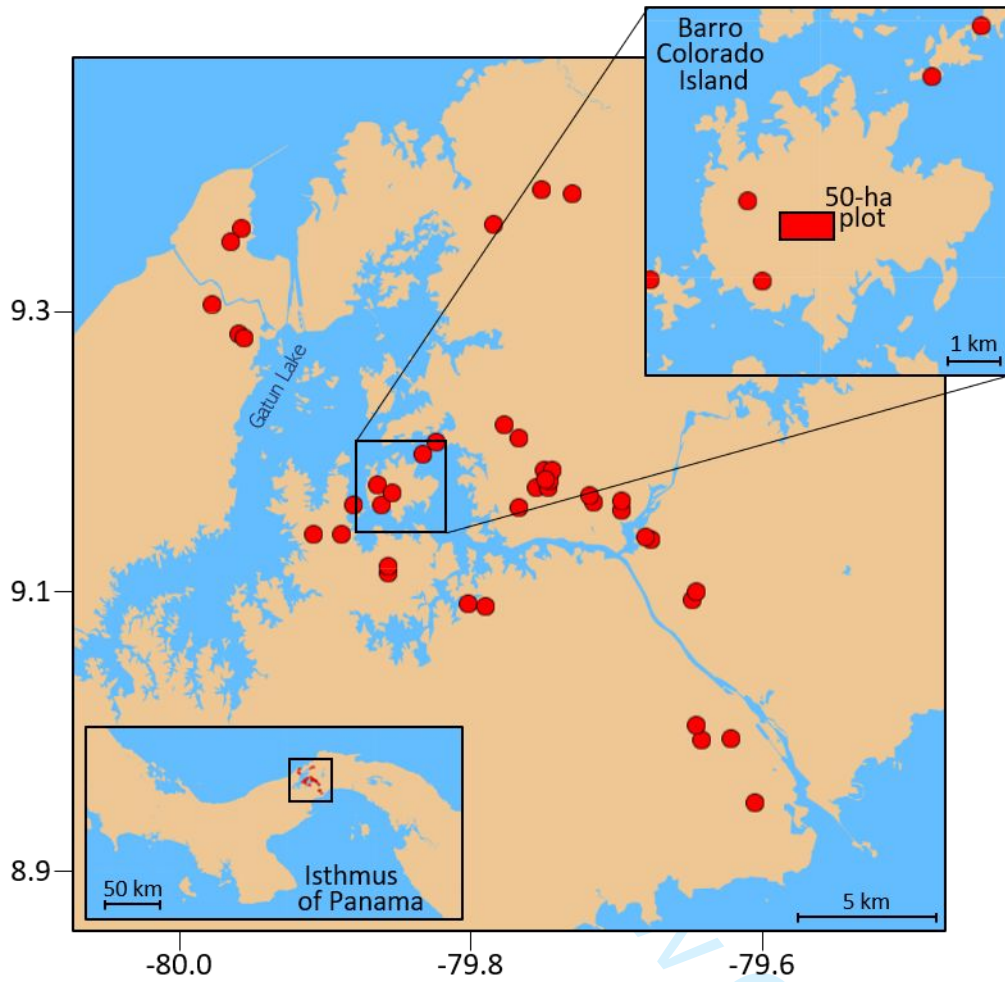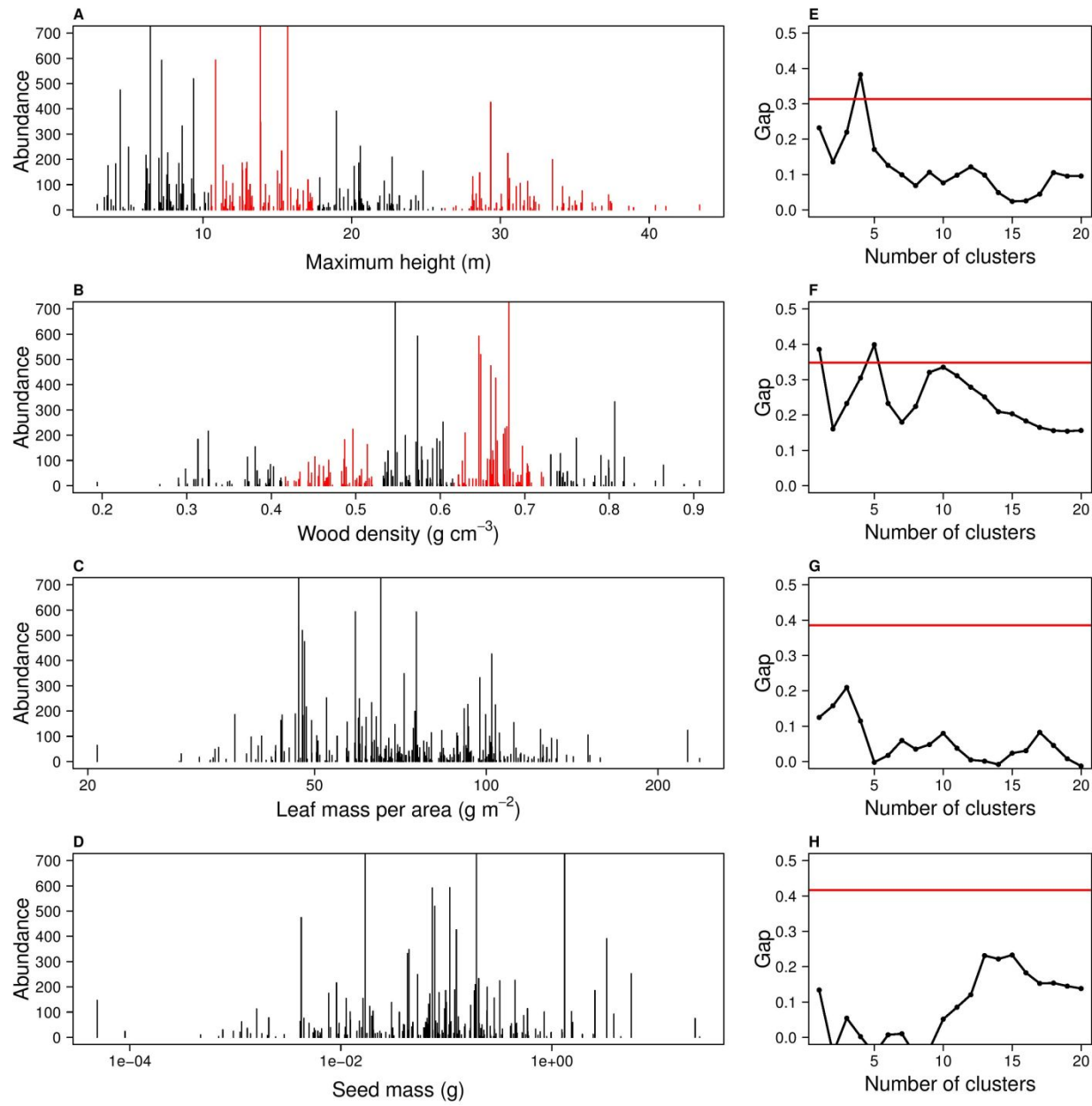