1 **Classification:** Original Article

2

3 **On the floristic identity of Amazonian vegetation types**

4 Ary T. Oliveira-Filho[1], Kyle G. Dexter[2,3], R. Toby Pennington[3,4], Marcelo F. Simon[5],

5 Marcelo L. Bueno[6] & Danilo M. Neves[1*]

6

7 *[1]Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte 31270–*

8 *090, Brazil. [2]School of GeoSciences, University of Edinburgh, Edinburgh EH9 3JN, UK.*

9 *[3]Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, UK. [4]Department of Geography,*

10 *University of Exeter, Exeter, EX4 4RJ. [5]EMBRAPA Recursos Genéticos e Biotecnologia,*

11 *Brasília 70770-200, Brazil. [6]Unidade Universitária de Mundo Novo, Universidade Estadual*

12 *de Mato Grosso do Sul, Mundo Novo 79980-000, Brazil.*

13

14 **Correspondence:** Danilo M. Neves, *Institute of Biological Sciences, Federal University of*

15 *Minas Gerais, Belo Horizonte 31270–090, Brazil.*

16 E-mail: dneves@icb.ufmg.br

17

22

23 **ABSTRACT**

24 The Amazon forest is far from uniform, containing different forest types and even savannas,

25 but quantitative analyses of this variation are lacking. Here, we applied ordination analyses

26 to test the floristic differentiation amongst Amazonian vegetation types using data for

27 virtually all known tree species occurring in the Amazon (8,224), distributed across 1,584

28 sites. We also performed multiple regressions to assess the role of climate and substrate in

29 shaping continental-scale patterns of community composition across Amazonia. We find that

30 the traditional classification of Amazonian vegetation types is consistent with quantitative

31 patterns of tree species composition, with high elevation and the extremes of substrate-related

32 factors underpinning the floristic segregation of environmentally "marginal" vegetation types

33 and *terra firme* forests (with climatic factors being relatively unimportant). These patterns

34 hold at continental-scales, with sites of similar vegetation types showing higher similarity

35 between them regardless of geographic distance, which contrasts with the idea of large-scale

36 variation amongst geographic regions (e.g., between the Guiana Shield and southwestern

37 Amazon) representing the dominant floristic pattern in the Amazon. In contrast to other

38 tropical biomes in South America, including the *Mata Atlântica* (second largest rain forest

39 biome in the neotropics), the main floristic units in the Amazon are not geographically

40 separated, but are edaphically driven and spatially interdigitated across Amazonia. Two thirds

41 of *terra firme* tree species are restricted to this vegetation type, whilst among marginal

42 vegetation types, only white-sand forests (*campinaranas*) have a substantial proportion of

43 restricted species, with other vegetation types sharing large numbers of species.

44

45 **Keywords:** community composition, edaphic conditions, environmentally marginal habitats,

46 ordination analysis, environmental gradients, *terra firme* forests, tree species, white-sand

47 forest.

48

INTRODUCTION

50

51 The Amazon forest, which spreads across the lowlands of the Amazon, Orinoco and other

52 northern drainages of South America, is the world's largest continuous expanse of tropical

53 rain forest, with an ever increasing number of described plant species (Cardoso *et al.*, 2017;

54 ter Steege *et al.*, 2016). Since the first scientific exploration of the 18th and 19th centuries, it

55 has been clear that the region is far from a continuous and undifferentiated rain forest, as

56 there are striking contrasts among forest physiognomies and even patches of savanna

57 vegetation.

58 Two main dichotomies in vegetation types have long been established, both using

59 divisions based upon which environments are interpreted to be more marginal (Salovaara *et*

60 *al.*, 2005). One contrasts upland, *terra firme* forests growing on flood-free interfluves

61 (literally solid or firm ground; *tierra firme* in Spanish) with those growing on the seasonally

62 inundated floodplains along wide and slow flowing, larger rivers (Luize *et al.*, 2018). The

63 other dichotomy contrasts both *terra firme* and flooded forests with forest occurring on

64 pockets of highly leached deposits of podzolized hypo-dystrophic white-sand (Adeney *et al.*,

65 2016). But while the environmental differences between *terra firme*, flooded and white-sand

66 forests are somewhat striking, the lines between these forests and the other vegetation types

67 in the Amazon are not always sharp, contributing to some nomenclatural confusion (Phillips

68 *et al.*, 2003).

69 There are several additional prominent vegetation types in the Amazon, growing on

70 and around rock outcrops and coastal sands. In both edaphic situations, the vegetation shows

71 a wide array of physiognomic expressions, including forests, dwarf-forests, scrublands and

72 bushlands, often mixed in mosaics. Rock outcrops are particularly evident across the chain of

73   sandstone highlands of the Guiana Shield (often referred to as *tepuis*; Berry & Riina, 2005;

74   Huber, 1997), and on the top of the numerous inselbergs of both the Brazilian and Guiana

75   Shields (Gröger, 2000; Raghoenandan, 2000). Coastal vegetation mosaics include extensive

76   tracts of mangrove forests that run almost uninterrupted from the Brazilian island of São Luís

77   to the Orinoco Delta in Venezuela, and may penetrate inland as far as 40 km where they

78   gradually blend with flooded forests (González, 2011; Nascimento *et al.*, 2013). Away from

79   the mangroves, pockets of stabilized coastal sands bear a mosaic of vegetation types usually

80   referred to as *restingas* and *matas de maré* in Brazil (Silva *et al.*, 2010).

81   Although virtually all the environmentally marginal vegetation types can include open

82   physiognomies that may resemble savannas, the Amazon is also home to savannas *sensu*

83   *stricto*, i.e. those associated with the existence of a dry season lasting for at least three

84   months, and a flammable grass ground layer that may trigger fire outbreaks (Huber, 1997).

85   Most of these savannas experience some form of waterlogging during the rainy season (see

86   Pennington *et al.*, 2006).

87   From previous studies, we know that tree species distribution patterns do match some

88   of these pre-defined vegetation types at local and regional scales (10 to 100,000 km^2; e.g.,

89   Draper *et al.*, 2018; Draper *et al.*, 2019; Duivenvoorden, 1995; Guitet *et al.*, 2015; Higgins *et*

90   *al.*, 2011; Pitman *et al.*, 2008; Scudeller, 2018; Stropp *et al.*, 2011; ter Steege *et al.*, 2000)

91   but, so far, this has not been scrutinized for the Amazon region as a whole and all its main

92   vegetation types. A common conclusion stemming from the few Amazon-wide floristic

93   studies is that distinct vegetation types in one region tend to resemble one another more

94   closely than they do the same vegetation types in other regions (Silva-Souza & Souza, 2020;

95   ter Steege *et al.*, 2006; Terborgh & Andresen, 1998). Such conclusions have led to a view of

96   floristic regionalization that has neglected the different vegetation types of Amazonia. For

97   example, there is a consistent west to east gradient in tree community composition, congruent

98   with an Amazon-wide variation in soil fertility and drought (soils in the eastern Amazon are

99   poorer and climate is drier; ter Steege *et al.*, 2006). These continental-scale analyses have

100   either been conducted at coarser taxonomic scales - at the family (Terborgh & Andresen,

101   1998) and genus-level (ter Steege *et al.*, 2006) - or have lumped taxa from distinct vegetation

102   types into large geographic 'grid cells' (e.g., 4° x 6° in ter Steege *et al.*, 2006; up to 20 km

103   distance in Silva-Souza & Souza, 2020). Here, we bring together the most comprehensive,

104   species-level dataset to date on the composition of tree communities across the entire

105   Amazon basin, where individual communities have been assigned *a priori* to one of the

106   predominant vegetation types in the Amazon.

107   Our objectives are three-fold. Firstly, we test the floristic differentiation of nine

108   vegetation types, following the classification system proposed by Oliveira-Filho (2015). We

109   predict that by using comprehensive, species-level tree community surveys, most (if not all)

110   vegetation types will show an Amazon-wide compositional consistency.

