

## SUPPORTING INFORMATION:

### **Klimes P. *et al.* Why are there more arboreal ant species in primary than secondary tropical forests?**

#### **Appendix S1**

SUPPLEMENTARY METHODS FOR THE ALGORITHM FOR PARTITIONING OF DIFFERENCE IN SPECIES RICHNESS BETWEEN PRIMARY AND SECONDARY FOREST TREE SETS.

To partition the factors contributing towards the differences in species richness in primary and secondary forest we took a rarefaction-based approach. Using an initial dataset of 0.32ha of primary forest and 0.32ha of secondary forest we rarefied a subset of the primary forest trees to match the secondary forest in terms of 1. Tree density, 2. Tree size distribution and 3. Tree taxonomic structure, including the frequencies species, genera and families within the whole dataset, and the distribution of individuals within species within tree size classes. Note that the steps were carried out additively in the order 1-3. All of the following algorithms were programmed in R 2.12.2 (see attached txt file **Appendix S2 for code**).

1. *Tree density*: First, we calculated the density of trees in an average 0.1ha plot in secondary forest (mean of three 0.1ha continuous plots= $D$ ). We then randomly selected  $D$  trees from the whole primary forest dataset and counted the total number of ant species inhabiting this subset of trees.

This was repeated 500 times and mean ant species richness calculated.

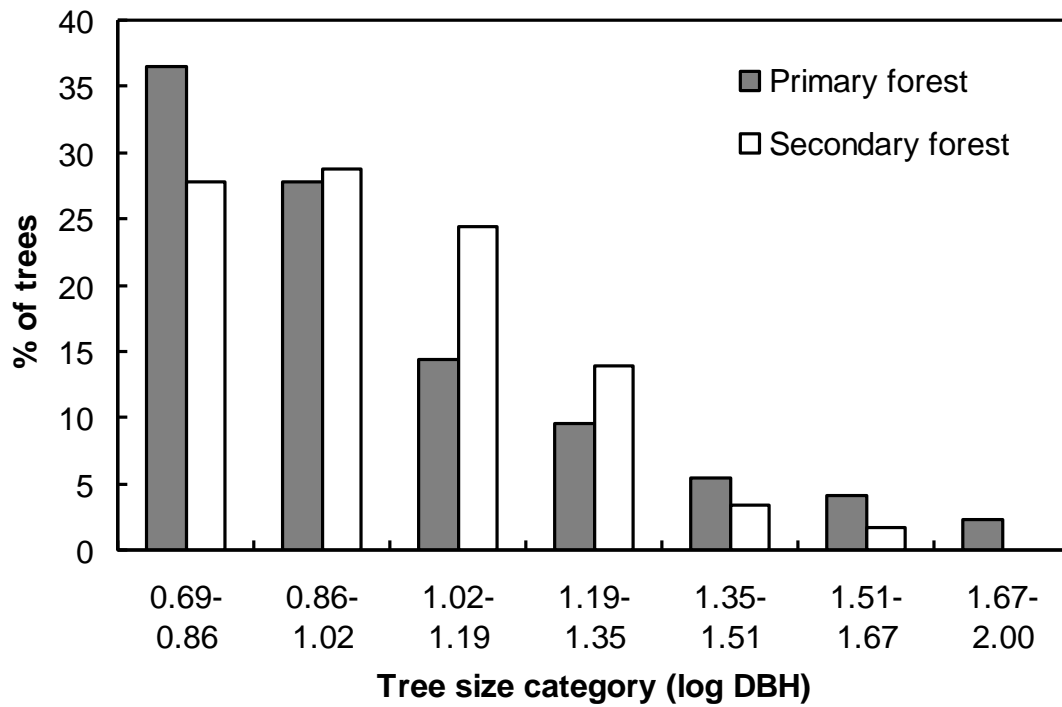
2. *Tree size distribution*: We generated a template size distribution from the secondary forest by selecting  $D$  trees from the whole secondary forest dataset and recording the abundances of these in logarithmic DBH size classes of trees (1,2...,  $i$ ) of  $d_1$  to  $d_i$  such that  $d_1 + d_2 + \dots + d_i = D$ . We then divided up all of the primary forest trees into the same size classes, and randomly drew  $d_1 + d_2 + \dots + d_i$  trees from each of these size classes. These trees were combined to form a simulated dataset and the species richness of ants associated with this rarefied community was calculated. We used

500 samplings of the secondary forest to form templates for rarefaction of the primary forest dataset. Mean ant species richness was calculated across all 500 simulated datasets.