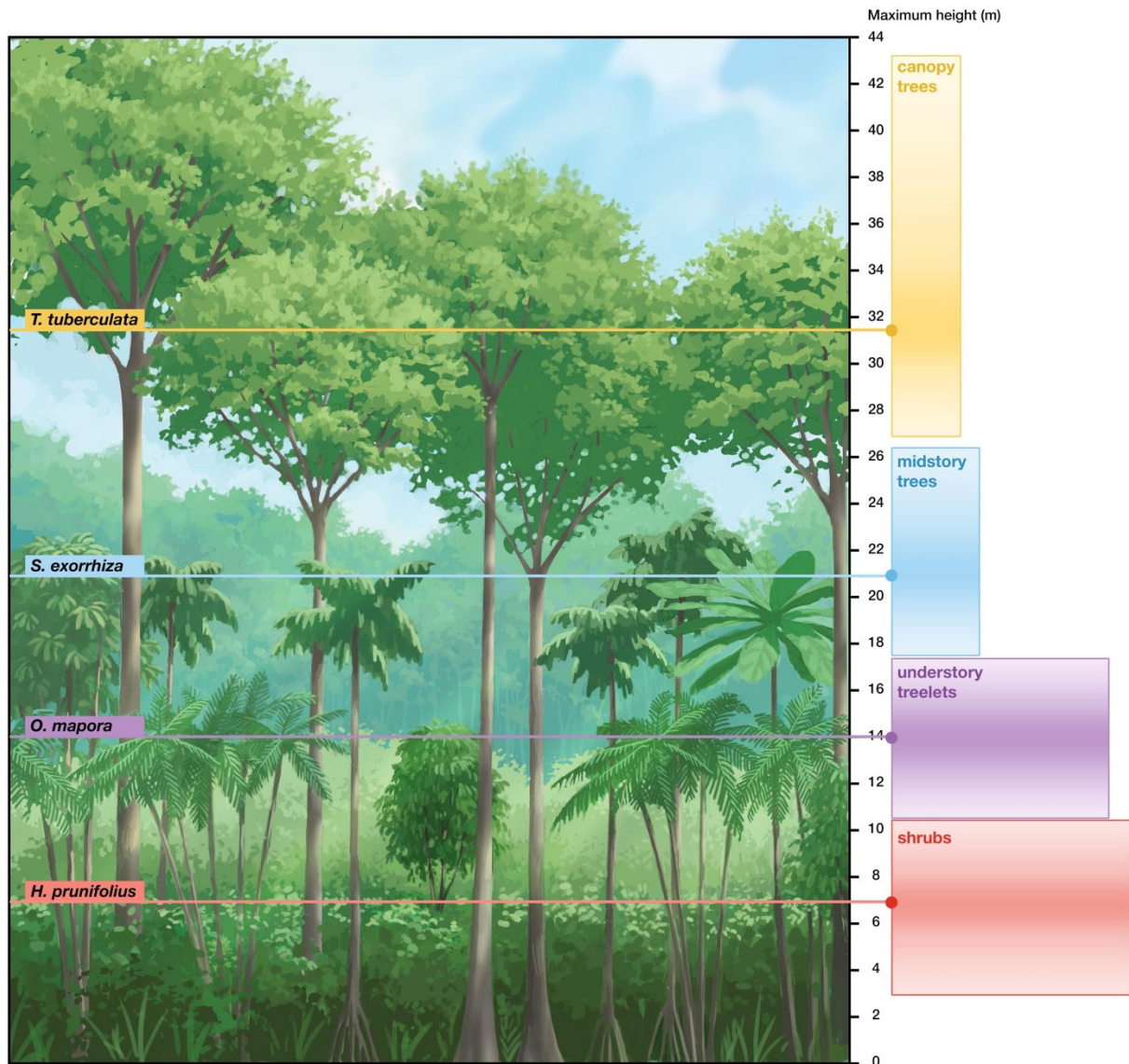
448   group.