111   Secondly, we test whether variation in edaphic and climatic conditions controls the

112   floristic differentiation between *terra firme* forest and other vegetation types, with marginal

113   vegetation types being placed in environments sometimes interpreted to be more stressful

114   (Salovaara *et al.*, 2005). We predict that the floristic segregation of lowland vegetation types

115   is primarily associated with edaphic factors (e.g., rockiness, sandiness, salinity, soil-

116   waterlogging), with climate being only important in segregating highland vegetation types

117   (e.g., montane forest, *tepuis*) from all others.

118   Finally, to give context to our results and to explore the floristic distinctiveness of

119   vegetation types, we also examine patterns of species shared amongst these vegetation types

120   and the proportion of species restricted to individual vegetation types.

121

122   **MATERIALS AND METHODS**

123

1. Study area

125 The Amazon forest, as circumscribed here (see outline in Figure 1), includes most of the

126 Amazon and Orinoco river basins (excluding the Andean headwaters of some rivers and the

127 mid-Orinoco Llanos), and the North Atlantic coastal river basins between the states of Delta

128 Amacuro, in Venezuela, and Maranhão, in Brazil. With regards to elevation, a maximum

129 altitude of 1,100m was established on the Andean flanks to exclude the complex and

130 extensive vegetation and environment gradients associated with the massive mountain chain.

131 No altitudinal limit was established, however, for the highlands of the Guiana Shield, which

132 are entirely embedded in the Amazon Province. This is a controversial issue in the sense that

133 some authors consider the Guiana Highlands as a separate biogeographic province (e.g.,

134 Cabrera & Willink, 1980; Cardoso *et al.*, 2017), supported by the high number of endemic

135 species, many of which are restricted to particular *tepuis* or highlands (Berry & Riina, 2005).

136 We based our decision on the following facts: (a) unlike the Andes, which make up a natural

137 limit, the Guiana Highlands are encircled and pervaded by Amazonian lowlands; (b) the

138 highest altitudes reached by the Guiana Highlands (2,500-3,000m) are modest compared to

139 those of the Andes; and (c) *tepuis* and highlands also share a considerable number of species

140 with lowland Amazonian vegetation types (Steyermark *et al.*, 1995-2005).

141

2. Nomenclature

143 The white-sand vegetation complex is particularly thorny when it comes to nomenclature.

144 The main reason for this is the remarkable variation in physiognomy, which ranges from

145 grass/shrublands to forests with slender-trunked trees and more open canopies compared to

146 those of adjacent *terra firme* forests growing on more clayey soils (Adeney *et al.*, 2016).

147 Throughout the Amazon, various local terms are also used to designate both the whole white-

148   sand complex and its physiognomic expressions, e.g. *bana*, *caatinga amazônica*, *campina*,

149   *campinarana*, *varillal* and *chamizal* (Demarchi *et al.*, 2018; Fine *et al.*, 2010; García-

150   Villacorta *et al.*, 2016; Stropp *et al.*, 2011). Following Daly *et al.* (2016), we here adopt

151   *campinarana*, because of its official use in Brazil (IBGE, 2012) and because it embraces the

152   whole array of physiognomies growing on podzolized sands liable to ground water saturation,

153   but conveniently excludes white-sand floodplain forests, which we distinguish in this paper.

154   Nomenclature for vegetation occurring on seasonal floodplains is less complex. We use

155   the prevailing nomenclature for two main seasonally flooded vegetation types in the Amazon,

156   distinguishing *igapó* and *várzea* forests, depending on the types of rivers along which they

157   occur (see Junk *et al.*, 2011; Kubitzki, 1987; Prance, 1979). *Várzea* forests are found along

158   rivers carrying copious quantities of sediments (and nutrients), mostly brought from the

159   Andes, with variation in the amount of clay resulting in waters that are many shades of

160   brown. Confusingly, these rivers are often called white-water rivers (*ríos de agua blanca,*

161   *rios de água branca*). In contrast, *igapó* forests are found along rivers with small amounts of

162   suspended mineral particles, which are called black or clear-water rivers (*ríos de agua negra*

163   *o clara, rios de água negra ou clara*). These rivers drain basins where white-sands or other

164   highly leached soils prevail (e.g., flowing from the Brazilian and Guiana Shields) and can

165   carry vast loads of humic acid colloids resulting from the arrested litter decay in these hypo-

166   dystrophic soils. A similar process takes place in black-water oxbow lakes severed from

167   white-water rivers as well as in narrower upstream floodplains throughout the basin. The

168   dichotomy of *várzea* and *igapó* falls short when it comes to rivers with "mixed" waters, and

169   rivers with temporal and spatial variations in suspended particles, of which the Casiquiare

170   Channel in Venezuela is an example. In both *igapós* and *várzeas*, vegetation structure varies,

171   from tall forests to floodplains with more open formations depending on local flooding

172   dynamics and related processes of either erosion or sedimentation (Kalliola *et al.*, 1992;

173   Luize *et al.*, 2018; Salo *et al.*, 1986; Worbes *et al.*, 1992). The timing and duration of

174   flooding in these forests can be variable, from once every few decades in rivers close to the

175   Andes (e.g. on the Manu River in Peru, pers. comm. John Terborgh) to multiple months

176   annually for the iconic *várzeas* and *igapós* along major rivers such as the Amazon and the

177   Rio Negro.

178

179   **3. Dataset**

180   We extracted the dataset from the NeoTropTree (NTT) database

181   (http://www.neotroptree.info/), which consists of tree species checklists (trees defined here as

182   freely standing woody plants >3 m in height) compiled for geo-referenced sites, from

183   southern Florida (U.S.A.) and Mexico to Patagonia in Argentina and Chile. NTT currently

184   holds 7,485 sites/checklists, 20,562 woody plant species and 1,206,314 occurrence records. A

185   site/checklist in NTT is defined by a single vegetation type, following the classification

186   system proposed by Oliveira-Filho (2015), contained in a circular area with a 10 km

187   diameter. Where two or more vegetation types co-occur in the area, there can be multiple

188   geographically overlapping sites in the NTT database.

189   The data were originally compiled from an extensive survey of published and

190   unpublished (e.g. PhD theses) literature, particularly those on woody plant community

191   surveys and floristic inventories. Additional occurrence records obtained from both major

192   herbaria and taxonomic monographs have been added to the checklists when they were

193   collected within the 10-km diameter of the original NTT site, and within the same vegetation

194   type. NTT does not include sites with an indication of high anthropogenic disturbance nor

195   those with low species richness, because this is often due to low sampling/collecting efforts,

196   which results in poor descriptive power. Thus, secondary forests, which might be considered

197   a distinct vegetation type, are not included in our study. Lowest species richness in the

198   Amazon dataset ranged from 20 species in savanna *s.s.* and *campinarana* to 100 in *terra*

199   *firme* forest, while plot size (in floristic surveys derived from plot data) ranged from 1 to 5

200   ha.

201   All species and their occurrence records were checked for taxonomic circumscriptions

202   and geographical distributions as accepted by the teams of specialists responsible for the

203   online projects *Flora do Brasil*, *Catalogue of the Vascular Plants of Ecuador*, *Peru Checklist*,

204   *Bolivia Catalogue* (available at http://floradobrasil.jbrj.gov.br/,

205   http://www.tropicos.org/Project/CE/, http://www.tropicos.org/Project/PEC, and

206   http://www.tropicos.org/Project/BC/, respectively) and published floras (Bernal *et al.*, 2016;

207   Boggan *et al.*, 1997; Cardoso *et al.*, 2017; Steyermark *et al.*, 1995-2005). We eventually

208   eliminated records for 111 species due to synonymy (59), invalid or dubious names (7),

209   incorrect growth habit (15) and incorrect distribution (30).

210   The final dataset contained presence/absence data for 8,224 tree species across 1,584

211   sites, with a total of 364,965 presences. Sites derived exclusively from herbarium data

212   represented 41% of the full matrix (654 sites), The dataset also included 23 environmental

213   variables (30 arc-sec resolution) for all its sites, derived from multiple sources. Procedures

214   and protocols concerning variables' sources and extraction are thoroughly detailed at

215   http://www.neotroptree.info/.