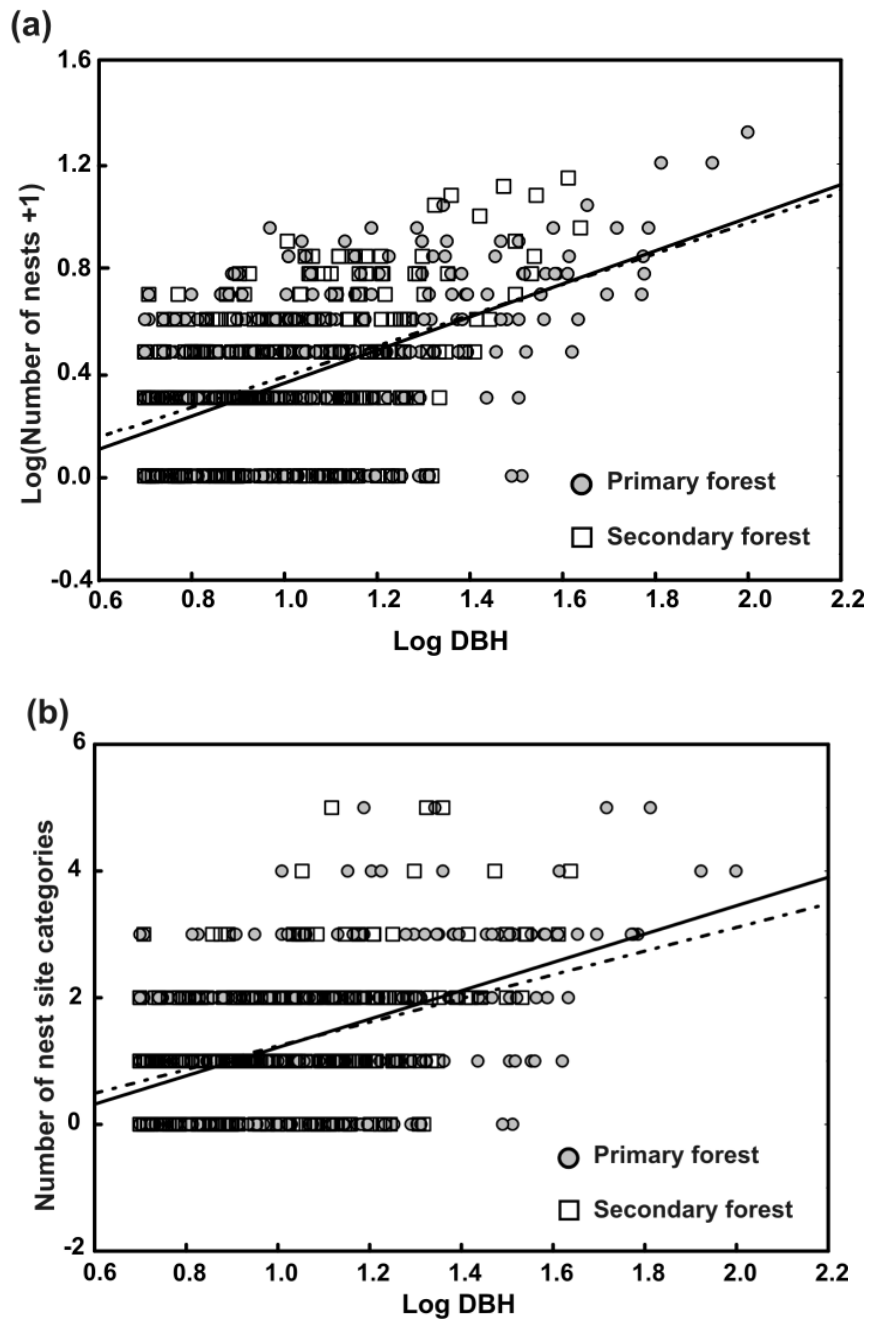
3. *Tree taxonomic structure*: D trees in total were drawn from the primary forest DBH classes as described above, with the added restriction that the abundance distribution structure within each DBH class, the number of species within each DBH class, and the total numbers of species, families and genera were to be matched as closely as possible to that of the template dataset. This was achieved by working through the abundances of tree species within the template (within a particular DBH class) and randomly selecting a species from the primary forest species pool that was at least that abundant. If the species was more abundant, a randomly selected subset of individuals was discarded to equilibrate abundances. In some cases there was no single species sufficiently abundant in the species pool, in which case the next most abundant species was selected. In these cases the total abundance within each DBH class was maintained by sequentially selecting other, less abundant species in the rarefied dataset, and replacing them with a species that was more abundant. In some rare cases there were no species remaining in the pool that were more abundant than any of those in the rarefied dataset, in which case it was necessary to select new singleton species to ensure that abundances within DBH classes were maintained. This matched relative abundance distributions and species richness within size classes, but did not match the overall taxonomy of the simulated datasets. To do this we used an iterative procedure that randomly selected a species occurrence from a particular size class, and replaced this occurrence with a randomly selected species from the primary source dataset. Species that already occurred within that size class were not made available for selection. The differences between the number of tree species, genera and families in the new dataset and template dataset were then compared (with equal weighting to all taxonomic levels). If the fit of the new dataset was better than that of the old, the new dataset was retained and used for the next iteration. If the fit was worse, the old dataset was used for the next iteration. This process stopped when either the fit was perfect, or when 3000 iterations had been