449

450   **Figure 1.**

451

452 **Figure 2.**

453

454 **Figure 3**.

1 **Appendix S1: Supplementary tables and figures to Counting niches:**
2 **Abundance-by-trait patterns reveal niche partitioning in a Neotropical forest**
3

4 **Journal: Ecology**
5

6 Rafael D'Andrea* [1,2,†], John Guittar[1,3,‡], Hector Figueroa[1,#], Richard Condit[4,&], S. J. Wright[4,$],
7 James P. O'Dwyer[2,§] & Annette Ostling[1,¶]
8

9 [1]Dept. Ecology & Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA
10 [2]Dept. Plant Biology, University of Illinois, Urbana-Champaign, IL, USA
11 [3]Dept. Plant Biology, Michigan State University, Hickory Corners, MI, USA
12 [4]Smithsonian Tropical Research Institute, Apartado 0843–03092, Balboa, Republic of Panama
13

14 [†]rdandrea@illinois.edu ORCID 0000-0001-9687-6178 (corresponding author),
15 [‡]guittarj@msu.edu, [#]hecfox@umich.edu, [&]conditr@gmai.com, [$]wrightj@si.edu ORCID 0000-
16 0003-4260-5676, [§]jodwyer@illinois.edu, [¶]aostling@umich.edu
17
18

19   **Table S1**. Clustering results are consistent across censuses. To save computational resources, traits
20   with insignificant results were tested every two censuses. RP: regional pool. RGR: relative dbh
21   growth rate. MRT: mortality.