216   We adopted the vegetation descriptors provided by NTT and based on Oliveira-Filho

217   (2015) to classify the sites into nine vegetation types: 776 *terra firme* forests, 171

218   *campinaranas*, 291 *várzeas*, 176 *igapós*, 55 rock outcrops, 36 *tepuis,* 29 coastal mosaics, 28

219   savannas *sensu stricto* (hereafter savanna *s.s.*) and 22 montane forests (Figure 1; Table S1).

220   All sites classified as *tepuis* and montane forests occur above 1,100m of altitude (see Study

221   Area), with *tepuis* differing from montane forests in their rocky soils and dwarfish

222   physiognomy. The map in Figure 1 was designed using the packages maptools (Bivand &

223   Lewin-Koh, 2017) and raster (Hijmans, 2016) in R Statistical Environment (R Core Team,

224   2018).

225   The NTT database also includes environmental variables for all its sites, derived from

226   multiple sources (at a 30 arc-second resolution). Altitude at the NTT site centre was used as

227   an integrative environmental variable. Variables representing average climate (mean annual

228   precipitation and temperature) as well as climate extremes (e.g., precipitation in driest month)

229   and seasonality (e.g., precipitation seasonality) were obtained from WorldClim 1.4 data

230   layers (Hijmans *et al.*, 2005). Frost frequency (days) and cloud interception (mm) were

231   obtained from interpolating known values as response variables (data obtained from 135 and

232   57 Brazilian Meteorological Stations measuring frost frequency and cloud interception,

233   respectively) with elevation, latitude and the WorldClim layers as predicting variables. Soil

234   coarseness (% sand) and soil fertility (% base saturation) and surface rockiness (% exposed

235   rock) were obtained from the Harmonized World Soil Database v 1.2 (available at

236   http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-

237   database-v12/en/) and ranked afterwards by mid-class percentage. The use of classes was

238   adopted because high local soil heterogeneity can make raw figures unrepresentative. Soil

239   Water Storage capacity (%) was obtained from the International Soil Moisture Network

240   (available at www.ipf.tuwien.ac.at/insitu/).

241

242   **4. Analyses of community composition**

243   We first explored the patterns of floristic differentiation amongst previously defined

244   vegetation types by performing non-metric multidimensional scaling (NMDS; McCune &

245   Grace, 2002), and tested its overall significance by applying an analysis of similarities

246   (ANOSIM; Clarke, 1993). Beforehand, we excluded 832 singletons (species found at a single

247   site), as they commonly increase the noise in ordination analyses without contributing

248 information (Lepš & Šmilauer, 2003), and then computed pairwise compositional distances

249 between all sites using Simpson distance as the dissimilarity metric (Simpson, 1960), which

250 describes community turnover without the influence of richness gradients (Baselga, 2010).

251   We used the vegetation types confirmed in the ordination analysis to produce sets of

252 diagnostic species based on their coefficient of fidelity (*phi*; Tichý and Chytrý, 2006). An

253 advantage of this coefficient is that they can take negative values, which expresses the fact

254 that a species tends to "avoid" a particular habitat and its environmental conditions (De

255 Cáceres *et al.*, 2008; De Cáceres and Legendre, 2009). In this study, diagnostic species

256 represent those statistically associated with one or more vegetation types so that their

257 presence in species lists may be a strong indicator of the vegetation types themselves.

258 Significance of *phi* was obtained via Monte Carlo permutations (999). Species' *phi* for each

259 vegetation type are provided as Supplementary Information (see Table S2).

260   We then used the major axes of compositional variation summarized by the ordination

261 analysis to test whether the observed patterns of floristic differentiation in the Amazon are

262 underpinned by increasingly stressful environmental conditions, segregating *terra firme*

263 forests from environmentally marginal vegetation types. First, we selected a subset of

264 significant environmental variables for each of the major NMDS axes through an AIC-based

265 forward selection method for generalized linear models, and then performed an additional

266 and progressive elimination of collinear variables based on their variance inflation factor

267 (VIF), informed by their ecological relevance, until maintaining only those with $VIF < 4$

268 (Quinn & Keough, 2002). We tested the significance of the selected environmental variables

269 by applying ANOVA permutation tests (999 permutations). We explored the results visually

270 by fitting the values of the most important environmental variables in ordination space

271 (NMDS). The variable selection, VIF and NMDS analyses were conducted using the vegan

272 (Oksanen *et al.*, 2016) and usdm (Naimi *et al.*, 2014) packages in the R Statistical

273 Environment.

274   Finally, we used a chord diagram to assess the patterns of compositional overlap

275 amongst Amazonian vegetation types. The chord diagram was designed using the D^3 Java

276 Environment (Bostock *et al.*, 2011; custom codes available at https://bl.ocks.org/nbremer).

277

278 **RESULTS**

279

280 The distribution of the sites in ordination space yielded by NMDS (K = 2; stress = 0.17;

281 Figure 2) largely segregated the previously defined Amazonian vegetation types (ANOSIM *R*

282 = 0.85; *P* = 0.001). We found a negligible decrease in stress values by adding a third NMDS

283 axis, and a high correlation between the distances summarized by the first two axes and the

284 full distance matrix (Pearson's *r* = 0.83). Thus, we focused subsequent analyses on the two-

285 dimensional ordination space, and the results are detailed below.

286   Axis 1 places both *terra firme* forests and *campinaranas* at intermediate scores, and is

287 congruent with two gradients: the first towards seasonally flooded forests (*várzeas* and

288 *igapós*), placed at one extreme, and the second towards montane forests and open formations

289 (coastal mosaic, rock outcrops, savanna and *tepuis*), at the other extreme. Axis 2 segregated

290 *várzeas*, *tepuis*, *terra firme* and montane forests at one extreme, and *igapós*, *campinaranas*,

291 and the remaining open formations at the other. It is worth noting that *campinaranas* seem to

292 be closer to savannas than to *terra firme* forests along this axis. In addition, the differentiation

293 between *terra firme* forest, savanna and the coastal mosaic is more nuanced, and suggests a

294 forest-to-savanna gradient. These patterns are robust to excluding sites (checklists) compiled

295 exclusively from herbarium data (654 sites; Figure S1).

296   The furthest extremes of substrate-related variables lead to distinct, environmentally

297   marginal vegetation types (Figure 2). The selected climatic and edaphic predictors account

298   for 72% and 62% of the variation in community composition summarized by the first two

299   NMDS axes, respectively (Table 1). An increase in sandiness was congruent with the floristic

300   differentiation of *campinaranas* from all other vegetation types, while an increase in soil

301   water storage capacity (a proxy of seasonal soil-waterlogging) was associated with the

302   floristic differentiation between seasonally flooded forests (*várzeas* and *igapós*) and all other

303   vegetation types. The somewhat nuanced differentiation between the two seasonally flooded

304   vegetation types is congruent with decreasing soil fertility from *várzeas* to *igapós*.

305   Precipitation seasonality was associated with the floristic differentiation of coastal mosaics

306   and savanna from all other vegetation types, with the segregation between these two being

307   associated with higher soil sandiness in coastal mosaics. High surface rockiness (a proxy of

308   soil water deficit) was congruent with the floristic segregation of forests associated with

309   *tepuis* and rock outcrops from all other vegetation types. *Tepuis* are also associated with

310   higher cloud interception, thus reflecting lower water deficit when compared to rock

311   outcrops. Finally, both *tepuis* and montane forests are found under lower mean annual

312   temperature, with intermediate conditions of cloud interception segregating montane forests

313   from *tepuis* (high cloud interception) and lowland *terra firme* forests (low cloud

314   interception).

