

This is the peer reviewed version of the following article: Paine, C. E. T., Baraloto, C., Díaz, S. (2015), Optimal strategies for sampling functional traits in species-rich forests. *Functional Ecology*, 29: 1325–1331, which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1111/1365-2435.12433/full>.

This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for self-archiving.

**TITLE**

Optimal strategies for sampling functional traits in species-rich forests

**AUTHORS AND AFFILIATIONS**

C. E. Timothy Paine (Corresponding author), [c.e.t.paine@stir.ac.uk](mailto:c.e.t.paine@stir.ac.uk)

- Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA, UK, phone: +44 (0)1786 467785, fax: +44 (0) 1786 467843

Christopher Baraloto, [Chris.Baraloto@ecofog.gf](mailto:Chris.Baraloto@ecofog.gf)

- INRA, UMR “Ecologie des Forêts de Guyane”, 97387, Kourou Cedex, French Guiana
- International Center of Tropical Botany, Department of Biology, Florida International University, Miami, Florida, 33199 USA

Sandra Diaz, [sdiaz@efn.uncor.edu](mailto:sdiaz@efn.uncor.edu)

- Instituto Multidisciplinario de Biología Vegetal (IMBIV-CONICET) and Departamento de Diversidad Biológica y Ecología, FCEfyN, Universidad Nacional de Córdoba, CC 495, 5000 Córdoba, Argentina.

**AUTHOR CONTRIBUTIONS**

All authors conceived the question, CETP and CB collected data with the BRIDGE team, CETP performed the simulations and wrote the first paper with contributions from all authors.

**RUNNING TITLE**

Optimal trait-sampling strategies

**KEYWORDS**

French Guiana, functional traits, plant traits, specific leaf area, wood density, sampling design, tropical forest

**WORD COUNTS**

Abstract: 156

Main text: 3072

Number of references: 34

Figures: 2

Tables: 1

38 **ABSTRACT**

39           Functional traits provide insight into a variety of ecological questions, yet the optimal  
40 sampling method to estimate the community-level distribution of plant functional trait values  
41 remains a subject of debate, especially in species-rich forests. We present a simulation analysis  
42 of the trait distribution of a set of nine completely sampled permanent plots in the lowland rain  
43 forests of French Guiana. Increased sampling intensity consistently improved accuracy in  
44 estimating community-weighted means and variances of functional trait values, whereas there  
45 was substantial variation among functional traits, and minor differences among sampling  
46 strategies. Thus, investment in intensified sampling yields a greater improvement in the accuracy  
47 of estimation than does an equivalent investment in sampling design complication. Notably,  
48 ‘taxon-free’ strategies frequently had greater accuracy than did abundance-based strategies,  
49 which had the additional cost of requiring botanical surveys. We conclude that there is no  
50 substitute for extensive field sampling to accurately characterize the distribution of functional  
51 trait values in species-rich forests.

52

53

54 **INTRODUCTION**

55           Functional traits impact plant fitness via their effects on recruitment, growth,  
56 reproduction and survival (Lavorel & Garnier 2002; Violle *et al.* 2007). Scaling from functional  
57 traits measured on individuals up to community-level distributions of trait values has provided  
58 insight into a variety of issues surrounding community assembly and ecosystem processes  
59 (McGill *et al.* 2006). Traits have been particularly useful in species-rich forests where the rarity  
60 of many species has stymied efforts to characterize their niche requirements (Baraloto, Paine,  
61 Poorter, *et al.* 2010; Fortunel *et al.* 2013). Standardized protocols for the measurement of  
62 functional traits and intensive efforts to compile trait data have fuelled explosive growth of  
63 functional trait databases (Kattge *et al.* 2011; Pérez-Harguindeguy *et al.* 2013), which in turn,  
64 support the development of trait-based dynamic global vegetation models (Scheiter, Langan, &  
65 Higgins 2013).

66           An investigator's choice of sampling strategy can have a major influence on the  
67 inferences he or she makes regarding the strength and extent of ecological processes. For  
68 example, common sampling designs for coring of trees can systematically bias estimations of  
69 forest growth and productivity (Nehrbass-Ahles *et al.* 2014). The increasing use of functional  
70 traits in plant ecology spurred Baraloto *et al.* (2010a) to assess trait-sampling strategies in  
71 species-rich forests. After measuring the traits on every individual tree in nine 1-ha plots of  
72 lowland tropical rain forest, they simulated performing four sampling strategies over a gradient  
73 of sampling intensity. Their study, however, suffered from a number of defects, which we  
74 remedy in the current contribution. First, their analysis did not assess interactions between  
75 strategy and intensity. Second, Baraloto *et al.* (2010a) classified the results of their simulations  
76 into successes and failures using [an arbitrary criterion](#), then analysed these binomally-distributed