Trait	Census	Species	Individuals	Clusters	Z-score	P-value
Maximum height	1982	252	21,423	4	2.6	0.005
	1985	253	21,670	4	2.6	0.003
	1990	254	22,340	4	2.5	0.003
	1995	257	21,341	4	2.4	0.011
	2000	259	20,620	4	2.3	0.012
	2005	260	20,177	4	2.1	0.024
	2010	260	20,631	4	2.2	0.014
	RP	242	13,743	1	3.4	0.001
Wood density	1982	223	17,332	1	2.8	0.001
	1985	223	17,902	1	2.8	0.001
	1990	224	18,597	1	2	0.025
	1995	227	17,646	5	2.2	0.023
	2000	229	17,139	5	2.1	0.017
	2005	229	16,874	5	1.9	0.035
	2010	229	17,325	5	1.7	0.055
	RP	219	9,273	20	1.6	0.058
Leaf mass per area	1982	243	18,478	18	-1	0.839
	1990	245	19,947	3	-0.4	0.618
	2000	250	18,368	3	0.2	0.396
	2010	251	18,437	3	1	0.15
	RP	234	9,528	3	1.5	0.072
Seed mass	1982	183	18,735	16	-0.2	0.548
	1990	184	19,707	13	0.3	0.396
	2000	185	18,259	15	0.2	0.377
	2010	185	18,154	13	0.2	0.407
	RP	177	10,667	3	1	0.158