conducted. This process was repeated 500 times for different secondary forest templates and mean ant species richness was calculated, as above.

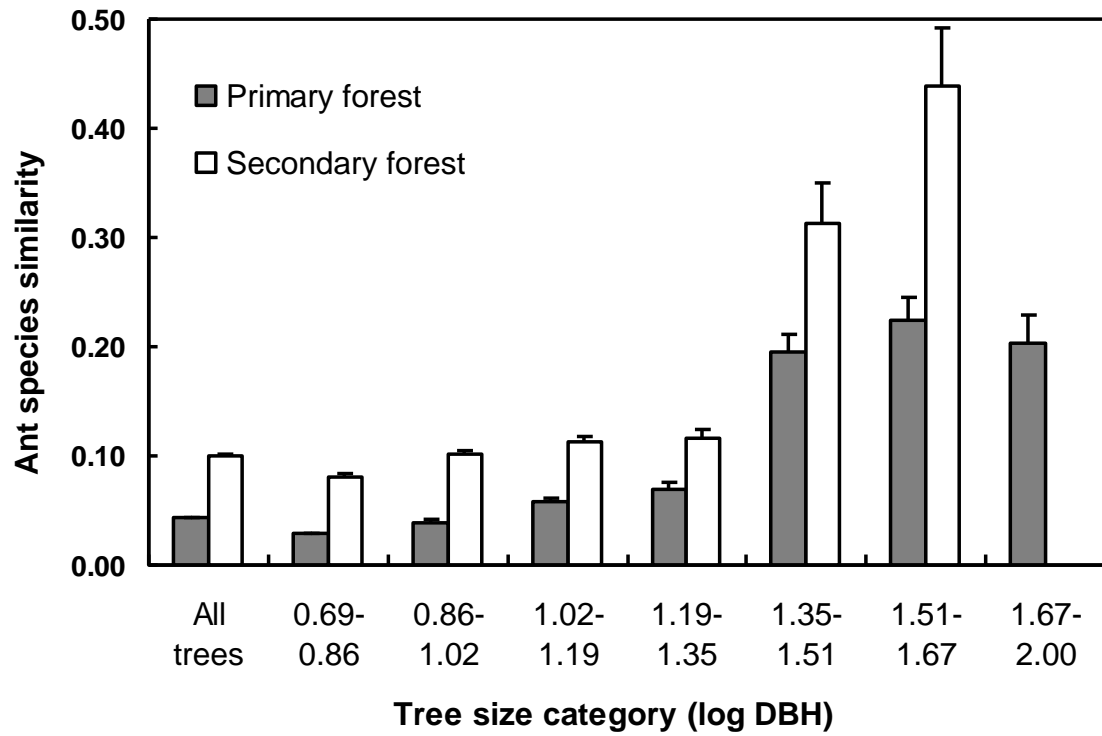
**Supplementary figures:**



**Fig. S1.** Distribution of trees among logarithmically scale size classes in primary and secondary forest (See Methods for definition of the classes). Frequencies of size classes are significantly different between both forest types ( $\chi^2 = 27.3$ , d.f. = 6,  $P < 0.001$ ).