77 variables as though they followed Gaussian distributions. Third, their analysis did not take into  
78 account variation among sites. Finally, Baraloto *et al.* (2010a) examined sampling strategies  
79 based on species abundance only in passing, although such strategies are widely used in  
80 functional trait studies in [extra](#)-tropical ecosystems (Conti & Díaz 2013; Fisichelli, Frelich, &  
81 Reich 2014; Lavorel *et al.* 2008; Pakeman & Quested 2007). This paper updates Baraloto *et al.*  
82 (2010a) with an expanded and refined response to the question: What sampling strategy yields  
83 the most accurate characterization of community-level trait distributions in species-rich forest  
84 communities?

85         We expand upon Baraloto *et al.* (2010a) by evaluating the interacting effects of sampling  
86 strategy, sampling intensity, and functional trait identity on the probability of accurately  
87 characterizing the community-level distribution of functional trait values. We characterize the  
88 distribution of trait values in terms of its mean and variance (Díaz *et al.* 2007; Violle *et al.* 2007).  
89 First, we obtain the true community-level distributions by assessing the trait values of every  
90 individual tree (Baraloto *et al.* 2010). Then, in simulation, we estimate the mean and variance of  
91 functional traits that result from applying each sampling strategy over a range of sampling  
92 intensities. We compare these estimates with the true values to determine the accuracy of  
93 estimation for every combination of strategy, intensity and trait. We include sampling strategies  
94 in which individuals are selected based upon their regional or local stem density or basal area, as  
95 measures of abundance, and we analyse their accuracy using mixed-effect models.

96

## 97 **METHODS**

98         We sampled functional traits on 4672 individual trees representing 668 species in nine 1-  
99 ha plots of lowland rain forest in French Guiana between November 2007 and September 2009

100 (Baraloto, Paine, Patiño, et al. 2010). In each plot, all trees >10 cm diameter at breast height  
101 (d.b.h.) were mapped, measured for d.b.h. and climbed to obtain a branch for leaf samples. For  
102 the current study, we used the same 10 functional traits as analysed by Baraloto *et al.* (2010a).  
103 [The eight](#) leaf traits are correlated with the global leaf economics spectrum (Díaz et al. 2004;  
104 Wright et al. 2004): chlorophyll content, concentrations of  $\delta^{13}\text{C}$ , N and  $\delta^{15}\text{N}$ , C:N ratio, thickness,  
105 toughness and specific leaf area (SLA), one [is](#) associated with the global wood economics  
106 spectrum (Chave et al. 2009): bole density, and one [is](#) of uncertain affinities (Baraloto, Paine,  
107 Poorter, et al. 2010): laminar surface area. The methods of trait sampling in the field and  
108 subsequent laboratory analysis are detailed in Baraloto *et al.* (2010a).

109 We investigated the accuracy of 12 plot-based strategies. [Three are based on abundance](#)  
110 [estimated as basal area, three are based on abundance estimated as number of stems, three are](#)  
111 [‘taxon-free’ \(sensu Lavorel et al. 2008\), and the final three have varying motivations.](#) The  
112 strategies vary widely in the amount of field data required prior to trait sampling, and each  
113 emphasizes different aspects of the tree community (Table 1). Most strategies were included  
114 because they have recently been used in published studies. The BA REGION and BA LOCAL  
115 strategies sample individuals on the basis of the basal area of their conspecifics in the region or  
116 local plot, respectively. BA INDIVIDUAL, on the other hand, samples individuals on the basis  
117 of their own basal area. It, therefore, tends to sample very large individuals. As such, BA  
118 REGION requires botanical inventories and d.b.h. measurements from multiple plots, whereas  
119 BA LOCAL and BA INDIVIDUAL require botanical inventories and d.b.h. measurements from  
120 a single local plot. Variants of these strategies have been applied in non-tropical communities  
121 (Conti & Díaz 2013; Fisichelli et al. 2014; Lavorel et al. 2008; Pakeman & Quested 2007). The  
122 commonness strategies (COMMON REGION and COMMON LOCAL) are similar, but are

123 based on stem density in the region or in the local plot, and require either regional or local  
124 botanical inventories, respectively. The SPECIES strategy flips COMMON LOCAL on its head,  
125 and weights species inversely by their local abundance, thus emphasizing relatively rare species.

126 The three taxon-free strategies require no botanical determinations. In CLUMP, three  
127 random [spatial](#) locations within the plot are selected, and trees were selected based on their  
128 proximity to these [foci](#). A variant of CLUMP has been used by the RAINFOR project (Patiño et  
129 al. 2009). The SIZE strategy stratifies the tree community into five equally spaced size-classes  
130 based on d.b.h, from which individuals were drawn at random (Poorter, Bongers, & Bongers  
131 2006). In RANDOM, individual trees are randomly selected from the plot. Although we have not  
132 seen this strategy used in the literature, it provides a basis for comparison with other more  
133 complicated strategies.