22
23

24

25

26   **Figure S1. Correspondence between quantitative height clusters and qualitative**
27   **classification**. Circle sizes are proportional to the numbers within, which show how many species
28   are classified in each combination of our four height clusters and the four qualitative height groups.
29   Largest values are close to the diagonal, indicating a good match between the two classifications.
30   Cramér's V, which measures the degree of association between two categorical variables, was 0.58
31   (0 = no association, 1 = perfect correspondence). Data for the qualitative groupings obtained from
32   Wright et al[35]. Note that this classification into height groups (x-axis) was based on typical height
33   of reproductive adults, not the species maximum height. As such, some species are classified one
34   height group below their corresponding maximum-height cluster, hence the bias towards
35   placement above the diagonal.
36
37   **Figure S2. Regional pool results.** Results displayed in analogous form to Fig. 2. Only maximum
38   height is clustered, with significance at $K = 1$ and $K = 4$. Wood density and LMA are marginally
39   significant at $K = 20$ and $K = 3$, respectively.
40
41   **Figures S3-S6. BCI results for all available censuses.** Pattern was consistent across censuses,
42   especially for maximum height (Fig S4). In wood density (Fig S5), the peak at $K = 5$ tops $K = 1$
43   as the highest gap index over time. LMA (Fig S6) and seed mass (Fig S7) results were obtained
44   for   alternate   censuses   to   save   computational   resources.
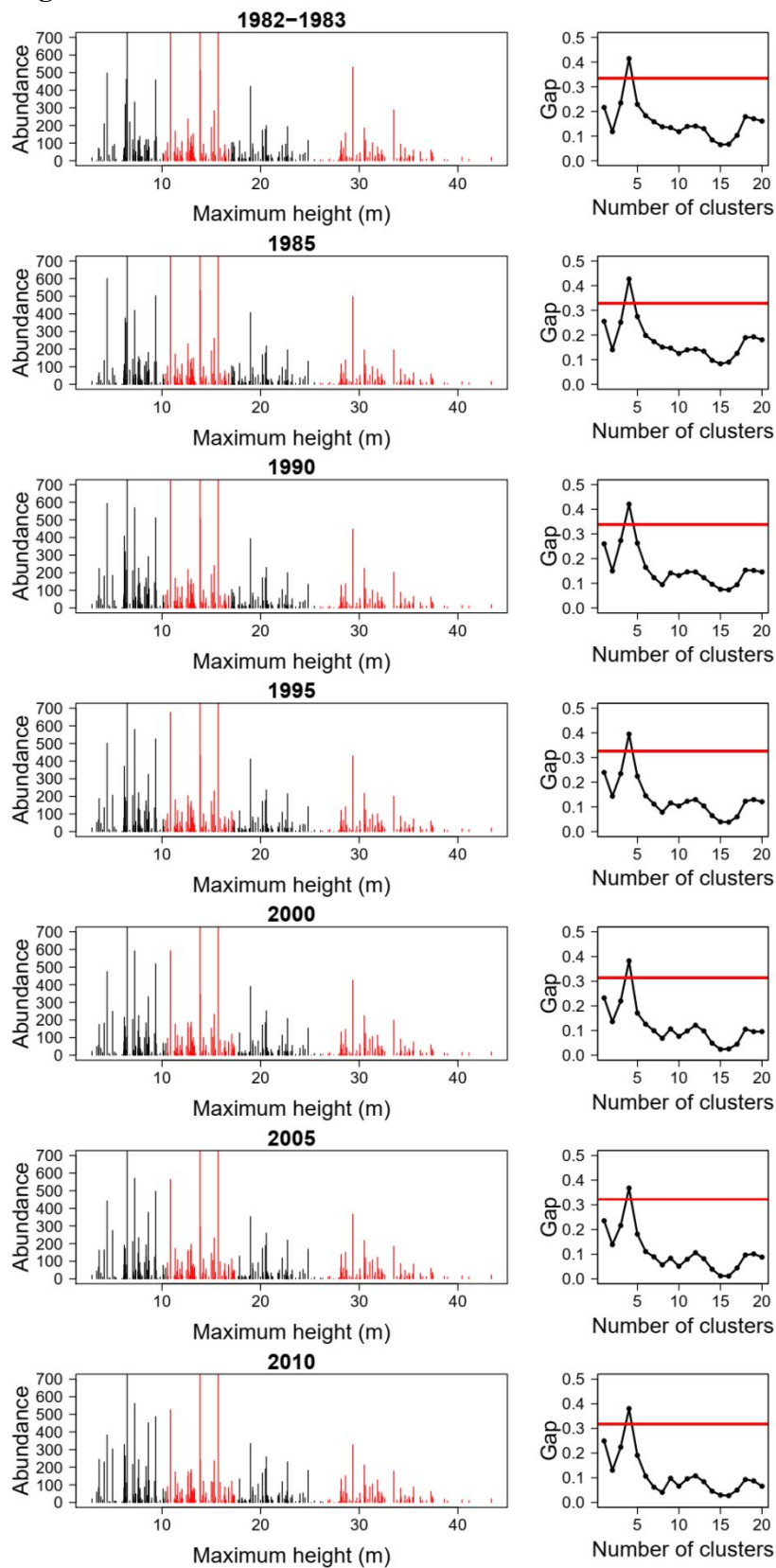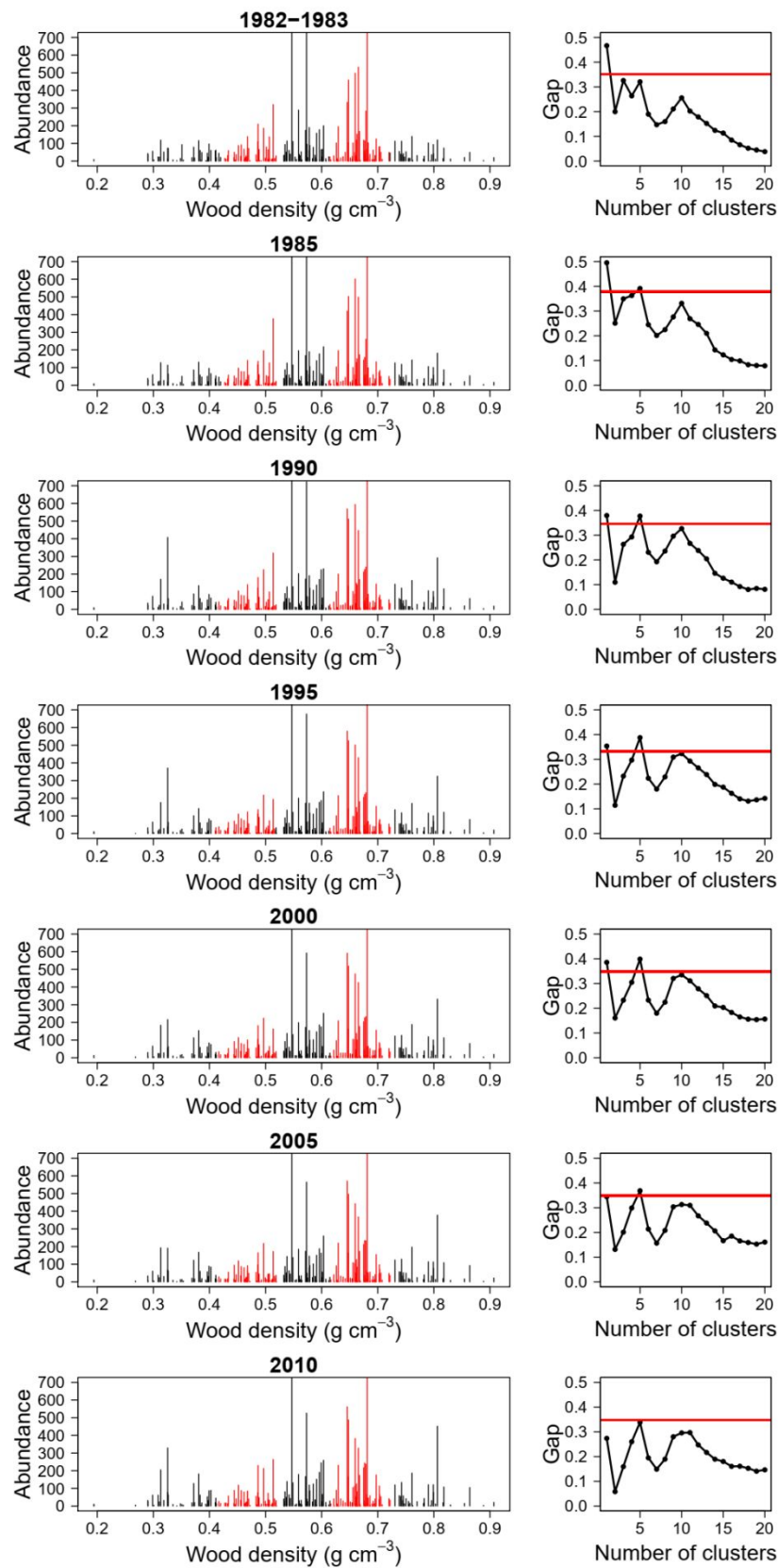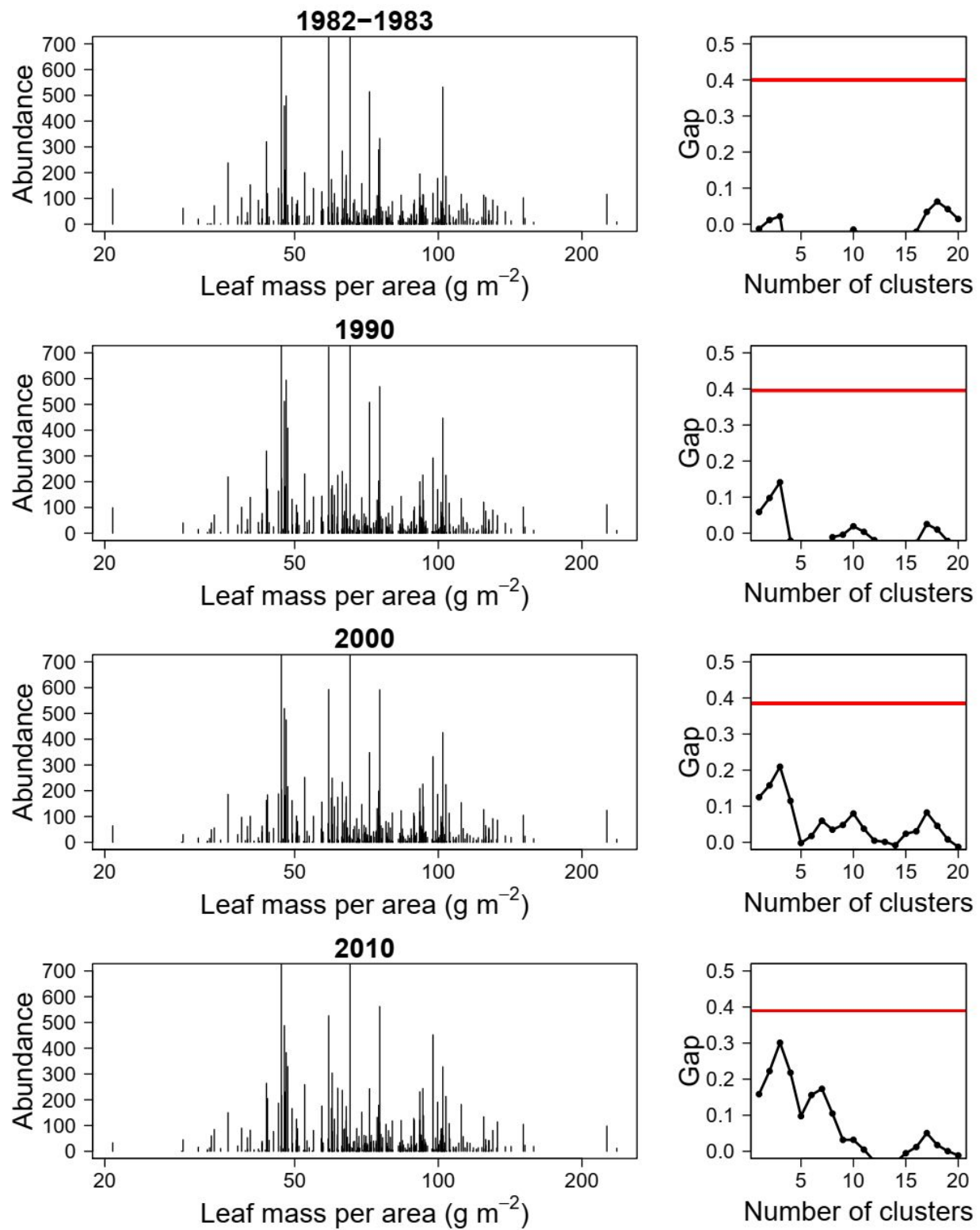
45   **Figure S1.**
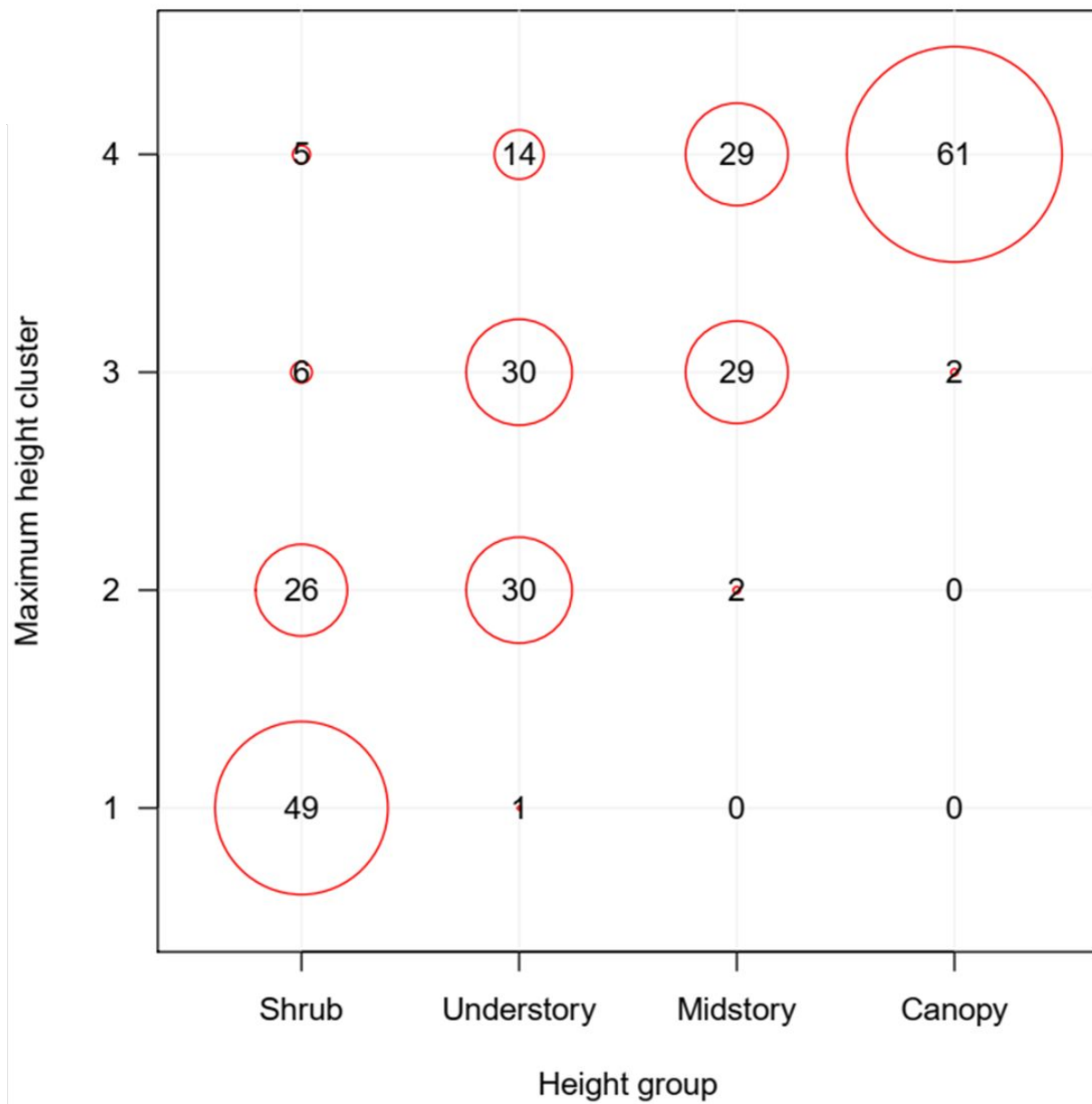
46
47