315   There are a considerable number of species restricted to *terra firme* forests in our

316   dataset (4,424 species), which far surpasses the number of species *terra firme* shares with

317   other vegetation types (2,032 species; Figure 3). There is also a high proportion of species in

318   *campinaranas* that are restricted to that vegetation type (42%). In contrast, the other seven

319   vegetation types have a low proportion of species restricted to them, ranging from 6% in

320   coastal mosaics and savanna to 25% in *tepuis*. Species shared between *terra firme* forest (the

321 largest species pool) and other marginal vegetation types are high. Among marginal

322 vegetation types, the number of species shared ranges from 18, between *igapó* and *tepui*, to

323 655 shared between *igapó* and *várzeas* (Figure 3). These results indicate that, apart from *terra*

324 *firme* forests, most of the tree flora of Amazonian vegetation types are shared among two or

325 more vegetation types, with their community compositions, which are distinct (Figure 2),

326 representing unique combinations of the Amazonian species pool.

327

328 **DISCUSSION**

329

330 **1. Continental-scale patterns**

331 The composition of the tree flora across the Amazon region shows variation congruent with

332 traditional vegetation classifications. The most species rich and geographically widespread

333 vegetation type is *terra firme* forest, while marginal vegetation types, such as *campinarana*,

334 savannas, *igapó* or *várzea*, diverge in species composition along distinct environmental

335 gradients. These marginal vegetation types house many tree species not found in *terra firme*

336 forest, yet surprising numbers of them are shared amongst the different marginal vegetation

337 types themselves, for example between *campinaranas* and rock outcrops.

338   The marginal vegetation types are placed at extreme values of the significant

339 environmental gradients, potentially indicating eco-physiological stress, and our results

340 highlighted that substrate, not climate, is the most important environmental driver controlling

341 the major axes of composition in Amazonian tree communities. Different from other forest

342 biomes in South America, where variation in temperature and water availability are clearly

343 the most important factors controlling continental-scale patterns of tree community

344 composition (e.g., in seasonally dry tropical forests (Neves *et al.*, 2015), or in the *Mata*

14

345   *Atlântica* (Neves *et al.*, 2017)), climatic conditions are relatively unimportant in Amazonia

346   (but see discussion for montane forests and *tepuis*).

347   Moreover, because these edaphic gradients are consistently important in segregating

348   Amazonian vegetation types from local to continental scales, our results run counter to

349   previous findings which have suggested that tree community composition in the Amazon is

350   primarily driven by Amazon-wide gradients in environmental conditions (e.g., precipitation

351   seasonality, soil fertility; Silva-Souza & Souza, 2020; ter Steege *et al.*, 2006). Previous

352   Amazon-wide studies analyzed tree species composition data without separating or

353   considering the different Amazonian vegetation types. If composition is summarized within

354   geographic grid cells (e.g., ter Steege *et al.*, 2006; Silva-Souza & Souza, 2020), then a given

355   grid cell may take on the compositional identity of the dominant vegetation type in the grid

356   cell, and if there are geographic gradients in the prevalence of vegetation types, the

357   geographic grid cell approach may lead to geographically-driven results, which mask

358   vegetation heterogeneity within grid cells. Our approach ensured that every sample unit (i.e.

359   site or community) represents only a single vegetation type, which is likely why we find a

360   clearer signal for vegetation type than for geography in our results. That these vegetation

361   types are floristically coherent across the Amazon basin also suggests that dispersal amongst

362   areas of the same vegetation type is not particularly limited by geographic distance, in

363   agreement with a recent study of several Amazonian tree genera (Dexter *et al.*, 2017).

364   Below we delve into the main floristic patterns observed in our results to discuss the

365   compositional identity and environmental distinctiveness of Amazonian vegetation types.

366   Because the dataset used in this study does not include sites with a high indication of

367   anthropogenic disturbance, we stress that analyses including community inventories (e.g.,

368   floristic checklists, plot data) from recently degraded areas, such as early-stage secondary

369   forests, may reveal additional vegetation types.

370

2. Forest types

Flooded forests share a similar environmental condition driving their compositional

distinction from *terra firme* forests: seasonal flooding, potentially combined with soil

waterlogging during the low water season. This, however, does not lead to homogeneous

stands of flooded forests throughout the Amazon and one of their main variations was

captured here: the floristic, edaphic and distributional differentiation of *igapós* and *várzeas*.

Várzea forests are more evenly distributed across major river basins in Amazonia, while most

igapó forests are concentrated in the Rio Negro and upper Orinoco River Basins where the

substrate is of highly leached and impoverished white-sands. Nonetheless, *igapó* forests are

also found in other Amazonian regions under similar edaphic conditions (Montero *et al.*,

2014; Wittman *et al.*, 2010). In addition, both types of flooded forests have species restricted

to them in our dataset (160 species restricted to *várzeas*, and 168 to *igapós*), though the

largest proportion of their species composition is either shared between them or with *terra*

firme forests (Figure 3; Scudeller, 2018).

The tree flora of montane forests in Amazonia is compositionally coherent with the

main floristic patterns described for Neotropical montane flora in general, such as the

presence of some genera that are rare to absent in the lowland flora, including *Bonnetia*,

Brunellia, *Drimys*, *Hedyosmum*, *Ilex*, *Laplacea*, *Meriania*, *Podocarpus*, *Symplocos* and

Weinmannia (Webster, 1995). The overall lack of these taxa in lowland Amazonia is likely

driven by temperature, an important environmental factor driving floristic differentiation

between montane and *terra firme* forests in our dataset. Nonetheless, variation in temperature

across the range and location of elevations sampled in our study is not large, thus explaining

the high proportion of tree species shared between montane forests and other Amazonian

habitats (88%), and supporting the claim that these forests should be treated as Amazonian

395 (contrasting with views in Cardoso *et al.*, 2017; Cabrera & Willink, 1980). *Igapó* forests, for

396 instance, share a lower proportion of tree species with other Amazonian habitats (82%), yet

397 *igapós* are consistently treated as Amazonian (Wittmann *et al.*, 2010). The fact is that many

398 lowland *terra firme* species do extend their distribution towards high altitudes (1,100-

399 3,000m), such as *Annona symphyocarpa*, *Coussapoa crassivenosa*, *Cyathea bipinnatifida*,

400 *Cyathea macrosora*, *Elaeoluma nuda*, *Hieronyma oblonga*, *Miconia dodecandra*, *Miconia*

401 *pseudocapsularis*, *Miconia punctata*, *Mollinedia ovata*, *Nectandra reticulata* and *Quiina*

402 *cruegeriana*, to cite a few diagnostic species of both *terra firme* and montane forests (Table

403 S2).

404   The scarcity of mineral nutrients in white-sand environments is probably the leading

405 environmental distinction of *campinaranas*, and plant species are known to have acquired

406 morphological, physiological and mutualistic traits to maximize both nutrient capture and

407 retention (Adeney *et al.*, 2016). This specialized flora explains much of the differentiation of

408 *campinaranas* from other vegetation types, which is evident in the high proportion of

409 restricted species (42%; Figure 3). This is almost twice the proportion of 23% of endemics in

410 western Amazonian *campinaranas* found by Garcia-Villacorta *et al.* (2016), but this is

411 probably explained by the fact that those authors worked with the whole spectrum of growth

412 habits, and not only trees, considered all available herbarium voucher data (not just those

413 collected near NTT sites as done here) and concentrated their efforts only in western

414 Amazonia.

415

416 **3. Open formations**

417 The coastal sand deposits along the Atlantic shores, covered by a mosaic of mangroves and

418 sandy beaches, represent another Amazonian vegetation type associated with white-sand

419 substrates. However, soils in these coastal mosaics are more fertile than in *campinaranas*, and

420   they are mostly found in the eastern Amazon, where precipitation seasonality is relatively

421   more pronounced. Nonetheless, these white-sand, seasonally dry coastal environments are not

422   too extensive (Cremers & Hoff, 2003; Silva *et al.*, 2010; González, 2011), nor do they seem

423   to be restrictive floristically, as 94% of species in coastal mosaics are also found in other

424   vegetation types. Accordingly, they have one of the lowest proportions of restricted tree

425   species in Amazonian habitats – only 15 tree species are restricted to coastal sand deposits in

426   this analysis, nine of which are typical of mangroves.