134 The final three strategies are more idiosyncratic. The ONE PER SP strategy samples a  
135 single individual from each species in each plot. Trait values for all individuals of each species in  
136 the plot are then set equal to the conspecifics of the sampled individuals (Kraft, Valencia, &  
137 Ackerly 2008). This strategy requires a full botanical inventory prior to sampling. The  
138 DATABASE strategy samples one individual per species from the region and applies that  
139 individual's trait values to all conspecifics (Paine et al. 2012; ter Steege et al. 2006). Note that  
140 our implementation of the DATABASE strategy may overestimate its accuracy, because the trait  
141 values we use come from a regional data base including only the plots under consideration,  
142 rather than from other regions, and because species–trait combinations for which no data were  
143 available were left as missing data, rather than being replaced with genus- or family level means  
144 (Baraloto, Paine, Patiño, et al. 2010). Finally, the basis of comparison for all sampling strategies  
145 was the BRIDGE strategy, in which every individual in the plot is sampled (Baraloto et al. 2012).

146 We simulated each strategy 500 times in each of the nine plots in the BRIDGE network.  
147 Sampling strategies, except those for which sampling intensity does not vary (BRIDGE,  
148 DATABASE and ONE PER SP), were executed over a gradient of sampling intensity (1, 2, 3, 5,  
149 10, 20, 30, 40 and 50% of individuals sampled per plot). Sampling intensities greater than 50%  
150 are impractical as they require essentially as much work as does 100% (BRIDGE) sampling. In  
151 contrast to Baraloto *et al.* (2010a), the sampling of individuals was probabilistic. Thus, in  
152 RANDOM, all individuals in a plot had an equivalent probability of being sampled (equal  
153 weights for all individuals), whereas unequal weights were used in all other strategies. In BA  
154 REGION, for example, an individual's probability of being sampled was proportional to the BA  
155 of its conspecifics in the region (Table 1).

156 We assessed the performance of each sampling strategy, sampling intensity and trait as  
157 the accuracy in recovering the true mean and variance of trait values. Baraloto *et al.* (2010),  
158 considered resampling events as successes (or failures) with respect to an arbitrary threshold. We,  
159 in contrast, define accuracy as the absolute value of the per cent difference between the trait  
160 values estimated in each re-sampling event to those derived from BRIDGE sampling. We  
161 analysed the performance of each of the nine strategies for which intensity was varied using a  
162 [linear mixed-effect model](#). Accuracy values were log-transformed for analysis to improve the  
163 normality of residuals, and back-transformed for presentation. Tests of significance are irrelevant  
164 in the analysis of simulation output. Rather, our primary interest was to assess the relative  
165 importance of variation among sampling strategies, functional traits, and sampling intensity in  
166 determining the accuracy of estimation of the mean and variance in trait values. Thus, we  
167 modelled the log-transformed accuracy of estimation as a normally distributed variable based on  
168 the three-way interaction of sampling intensity, strategy and trait. Plots were included as a

169 random effect. We calculated the percentage of variance in the accuracy of estimations explained  
170 by sampling strategy, intensity and trait identity using the method of Nakagawa and Schielzeth  
171 (2013). Simulations and mixed-effect models were implemented in R 3.1.1 (R Core  
172 Development Team 2014) using the lme4 package (Bates et al. 2014). The R script used to  
173 conduct the simulations is available as Appendix S1 in Supporting Information.

174

## 175 **RESULTS**

176         Sampling intensity, sampling strategy, and the trait under consideration affected the  
177 probability of accurately estimating the true mean and variance in functional trait values (Fig. 1).  
178 Executed at low sampling intensities, all strategies yielded inaccurate estimates of trait means  
179 and variances, but increased intensity consistently increased the accuracy of estimation. Over a  
180 broad range of sampling intensity, the taxon-free strategies of CLUMP, SIZE and RANDOM  
181 had consistently superior performance in estimating both the mean and variance of most traits.  
182 Contrastingly, the strategies based upon basal area or commonness, whether assessed locally or  
183 regionally, performed less well. In general, performing botanical inventories prior to trait  
184 sampling did not improve the accuracy of estimating the mean or, especially, the variance in trait  
185 values. The error in estimates of the trait variances exceeded that of trait means by an order of  
186 magnitude, on average. Despite sampling of every species, and approximately one-third of the  
187 individuals in a plot, the ONE PER SP strategy performed no better than sampling an equivalent  
188 number of randomly selected individuals. Notably, the DATABASE strategy estimated means  
189 and variances approximately as well as field-based sampling strategies that were executed at  
190 10% sampling intensity. There was considerable variation among traits in the probability of  
191 correctly estimating their mean and variance, as certain traits were recalcitrant to intensified

192 sampling. Accurate estimates of mean leaf toughness and  $\delta^{15}\text{N}$  were achieved only rarely,  
193 regardless of the sampling intensity or strategy employed.