48   **Figure S2.**
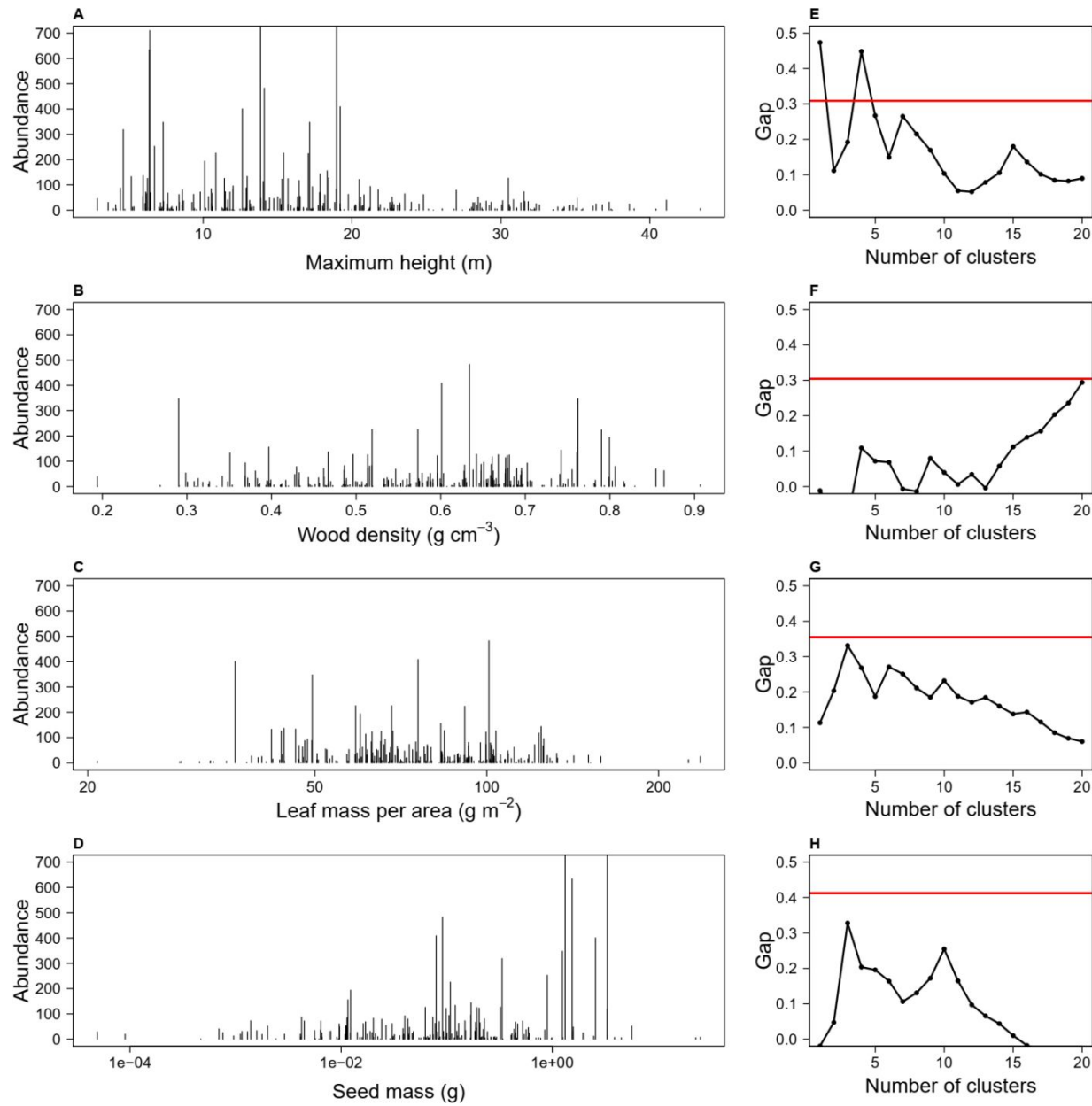
49
50

51 **Figure S3.**

54   **Figure S4.**

55

56   **Figure S5.**

59   **Figure S6.**