427   Water deficit intervals, mediated by climate, substrate or both, drives tree community

428   differentiation in two other environments: savannas and rock outcrops. Most savannas are

429   found where the dry season is longest in the Amazon, and where fire outbreaks may occur in

430   the dry season. Interestingly, many of these savannas are hyper-seasonal (*sensu* Sarmiento,

431   1983), in that they also face some form of water excess in the rainy season, mostly due to soil

432   waterlogging either caused by poor drainage or floods, as in the Bolivian Llanos de Moxos

433   and in the Brazilian estuarine island of Marajó. Nevertheless, there are also non-hyper-

434   seasonal savannas, particularly on hills with shallow soils in the Brazilian state of Pará, where

435   the flora shares a great number of species with that of the Cerrado savannas in Central Brazil

436   (Devecchi, *et al.*, 2020). In fact, the tree flora of most Amazonian savannas does show some

437   floristic affinity with the savannas of either the Brazilian Cerrados or the Venezuelan Llanos

438   (Buzatti *et al.*, 2018; Ratter *et al.*, 2006; see also Devecchi, *et al.* 2020 for comparisons

439   between all plant life-forms).

440   Rock outcrops are another common feature in the Amazon that may experience local

441   water shortage, even in everwet areas, because rainwater is promptly drained from the

442   substrate. Rock outcrops are particularly common across inselbergs on both the Guiana and

443   Brazilian Crystalline Shields, where they host tree species not found elsewhere in the

444   Amazon (Gröger, 2000; Raghoenandan, 2000). The xeric nature of rock outcrops is

445 confirmed by the disjunct occurrences of a considerable number of species that are also

446 typical of seasonally dry deciduous forests outside of the Amazon, such as *Aspidosperma*

447 *cuspa*, *Brasiliopuntia brasiliensis*, *Bursera simaruba*, *Cereus hexagonus*, *Guapira cuspidata*,

448 *Senegalia riparia* and *Vachellia farnesiana*. Apart from rock outcrops, another substrate

449 related to periods of water deficit in the Amazon is the hardened surface of some mudflats,

450 such as those found on the Guyanese coastal plains and on the meanders of the Lower

451 Amazon and mid-Marañon rivers. This condition probably explains the occurrence of a

452 typical savanna vegetation and flora there, even under year-round ever-wet climates.

453 The *tepuis* are formed of the steep slopes and plateaus of the massive Paleozoic

454 sandstones atop the Guiana Shield. Environmental factors along altitudinal gradients are

455 hardly ever easily summarized by one variable, and include factors such as increasing cloud

456 interception (an important factor in our models) and declining temperatures towards higher

457 altitudes. Additional sources of environmental heterogeneity include slope, aspect, and

458 surface rockiness, and there is usually a local combination of stressful factors at play. In

459 general, scrubs and savannas prevail on the shallow soils and bare rocks at the summits of

460 *tepuis* (Huber, 2005), where cloud interception is the major source of water, and are replaced

461 by montane forests in colder environments with deeper soils.

462 It is important to bear in mind that, different from forest types, the open formations in

463 our analyses do not represent a tree-dominated habitat. Therefore, our comparisons are based

464 on the few tree species that occur in these plant communities, while herbs, forbs and shrubs

465 are not included. If data for the whole plant community were available, we would expect

466 even higher dissimilarities between these open habitats and other tree-dominated habitats,

467 with increased number of endemic species in the former.

468

469 **4. Conclusions**

470   Our findings show that the traditional classification of Amazonian vegetation formations is

471   consistent with quantitative patterns of tree species distribution. We also demonstrate how the

472   *terra firme* forest is the core vegetation type from which the eight marginal habitats

473   differentiate floristically in a manner consistent with more extreme environmental conditions.

474   These patterns, which have been previously described at a regional scale, are documented

475   here for the first time across the entire Amazon Basin.

476   In addition, we show that a large proportion of tree species found in the eight marginal

477   vegetation types are shared amongst each other and with *terra firme* forests. In fact, apart

478   from *terra firme* forests and *campinaranas*, there is a small percentage of tree species

479   restricted to a single vegetation type in the Amazon. Nonetheless, if future conservation

480   strategies aim to protect the full set of tree species in the Amazon, they must consider the

481   identity and distribution of the multiple vegetation types there, as well as their current status

482   of conservation. Many of the localities in our analyses may have been impacted by the recent

483   increase in deforestation and forest fires in the Amazon, especially those found across the

484   south and eastern borders of the Brazilian Amazon – a region known as the "arc of

485   deforestation" (Soares-Filho *et al.*, 2006).

486

487 **Table 1.** Significant climatic and edaphic predictors large-scale gradients of tree community

488 composition in the Amazon. Values under NMDS1 and NMDS2 represent the coefficients of

489 determination (adjusted R^2, and their respective *P*-values) of generalized linear models

490 (GLMs) between the first two axes of a Non-metric Multi-Dimensional Scaling and

491 environmental variables. Values in the last row represent coefficients of determination of

492 GLM-based multiple regressions between each NMDS axis and all significant variables. VIF

493 = variation inflation factor, as a measure of collinearity between all variables in the analyses

494 (variables were progressively eliminated until VIF < 4).

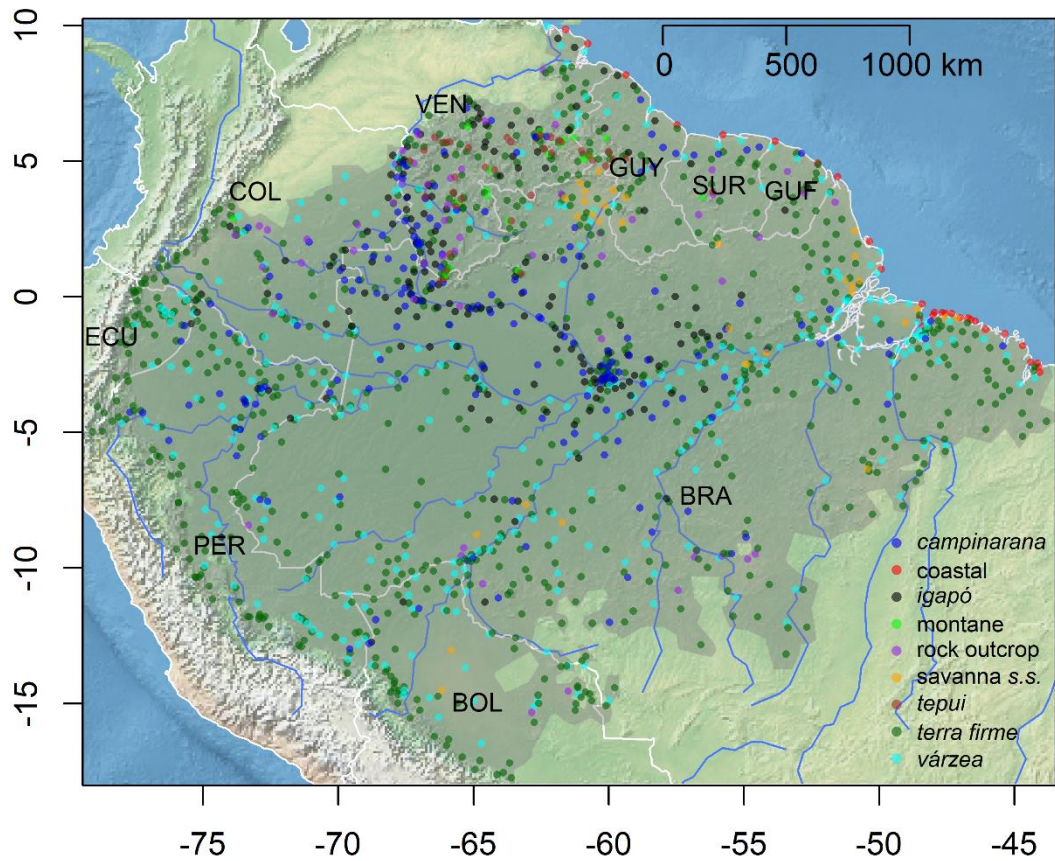|	NMDS1	*P*	NMDS2	*P*	VIF
Cloud Interception	0.224	< 0.0001	0.059	< 0.0001	2.398
Mean Annual Temperature	0.228	< 0.0001	0.225	< 0.0001	2.812
Precipitation Seasonality	0.041	< 0.0001	-	-	1.046
Sandiness	0.051	< 0.0001	0.294	< 0.0001	1.990
Soil Fertility	-	-	0.207	< 0.0001	2.087
Soil Water Storage capacity	0.656	< 0.0001	0.013	< 0.0001	3.095
Surface Rockiness	0.327	< 0.0001	0.019	< 0.0001	2.534
All variables	0.723	< 0.0001	0.623	< 0.0001	-

495

496 **FIGURE LEGENDS**

497 **Figure 1.** Distribution of the 1,584 Amazonian sites used in the analyses with their *a priori*

498 classification into nine vegetation types. Blue and white contours illustrate major rivers and

499 national borders, respectively. Our delimitation of the Amazon is outlined in a darker, gray-

500 green colour.