194 We partitioned the variance in the accuracy of estimating means and variances of  
195 functional trait values to determine the relative importance of sampling strategy, sampling  
196 intensity and trait identity (Fig. 2). The most important determinants of accurately estimating the  
197 mean and variance in trait values were the trait under investigation and sampling intensity,  
198 respectively. Sampling strategy, in contrast, was far less important in determining the accuracy  
199 of estimation. Interactions among sampling intensity, strategy, and trait were weak, implying that  
200 increasing sampling intensity increases the probability of accurate estimation for all traits and  
201 strategies, not only some. Similarly, sampling strategies performed equivalently well (or poorly)  
202 in estimating mean and variance for all traits. Finally, sampling strategy, intensity and trait  
203 together jointly explained more variance in the accuracy of estimation for means, whereas the  
204 percentage of variance that remained unexplained was greater for trait variance.

205

## 206 **DISCUSSION**

### 207 *Investment in sampling strategy and sampling intensity*

208 By assessing the interactions between sampling strategy, sampling intensity and trait identity, we  
209 dissected the relative contributions of each to the accurate estimation of trait distributions.

210 Overall, sampling intensity was the single best predictor of accuracy in the estimation of the  
211 mean and variance of trait values (Fig. 2). Statistical interactions with trait identity and sampling  
212 strategy were weak, indicating that more-intense sampling consistently yielded more accurate  
213 estimations. Moreover, sampling intensity explained far more variance in the probability of  
214 accurate estimation than did sampling strategy, indicating that an investment in sampling

215 intensity would yield a greater improvement in the accuracy of estimation than would an  
216 equivalent investment in a complicated sampling design. Sampling intensity was an especially  
217 strong predictor of the variance in trait values (Fig. 2). In other words, intense sampling is  
218 essential for accurately estimating trait variance. *It is not possible to advocate a minimal*  
219 *intensity of trait sampling for all trait-based studies, however, as their tolerance of inaccuracy in*  
220 *functional trait estimation will vary, depending on their objectives. Nevertheless, sampling just*  
221 *10% of individuals is likely to yield sufficiently small estimation errors in most cases.*

222         Traits varied widely in terms of the accuracy at which their distributions could be  
223 estimated. Means and variances for some, such as foliar thickness and  $\delta^{13}\text{C}$ , were accurately  
224 estimated by sampling relatively few individuals. Others, including foliar toughness and  $\delta^{15}\text{N}$ ,  
225 defied accurate estimation. Accordingly, trait identity affected the accuracy of estimation of trait  
226 means (Fig. 2). Interestingly, among-trait variation was much less in the estimation of trait  
227 variance. Similar variation among traits has been reported for temperate herbaceous communities  
228 (Pakeman & Quested 2007, Lavorel *et al.* 2008). The reasons behind this variation among traits  
229 remain unclear. One possibility is that leaf toughness and  $\delta^{15}\text{N}$  are more sensitive than other  
230 traits we analysed to extrinsic sources of error, which affect their measurement, and thereby  
231 complicate efforts to estimate their distributions. Variation in soil isotopic signatures can cause  
232 substantial variation in  $\delta^{15}\text{N}$  (Nardoto *et al.* 2014), whereas leaf toughness, in this study, was  
233 prone to measurement error arising from variation in the distance of the punch test to veins in the  
234 leaf blade (Baraloto *et al.* 2010a). Though reducing sources of extrinsic variation may increase  
235 the accuracy of estimating trait values, *we do not know why the sampling necessary to obtain*  
236 *accurate estimates varies among traits.*

237 ***Taxon-free strategies have superior performance***

238 Our simulations indicate that simple, taxon-free, sampling strategies can often  
239 characterise community-level functional trait distributions more accurately than can more-  
240 intricate strategies. Notably, strategies based on stem density or basal area of species, whether  
241 assessed locally or regionally, performed less well than did taxon-free sampling strategies for a  
242 given level of sampling intensity (Fig. 1). Abundance-based sampling also implies a greater  
243 commitment of resources to field sampling, because it must be preceded by local or regional  
244 botanical inventories (Table 1). The SIZE strategy, which entailed the random sampling of size-  
245 stratified individuals, yielded the least-biased estimates of community-level trait distributions,  
246 with CLUMP and RANDOM running closely behind. These results are consistent with Baraloto  
247 *et al.* (2010a), though that study examined only one, slightly different, abundance-based strategy.  
248 They also agree with a recent examination of sampling designs for dendrochronology studies, in  
249 which random sampling yielded the least-biased estimates of tree growth and productivity  
250 (Nehrbass-Ahles *et al.* 2014).