**Fig. S2.** Relationships of tree size (logDBH) and number of nests (a), and nest site richness (N of nest site categories per tree) (b), in primary (circles and full line) and secondary forest (squares and dashed line). There is no significant difference in the regression slopes between both forests (see Table S3 for the statistic).



**Fig. S3.** Mean similarity in ant species composition between pairs of trees from the same forest type (primary forest: grey, secondary forest: white columns). The mean Sørensen similarity index ( $\pm$  SE) is reported for all sampled trees (with DBH  $\geq$  5 cm) and for all trees within each size class.

## Supplementary tables:

**Table S1.** Stepwise regression of ant species richness (i.e. the number of nesting species per tree) on tree traits related to their size: tree diameter at breast height (DBH), fresh leaf weight (LeafW), tree height (TreeH), trunk height (TrunkH), crown width (CrownW), and crown height (CrownH). DBH is the best predictor accounting for 85% of the variance in ant species richness explained by the entire model; b\* refers to standardized regression coefficient.

Log (n of ants species + 1)	Step in:	R	R <sup>2</sup>	F (in)	P	b*
<b>Log(DBH)</b>	<b>1</b>	<b>0.473</b>	<b>0.223</b>	<b>196.105</b>	<b>&lt;0.001</b>	<b>0.339</b>
<b>Log (LeafW+1)</b>	<b>2</b>	<b>0.497</b>	<b>0.247</b>	<b>21.469</b>	<b>&lt;0.001</b>	<b>0.146</b>
<b>Log (TrunkH+1)</b>	<b>3</b>	<b>0.507</b>	<b>0.257</b>	<b>9.249</b>	<b>0.002</b>	<b>-0.121</b>
<b>Log (CrownW+1)</b>	<b>4</b>	<b>0.512</b>	<b>0.262</b>	<b>4.606</b>	<b>0.032</b>	<b>0.126</b>
Log (CrownH+1)	Not selected	-	-	-	-	-
Log(TreeH)	Not selected	-	-	-	-	-

Whole model: R = 0.512, R<sup>2</sup> = 0.262, F<sub>(5,679)</sub> = 60.318 p < 0.001.





**Table S3:** Effect of forest type (primary, secondary), tree size (DBH) and their interaction on ant species richness (a), the number of their nests (b) and nest site richness (c) tested by GLM model of homogeneity of slopes. Only the effect of DBH is significant in all analyses ( $P < 0.05$ , in bold).

a)

Log (ants species number + 1)	SS	Degr. of Freedom	MS	F	<i>P</i>
Forest type	0.060	1	0.060	1.270	0.260
<b>Log(DBH)</b>	<b>7.581</b>	<b>1</b>	<b>7.581</b>	<b>161.25</b>	<b>&lt; 0.001</b>
Forest * Log(DBH)	0.042	1	0.042	0.900	0.343
Error	31.968	680	0.047		

b)

Log (nest number + 1)	SS	Degr. of Freedom	MS	F	<i>P</i>
Forest type	0.039	1	0.039	0.650	0.420
<b>Log(DBH)</b>	<b>12.743</b>	<b>1</b>	<b>12.743</b>	<b>210.16</b>	<b>&lt; 0.001</b>
Forest * Log(DBH)	0.017	1	0.017	0.285	0.593
Error	41.235	680	0.060		

c)

Number of nest site categories	SS	Degr. of Freedom	MS	F	<i>P</i>
Forest type	1.22	1	1.22	1.39	0.238
<b>Log(DBH)</b>	<b>142.96</b>	<b>1</b>	<b>142.96</b>	<b>162.12</b>	<b>&lt; 0.001</b>
Forest * Log(DBH)	1.14	1	1.14	1.29	0.255
Error	599.67	680	0.88		