501 **Figure 2.** Ordination of 1,584 tree communities in the Amazon inferred from non-metric

502 multidimensional scaling of their species composition. Colors indicate the *a priori*

503 classification into nine main vegetation types, and darker shades in each color indicate

504 overlapping circles (i.e., two or more sites show high similarity in species composition).

505 PrecSeas = precipitation seasonality; CloudItcp = cloud interception; SoilFert = soil fertility;

506 SWS = soil water storage capacity; TempAnn = mean annual temperature.
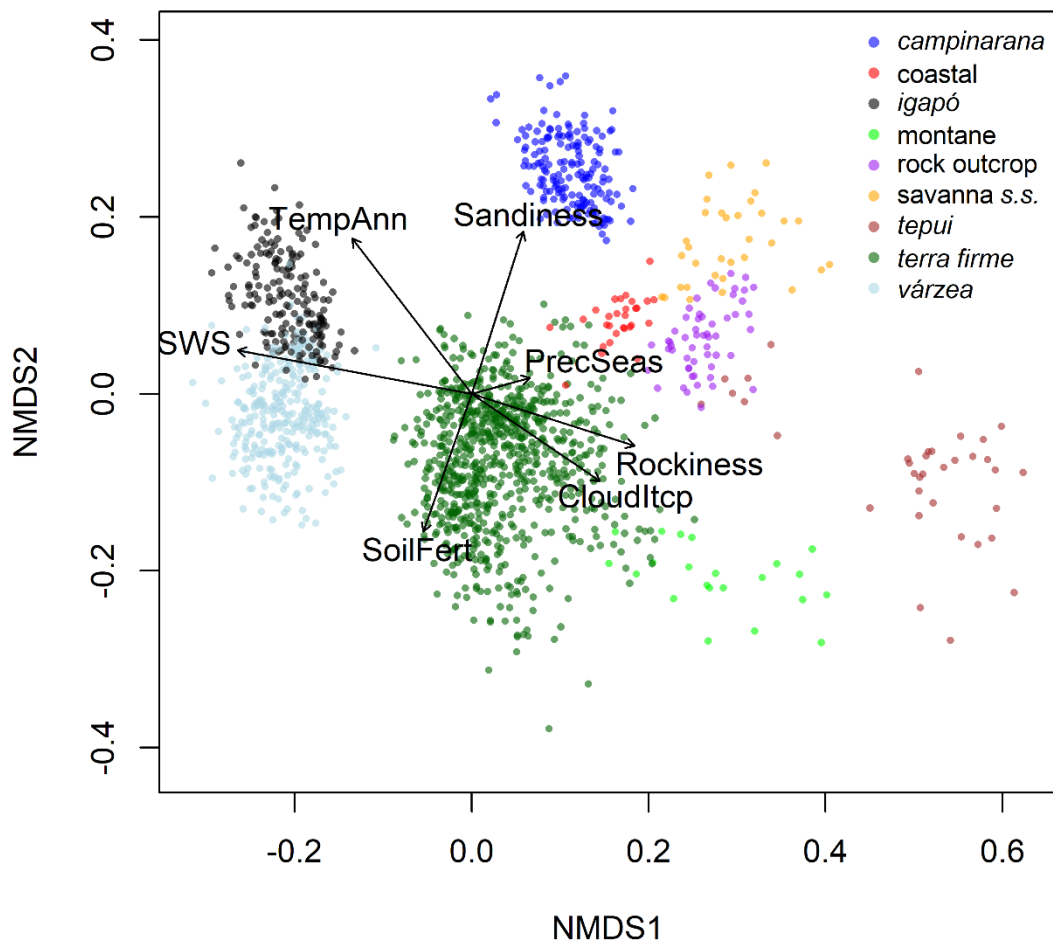
507 **Figure 3.** Overlap in tree species composition amongst Amazonian vegetation types. Values

508 in white express the number of species that are shared amongst vegetation types or restricted

509 to a given vegetation type, respectively. Chord width is proportional to the number of shared

510 species.

511

512

Figure 1. Distribution of the 1,584 Amazonian sites used in the analyses with their *a priori*

classification into nine vegetation types. Blue and white contours illustrate major rivers and

national borders, respectively. Our delimitation of the Amazon is outlined in a darker, gray-

green colour.

Figure 2. Ordination of 1,584 tree communities in the Amazon inferred from non-metric

multidimensional scaling of their species composition. Colors indicate the *a priori*

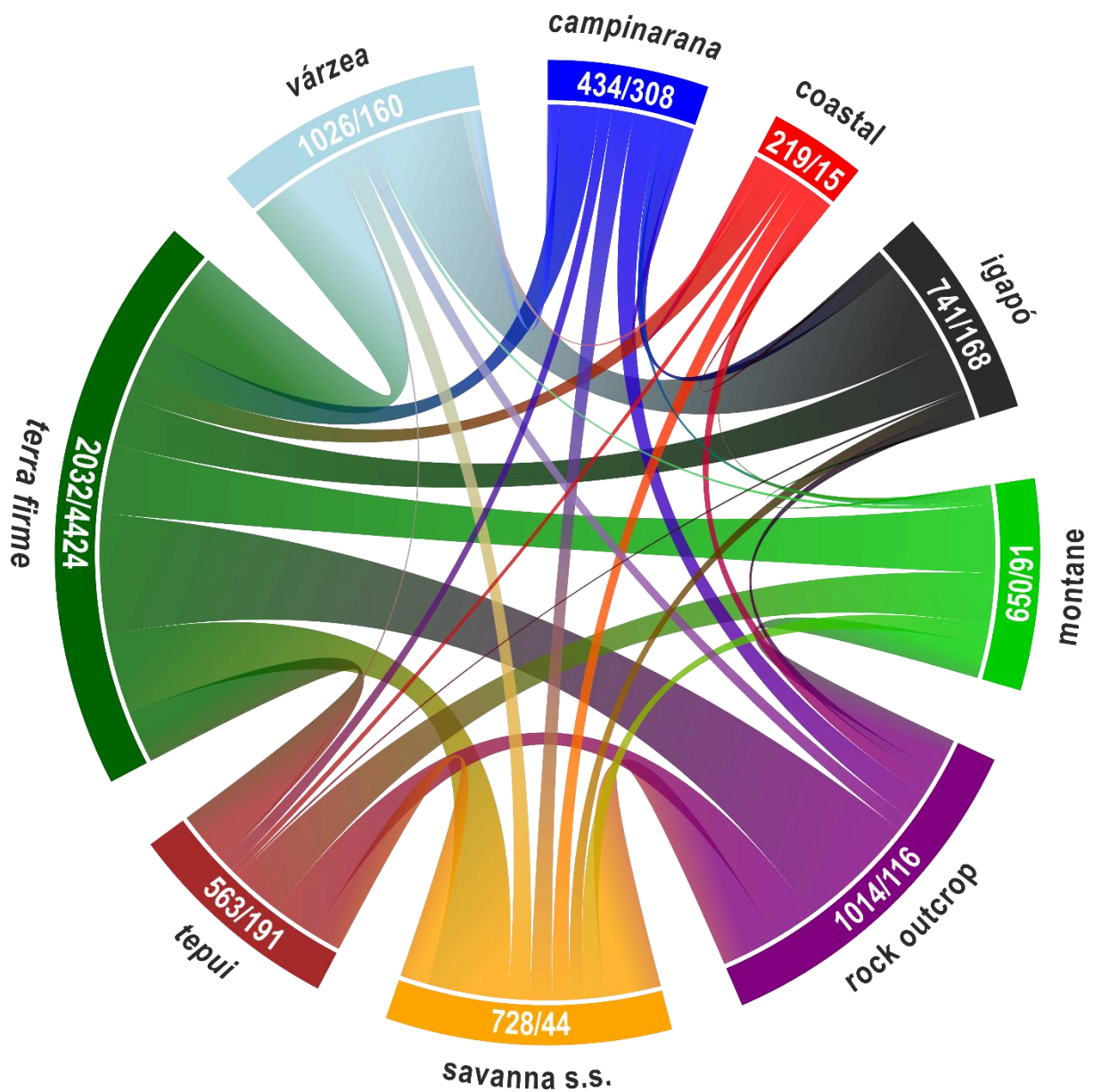classification into nine main vegetation types, and darker shades in each color indicate