251 Therefore, we advocate the use of taxon-free sampling techniques for estimating  
252 community-level trait distributions in species-rich forests, though they have relatively poor  
253 performance in Alpine grasslands (Lavorel *et al.* 2008). As a practical matter, concentrating field  
254 sampling at a few **spatial** foci, as in CLUMP, is likely to simplify logistics, accounts for spatial  
255 variation in species composition, and may provide a minor improvement in accuracy over  
256 random sampling.

257 What explains the poor performance of abundance-based strategies? In diverse  
258 ecosystems, such as those studied here, even the most common species represent a small  
259 proportion of individuals. For example, the most common tree species in the forests we studied,

260 *Eschweilera coriacea* (Lecythidaceae), represented only 3.4% of all observed individuals.  
261 Across our nine forests, the ten most-common species together represent only 20% of individuals.  
262 Contrastingly, at the scale of the Amazon basin, *E. coriacea* and 226 other ‘hyperdominants’  
263 (1.4% of the approximately 16,000 species of Neotropical trees), represent 50% of all individual  
264 trees. Their broad-scale dominance is attributable to their wide spatial ranges, however, rather  
265 than to numerical abundance at local sites (ter Steege *et al.* 2013). In less species-rich plant  
266 communities, or those with a less-equitable distribution of individuals among species,  
267 abundance-based sampling strategies may be more appropriate. The accuracy of abundance-  
268 based sampling strategies would also be reduced if the most-common species did not have the  
269 most-typical trait values. This situation should be uncommon, because environmental filtering,  
270 which leads to the sorting of species along environmental gradients (Grime 2006), is strong in  
271 Neotropical forests (Kraft *et al.* 2008; Paine *et al.* 2011). In a recent study of French Guiana trees,  
272 species abundance was negatively correlated with the distinctiveness of trait values, as predicted  
273 by environmental filtering (Mouillot *et al.* 2013). [Mouillot \*et al.\*](#) found similar relationships for  
274 herbaceous plants in the Alps, and fish on coral reefs in the Pacific. Abundance-based trait  
275 sampling strategies therefore yield poor results in diverse ecosystems not because common  
276 species have atypical traits, but rather because even the common species in such species-rich  
277 ecosystems are represented by few individuals. They should be more accurate in ecosystems in  
278 which more individuals are concentrated among fewer species.

279 We reiterate the caution sounded by Baraloto *et al.* (2010a) regarding the use of database-  
280 derived trait values to estimate community-level trait distributions. They are tempting because  
281 the cost associated with extracting trait data from a database is minuscule, compared to that  
282 incurred through even low-intensity field work (Baraloto *et al.* 2010a). Moreover, global trait

283 databases such as TRY (Kattge *et al.* 2011) are essential for broad-scale comparative studies. At  
284 the relatively small scale of this study, however, the DATABASE strategy rarely yielded  
285 accurate estimations of means or variances. Sampling as few as 10% of individuals was typically  
286 sufficient to yield more accurate estimates than those obtained through DATABASE lookups.  
287 The ONE PER SP strategy (Kraft *et al.* 2008) yielded good performance in almost all cases (Fig.  
288 1). In diverse ecosystems, however, it could require sampling of half of all individuals (Pitman *et*  
289 *al.* 2001). We conclude that, despite the commitment of time and money required, there is no  
290 substitute for extensive field sampling to accurately characterize the distribution of trait values in  
291 species-rich forests.

292

#### 293 **ACKNOWLEDGEMENTS**

294 The authors declare no conflicts of interest. We thank Angela Moles and an anonymous reviewer  
295 for their insightful comments on an earlier version of the manuscript. SD was supported by The  
296 Leverhulme Trust, United Kingdom, and by CONICET and FONCyT, Argentina. The Agence  
297 National de la Recherche, France, supported this research through a grant from its Biodiversité  
298 section, as well as an “Investissement d’Avenir” grant (CEBA, ref. ANR-10-LABX-25-01).

299

#### 300 **DATA ACCESSIBILITY**

301 Functional trait data are deposited in the TRY Global Functional Trait Database. ([http://try-  
302 db.org](http://try-<br/>302 db.org)).