overlapping circles (i.e., two or more sites show high similarity in species composition).

PrecSeas = precipitation seasonality; CloudItcp = cloud interception; SoilFert = soil fertility;

SWS = soil water storage capacity; TempAnn = mean annual temperature.

Figure 3. Overlap in tree species composition amongst Amazonian vegetation types. Values in white express the number of species that are shared amongst vegetation types or restricted to a given vegetation type, respectively. Chord width is proportional to the number of shared species.

Author contributions: A.O.F compiled the database, conceived the idea and designed the manuscript; D.M.N. analysed the data; A.O.F. and D.M.N. led the writing with substantial input from R.T.P., K.G.D. and M.F.S. All authors commented on the manuscript and approved the final version.

544 **CONFLICT OF INTEREST**

545 The authors declare no conflict of interest.

546

547 **DATA AVAILABILITY STATEMENT**

548 Presence/absence data for the 8,224 tree species found across the 1,584 Amazonian

549 communities were extracted from the NeoTropTree database (available at

550 http://www.neotroptree.info/data). Bioclimatic variables and altitude were obtained from

551 WorldClim 1.4 data layers (available at: http://www.worldclim.org/download). Soil variables

552 were obtained from the Harmonized World Soil Database v 1.2 (available at:

553 http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-

554 database-v12/en/). Soil Water Storage capacity was obtained from the International Soil

555 Moisture Network (available at https://ismn.geo.tuwien.ac.at/en/).

556

REFERENCES

Adeney, J. M., Christensen, N. L., Vicentini, A., & Cohn-Haft, M. (2016). White-sand

    ecosystems in Amazonia. *Biotropica*, **48**, 7-23.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity.

    Global Ecology and Biogeography, **19**, 134–143.

Bernal, R., Gradstein, S. R., & Celis, M. (2016). *Catálogo de plantas y líquenes de*

    Colombia. Bogotá: Universidad Nacional de Colombia.

Berry, P. E., & Riina R. (2005). Insights into the diversity of the Pantepui flora and the

    biogeographic complexity of the Guayana Shield. *Biologiske Skrifter*, **55**,145-167.

Bivand, R., & Lewin-Koh, N. (2017). *maptools: tools for reading and handling spatial*

    objects. R package version 0.9-2. https://CRAN.R-project.org/package=maptools

Boggan, J., Funk, V., Kelloff, C., Hoff, M., Cremers, G., & Feuillet, C. (1997). *Checklist of*

    the plants of the Guianas (Guyana, Surinam, French Guiana) (2nd ed.). Georgetown,

    Guyana: Centre for the Study of Biological Diversity, University of Guyana.

Bostock, M., Ogievetsky, V., & Heer, J. (2011). D³ Data-Driven Documents. *IEEE*

    Transactions on Visualization and Computer Graphics, **17**, 2301-2309.

Breckle, S.-W. (2002). *Walter's vegetation of the earth, the ecological systems of the geo-*

    biosphere (4th ed.) Berlin: Springer.

Bush, M. B., McMichael, C. H., Raczka, M. F., de Toledo, M. B., Power, M. J., Mayle, F. E.,

    & de Oliveira, P. E. (2014). The Holocene of the Amazon. In I. de Souza Carvalho, M.

    J. Garcia, O. Strohschoen, & C. C. Lana (Eds.). *Paleontologia: cenários de vida –*

    Paleoclimas (pp. 387-396). Rio de Janeiro: Interciência.

Buzatti, R. S. O., Pfeilsticker, T. R., Magalhaes, R. F., Bueno, M. L., Lemos-Filho, J. P., &

    Lovato, M. B. (2018). Genetic and historical colonization analyses of an endemic

581   savanna tree, *Qualea grandiflora*, reveal ancient connections between Amazonian

582   savannas and Cerrado core. *Frontiers in Plant Science*, **9**, 981.

583 Cabrera, A. L., & Willink, A. (1980). *Biogeografia de America Latina.* Washington, DC:

584   Oganization of American States.

585 Cardoso, D., Särkinen,T., Alexander, S., Amorim, A. M., Bittrich, V., Celis, M., Daly, D. C.,

586   Fiaschi, P., Funk, V. A., Giacomin, L. L., Goldenberg, R., Heiden, G., Iganci, J.,

587   Kelloff, C. L., Knapp, S., Lima, H. C., Machado, A. F. P., Santos, R. M., Silva, R. M.,

588   Michelangeli, F. A., Mitchell, J., Moonlight, P., Moraes, P. L. R., Mori, S. A., Nunes,

589   T. S., Pennington, T. D., Pirani, J. R., Prance, G. T., Queiroz, L. P., Rapini, A., Riina,

590   R., Rincon, C. A. V., Roque, N., Shimizu, G., Sobral, M., Stehmann, J. R., Stevens, W.

591   D., Taylor, C. M., Trovó, M., van den Berg, C., van der Werff, H., Viana P. L.,

592   Zartman, C. E., & Forzza, R. C. (2017). Amazon plant diversity revealed by a

593   taxonomically verified species list. *Proceedings of the National Academy of Sciences of*

594   *the United States of America*, **40**, 10695–10700.

595 Clarke, K. R. (1993). Non-parametric multivariate analysis of changes in community

596   structure. *Australian Journal of Ecology*, **18**, 117–143.

597 Coomes, D. A., & Grubb, P. J. (1996). Amazonian caatinga and related communities at La

598   Esmeralda, Venezuela: forest structure, physiognomy and floristics, and control by soil

599   factors. *Vegetatio*, **122**, 167-191.

600 Cremers, J. C., & Hoff, M. (2003). *Guide de la flor des bords de mer de Guyane française.*

601   Paris : Muséum National d'Histoires Naturelle, Patrimoines Naturels 59.

602 Daly, D. C., Silveira, M., Medeiros, H., Castro, W., & Obermüller, F. A. (2016). The white-

603   sand vegetation of Acre, Brazil. *Biotropica*, **48**, 81-89.

604   Dapporto, L.; Ramazzotti, M.; Fattorini, S.; Vila, R.; Talavera, G., & Dennis, R.H.L. (2015).

605   *recluster: ordination methods for the analysis of beta-diversity indices*. R package

606   version 2.8. Available at: https://CRAN.R-project.org/package=recluster.

607   De Cáceres, M., Font, X., & Oliva, F. Assessing species diagnostic value in large data sets: A

608   comparison between phi-coefficient and Ochiai index. *Journal of Vegetation Science*,

609   **19**, 779–788.

610   De Cáceres; M. & Legendre, P. (2009) Associations between species and groups of sites:

611   indices and statistical inference. *Ecology*, **90**, 3566-3574.

612   Demarchi, L. O., Scudeller, V. Z., Moura, L. C., Dias-Terceiro, R. G., Lopes, A., Wittmann,

613   F. K., & Piedade, M. T. F. (2018). Floristic composition, structure and soil-vegetation

614   relations in three white-sand soil patches in central Amazonia. *Acta Amazonica*, **48**, 46-

615   56.