303

#### 304 **LITERATURE CITED**

- 305 Albert, C.H., Thuiller, W., Yoccoz, N.G., Douzet, R., Aubert, S. & Lavorel, S. (2010) A multi-  
306 trait approach reveals the structure and the relative importance of intra-vs. interspecific  
307 variability in plant traits. *Functional Ecology*, **24**, 1192–1201.
- 308 Baraloto, C., Hardy, O.J.O., Paine, C.E.T., Dexter, K.G., Cruaud, C., Dunning, L.T., Gonzalez,  
309 M.-A.A., Molino, J.-F., Sabatier, D., Savolainen, V. & Chave, J. (2012) Using functional  
310 traits and phylogenetic trees to examine the assembly of tropical tree communities. *Journal*  
311 *of Ecology*, **100**, 690–701.
- 312 Baraloto, C., Paine, C.E.T., Patiño, S., Bonal, D., Hérault, B. & Chave, J. (2010) Functional trait  
313 variation and sampling strategies in species-rich plant communities. *Functional Ecology*, **24**,  
314 208–216.
- 315 Baraloto, C., Paine, C.E.T., Poorter, L., Beauchene, J., Bonal, D., Domenach, A.-M., Hérault, B.,  
316 Patiño, S., Roggy, J.-C. & Chave, J. (2010) Decoupled leaf and stem economics in rain  
317 forest trees. *Ecology Letters*, **13**, 1338–1347.
- 318 Bates, D.M., Maechler, M., Bolker, B.M. & Walker, S. (2014) lme4: Linear mixed-effects  
319 models using Eigen and S4.
- 320 Chave, J., Coomes, D.A., Jansen, S., Lewis, S.L., Swenson, N.G. & Zanne, A.E. (2009) Towards  
321 a worldwide wood economics spectrum. *Ecology letters*, **12**, 351–66.
- 322 Conti, G. & Díaz, S. (2013) Plant functional diversity and carbon storage - an empirical test in  
323 semi-arid forest ecosystems (S Lavorel, Ed.). *Journal of Ecology*, **101**, 18–28.
- 324 Díaz, S., Hodgson, J.G., Thompson, K., Cabido, M., Cornelissen, J.H.C., Jalili, A., Montserrat-  
325 Martí, G., Grime, J., Zarrinkamar, F., Asri, Y., Band, S.R., Basconcelo, S., Castro-Díez, P.,  
326 Funes, G., Hamzehee, B., Khoshnevi, M., Pérez-Harguindeguy, N., Pérez-Rantomé, M.C.,  
327 Shirvany, A., Vendramini, F., Yazdani, S., Abbas-Azimi, R., Bogaard, A., Boustani, S.,  
328 Charles, M., Dehghan, M., de Torres-Espuny, L., Falczuk, V., Guerrero-Campo, J., Hynd,  
329 A., Jones, G., Kowsary, E., Kazemi-Saeed, F., Maestro-Martínez, M., Romo-Díez, A., Shaw,  
330 S., Siavash, B., Villar-Salvador, P. & Zak, M.R. (2004) The plant traits that drive  
331 ecosystems: Evidence from three continents. *Journal of Vegetation Science*, **15**, 295.
- 332 Díaz, S., Lavorel, S., Bello, F. de, Quétier, F., Grigulis, K. & Robson, T.M. (2007) Incorporating  
333 plant functional diversity effects in ecosystem service assessments. *Proceedings of the*  
334 *National Academy of Sciences of the United States of America*, **104**, 20684–9.
- 335 Fisichelli, N. A., Frelich, L.E. & Reich, P.B. (2014) Temperate tree expansion into adjacent  
336 boreal forest patches facilitated by warmer temperatures. *Ecography*, **37**, 152–161.
- 337 Fortunel, C., Paine, C.E.T., Fine, P.V. A., Kraft, N.J.B. & Baraloto, C. (2013) Environmental  
338 factors predict community functional composition in Amazonian forests. *Journal of Ecology*,  
339 **102**, 145–155.
- 340 Grime, J. (2006) Trait convergence and trait divergence in herbaceous plant communities:  
341 mechanisms and consequences. *Journal of Vegetation Science*, **17**, 255–260.
- 342 Kattge, J., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönsch, G., Garnier, E., Westoby,  
343 M., Reich, P.B. & Wright, I.J. (2011) TRY--a global database of plant traits. *Global Change*  
344 *Biology*, **17**, 2905–2935.
- 345 Kraft, N.J.B., Valencia, R. & Ackerly, D.D. (2008) Functional traits and niche-based tree  
346 community assembly in an Amazonian forest. *Science*, **322**, 580–582.
- 347 Lavorel, S. & Garnier, E. (2002) Predicting changes in community composition and ecosystem  
348 functioning from plant traits: revisiting the Holy Grail. *Functional Ecology*, **16**, 545–556.

- 349 Lavorel, S., Grigulis, K., McIntyre, S., Williams, N.S.G., Garden, D., Dorrough, J., Berman, S.,  
350 Quétier, F., Thébault, A. & Bonis, A. (2008) Assessing functional diversity in the field –  
351 methodology matters! *Functional Ecology*, **22**, 134–147.
- 352 McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology  
353 from functional traits. *Trends in Ecology & Evolution*, **21**, 178–185.
- 354 Mouillot, D., Bellwood, D.R., Baraloto, C., Chave, J., Galzin, R., Harmelin-Vivien, M., Kulbicki,  
355 M., Lavergne, S.S., Lavorel, S., Mouquet, N., Paine, C.E.T., Renaud, J. & Thuiller, W.  
356 (2013) Rare species support vulnerable functions in high-diversity ecosystems (GM Mace,  
357 Ed.). *PLoS Biology*, **11**, e1001569.
- 358 Nardoto, G.B., Quesada, C.A., Patiño, S., Saiz, G., Baker, T.R., Schwarz, M., Schrodt, F.,  
359 Feldpausch, T.R., Domingues, T.F., Marimon, B.S. et al. (2014) Basin-wide variations in  
360 Amazon forest nitrogen-cycling characteristics as inferred from plant and soil 15N:14N  
361 measurements. *Plant Ecology & Diversity*, **7**, 173–187.
- 362 Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R<sup>2</sup> from  
363 generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133–142.
- 364 Nehrbass-Ahles, C., Babst, F., Klesse, S., Nötzli, M., Bouriaud, O., Neukom, R., Dobbertin, M.  
365 & Frank, D. (2014) The influence of sampling design on tree-ring based quantification of  
366 forest growth. *Global Change Biology*, **20**, 2867–2885.
- 367 Paine, C.E.T., Baraloto, C., Chave, J. & Hérault, B. (2011) Functional traits of individual trees  
368 reveal ecological constraints on community assembly in tropical rain forests. *Oikos*, **120**,  
369 720–727.
- 370 Paine, C.E.T., Baraloto, C., Díaz, S. (2015) Data from The BRIDGE database. TRY Global  
371 Functional Trait Database. (<http://try-db.org>).
- 372 Paine, C.E.T., Norden, N., Chave, J., Forget, P.-M., Fortunel, C., Dexter, K.G. & Baraloto, C.  
373 (2012) Phylogenetic density dependence and environmental filtering predict seedling  
374 mortality in a tropical forest. *Ecology Letters*, **15**, 34–41.
- 375 Pakeman, R.J. & Quested, H.M. (2007) Sampling plant functional traits: What proportion of the  
376 species need to be measured? *Applied Vegetation Science*, **10**, 91–96.
- 377 Patiño, S., Lloyd, J., Paiva, R., Baker, T.R., Quesada, C.A., Mercado, L.M., Schmerler, J.,  
378 Schwarz, M., Santos, A. J.B., Aguilar, A., Czimeczik, C.I., Gallo, J., Horna, V., Hoyos, E.J.,  
379 Jiménez, E., Palomino, W., Peacock, J., Peña Cruz, A., Sarmiento, C., Sota, A., Turriago,  
380 J.D., Villanueva, B., Vitzthum, P., Alvarez Dávila, E., Arroyo, L.L., Baraloto, C., Bonal, D.,  
381 Chave, J., Costa, A.C.L., Herrera, R., Higuchi, N., Killeen, T.J., Leal, E., Luizão, F.J., Meir,  
382 P., Monteagudo, A., Neill, D.A., Núñez Vargas, P., Peñuela, M.C., Pitman, N.C.A., Priante  
383 Filho, N., Prieto, A., Panfil, S.N., Rudas, A., Salamão, R., Silva, N., Silveira, M., Soares  
384 deAlmeida, S., Torres-Lezama, A., Vásquez-Martínez, R., Vieira, I.C.G., Malhi, Y. &  
385 Phillips, O.L. (2009) Branch xylem density variations across the Amazon Basin.  
386 *Biogeosciences*, **6**, 545–568.
- 387 Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S. & Poorter, H. (2013) New handbook  
388 for standardised measurement of plant functional traits worldwide. *Australian Journal of*  
389 *Botany*, **61**, 167–234.
- 390 Pitman, N.C. A., Terborgh, J.W., Silman, M.R., Núñez Vargas, P., Neill, D.A., Cerón, C.E.,  
391 Palacios, W. A. & Aulestia, M. (2001) Dominance and distribution of tree species in upper  
392 Amazonian terra firme forests. *Ecology*, **82**, 2101–2117.
- 393 Poorter, L., Bongers, L. & Bongers, F. (2006) Architecture of 54 moist-forest tree species: traits,  
394 trade-offs, and functional groups. *Ecology*, **87**, 1289–301.

- 395 R Core Development Team. (2014) R: A Language and Environment for Statistical Computing.  
396 Scheiter, S., Langan, L. & Higgins, S. (2013) Next-generation dynamic global vegetation  
397 models: learning from community ecology. *New Phytologist*, **198**, 957–969.  
398 ter Steege, H., Pitman, N.C.A., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.-F.,  
399 Prévost, M.-F., Spichiger, R., Castellanos, H., von Hildebrand, P., Vásquez, R. & Vásquez-  
400 Martínez, R. (2006) Continental-scale patterns of canopy tree composition and function  
401 across Amazonia. *Nature*, **443**, 444–447.  
402 ter Steege, H., Pitman, N.C.A., Sabatier, D., Baraloto, C., et al. (2013) Hyperdominance in the  
403 Amazonian Tree Flora. *Science*, **342**, 1243092–1243092.  
404 Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007)  
405 Let the concept of trait be functional! *Oikos*, **116**, 882–892.  
406 Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares,  
407 J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias,  
408 J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C.H., Midgley, J.J., Navas, M.-L.,  
409 Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L.D., Pyankov, V.I.,  
410 Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. & Villar, R. (2004) The  
411 worldwide leaf economics spectrum. *Nature*, **428**, 821–827.  
412

## 413 SUPPORTING INFORMATION

414 [Additional supporting information may be found in the online version of this article.](#)

415 [Appendix S1 R script for all analyses](#)

416 Please note: Wiley Blackwell are not responsible for the content or functionality of any  
417 supporting information supplied by the authors. Any queries (other than missing material) should  
418 be directed to the corresponding author for the article.

419

420 TABLES

421 **Table 1** Trait sampling strategies investigated in this study.

Type	Name	How are individuals sampled?	Weighting*	Requirement*	Intensity
Abundance-based	BA REGION	Proportional to species basal area in the region	Overweights species that are regionally common, widespread and large-statured	Botanical inventories and d.b.h. from the region	1-50%
	BA LOCAL	Proportional to species basal area in the local plot	Overweights species that are locally common and large-statured	Botanical inventory and d.b.h. from local plot	1-50%
	BA INDIVIDUAL	Proportional to individual basal area	Overweights species with large-statured individuals, regardless of abundance	Botanical inventory and d.b.h. from local plot	1-50%
	COMMON REGION	Proportional to regional abundance	Overweights species that are common and widespread, regardless of stature	Botanical inventories from the region	1-50%
	COMMON LOCAL	Proportional to local abundance	Overweights species that are locally common, regardless of stature or regional abundance	Botanical inventory of local plot	1-50%
	SPECIES	Inversely proportional to local abundance	Overweights species that are locally rare	Botanical inventory of local plot	1-50%
Taxon-free	RANDOM	Randomly	Accurately reflects local species composition	None	1-50%
	SIZE	Stratify individuals by stature. Draw individuals randomly from strata	Overweights species with large individuals	d.b.h. from local plot, but no botanical inventory	1-50%
	CLUMP	Rank individuals by distance to three randomly located foci. Draw individuals based on their proximity to nearest focus.	Accurately reflects local relative abundance. Accounts for geographic variation in species composition	None	1-50%
Special	ONE PER SP	Draw one individual from the	Overweights species that are locally rare	Botanical inventory of	~33%

---

	plot to represent each species		local plot	
DATABASE	Draw one individual from the region to represent each species	Overweights species that are locally rare	Regional checklist of species	100%
BRIDGE	Select all individuals	Accurately reflects local species composition	Botanical inventory of local plot	100%

---

422 Notes:

423 Weighting is defined with respect to drawing individuals at random.

424 Requirement indicates field work required prior to trait sampling.

425 **FIGURES**426 ***Figure 1***

427           The probability of correctly estimating the community mean or variance in functional  
428 trait values increases with increasing sampling intensity and varies among sampling strategies  
429 and the trait under consideration. Lines show predictions from linear mixed-effect models,  
430 applied to simulation output. Larger values represent less-accurate estimates of the true mean and  
431 variance. Note that the X-axis is log-transformed. Results from the ONE PER SP, DATABASE  
432 and BRIDGE strategies are shown as points because they are only defined at a single sampling  
433 intensity (Table 1). Vertical arrows indicate that the estimation error of DATABASE and ONE  
434 PER SP exceed the plotted region for Leaf  $\delta^{15}\text{N}$ .

435

436 ***Figure 2***

437           The per cent variance in the probability of accurately estimating community mean and  
438 variance in trait values explained by sampling intensity (I), sampling strategy (S), the sampled  
439 trait (T) and their interactions. Note that the per cent variance explained by some interactions is  
440 so slight that their bars are scarcely visible.