616   Devecchi, M. F., Lovo, J., Moro, M. F., Andrino, C. O., Barbosa-Silva, R. G., Viana, P. L.,

617   Giulietti, A. M., Antar, G., Watanabe, M. T. C., & Zappi, D. C. (2020) Beyond forests

618   in the Amazon: biogeography and floristic relationships of the Amazonian savannas.

619   *Botanical Journal of the Linnean Society*, **193**, 478–503.

620   Dexter, K. G., Lavin, M., Torke, B. M., Twyford, A. D., Kursar, T. A., Coley, P. D., Drake,

621   C., Hollands, R., & Pennington, R. T. (2017). Dispersal assembly of rain forest tree

622   communities across the Amazon basin. *Proceedings of the National Academy of*

623   *Sciences of the United States of America*, **114**, 2645-2650.

624

625   Draper, F. C., Honorio Coronado, E. N., Roucoux, K. H., Lawson, I. T., Pitman, N. C. A.,

626   Fine, P. V., Phillips, O. L., Torres Montenegro, L. A., Valderrama Sandoval, E.,

627   Mesones, I., García- Villacorta, R., Arévalo, F. R. R., & Baker, T. R. (2018). Peatland

628   forests are the least diverse tree communities documented in Amazonia, but contribute

629   to high regional beta- diversity. *Ecography*, **41**, 1256-1269.

630 Draper, F. C., Asner, G. P., Honorio Coronado, E. N., Baker, T. R., García- Villacorta, R.,

631   Pitman, N. C. A., Fine, P. V. A., Phillips, O. L., Zárate Gómez, R., Amasifuén Guerra,

632   C. A., Flores Arévalo, M., Vásquez Martínez, R., Brienen, R. J. W., Monteagudo-

633   Mendoza, A., Torres Montenegro, L. A., Valderrama Sandoval, E., Roucoux, K. H.,

634   Ramírez Arévalo, F. R., Mesones Acuy, Í., Del Aguila Pasquel, J., Tagle Casapia, X.,

635   Flores Llampazo, G., Corrales Medina, M., Reyna Huaymacari, J., & Baraloto, C.

636   (2019). Dominant tree species drive beta diversity patterns in western Amazonia.

637   *Ecology*, **100**, e02636.

638

639 Duivenvoorden, J. F. (1995). Tree species composition and rain forest-environment

640   relationships in the midlle Caqueta area, Colombia, NW Amazonia. *Vegetatio*, **120**, 91-

641   113.

642 Fine, P. V. A., García-Villacorta, R., Pitman, N. C. A., Mesones, I., & Kembel, S. W. (2010).

643   A floristic study of the white-sand forests of Peru. *Annals of the Missouri Botanical*

644   *Garden*, **97**, 283-305.

645 García-Villacorta, R., Dexter, K. G., & Pennington, T. (2016). Amazonian white-sand forests

646   show strong floristic links with surrounding oligotrophic habitats and the Guiana

647   Shield. *Biotropica*, **48**, 47-57.

648 González, V. (2011). Los bosques del Delta del Orinoco. *BioLlania*, **10**, 197-240.

649 Gröger, A. (2000). Flora and vegetation of inselbergs in southern Venezuela. In S.

650   Porembski, & W. Barthlott (Eds.). *Inselbergs - biotic diversity of isolated rock outcrops*

651   *in tropical and temperate regions* (pp. 291-314). Berlin: Springer-Verlag, Ecological

652   Studies.

653   Guitet, S., Pélissier, R., Brunaux, O., Jaouen, G., & Sabatier, D. (2015). Geomorphological

654   landscape features explain floristic patterns in French Guiana rainforest. *Biodiversity*

655   *and Conservation*, **24**, 1215-1237.

656   Haffer, J. (1993). Time's cycle and time's arrow in the history of Amazonia. *Biogeographica*,

657   **69**, 15-45.

658   Higgins, M. A., Ruokolainen, K., Tuomisto, H., Llerena, N., Cardenas, G., Phillips, O. L.,

659   Vásquez, R., & Räsänen, M. (2011). Geological control of floristic composition in

660   Amazonian forests. *Journal of Biogeography*, **38**, 2136-2149.

661   Hijmans, R. J. (2016). *raster: geographic data analysis and modeling.* R package version

662   2.5-8. https://CRAN.R-project.org/package=raster

663   Huber, O. (1997). Pantepui region of Venezuela. In S. L. D. Davis, V. H. Heywood, O.

664   Herrera-MacBryde, J. Villa-Lobos, & A.C. Hamilton (Eds.), *Centres of plant diversity:*

665   *a guide and strategy for their conservation* (Vol. 3, pp. 308-311). Cambridge, UK:

666   World Widelife Fundation for Narure (WWF) and The World Conservation Union

667   (IUCN), IUCN Publications Unit.

668   Huber, O. (2005). Diversity of vegetation types in the Guyana region: an overview.

669   *Biologiske Skrifter*, **55**, 169-188.

670   IBGE (2012). *Manual técnico da vegetação brasileira: sistema fitogeográfico, inventário das*

671   *formações florestais e campestres, técnicas e manejo de coleções botânicas,*

672   *procedimentos para mapeamentos.* Rio de Janeiro: Instituto Brasileiro de Geografia e

673   Estatística (IBGE), Manuais Técnicos de Geociências.

674   Junk, W. J., Piedade, M. T. F., Schöngart, J., Cohn-Haft, M., Adeney, J. M., & Wittmann, F.

675   (2011). A classification of major naturally-occurring Amazonian lowland wetlands.

676   *Wetlands*, **31**, 623-640.

677 Kalliola, R., Salo, J., Puhakka, M., Rajasilta, M., Häme, T., Neller, R. J., Räsänen, M. E., &

678   Danjoy Arias, W. A. (1992). Upper Amazon channel migration: Implications for

679   vegetation perturbance and succession using bitemporal Landsat MSS images.

680   *Naturwissenschaften*, **79**, 75-79.

681 Kubitzki, K. (1987). The ecogeographical differentiation of Amazonian inundation forests.

682   *Plant Systematics and Evolution*, **162**, 285-304.

683 Kubitzki, K. (1990). The psammonophilous flora of northern South America. *Memoirs of the*

684   *New York Botanical Garden*, **64**, 248-253.

685 Luize, B. G., Magalhães, J. L. L., Queiroz, H., Lopes, M. A., Venticinque, E. M., Novo, E.

686   M. L. M. & Silva, T. S. F. (2018). The tree species pool of Amazonian wetland forests:

687   Which species can assemble in periodically waterlogged habitats? *PLoS ONE*, **13**,

688   e0198130.

689 McCune, B. & Grace, J. B. (2002). *Analysis of ecological communities.* Gleneden Beach,

690   OR: MjM Software Design.

691 Melack, J. M., & Hess, L. L. (2010). Remote sensing of the distribution and extent of

692   wetlands in the Amazon Basin. *Ecological Studies, Analysis and Synthesis*, **210**, 43-59.

693 Montero, J. C., Piedade, M. T. F., & Wittman, F. (2014). Floristic variation across 600 km of

694   inundation forests (Igapó) along the Negro River, Central Amazonia. *Hydrobiologia*,

695   **729**, 229-246.

696 Naimi, B; Hamm, Na; Groen, T.A.; Skidmore, A.K., & Toxopeus, A.G. (2014). Where is

697   positional uncertainty a problem for species distribution modelling. *Ecography* **37**, 191-

698   203.

699 Nascimento, W. R., Souza-Filho, P. W. N., Proisy, C., Lucas, R. M., & Rosenqvist, A.

700   (2013). Mapping changes in the largest continuous Amazonian mangrove belt using

701   object-based classification of multisensor satellite imagery. *Estuarine, Coastal and*

702   *Shelf Science*, **117**, 83-93.

703   Neves, D. M., Dexter, K. G., Pennington, R. T., Bueno, M. L., & Oliveira-Filho, A. T.

704   (2015). Environmental and historical controls of floristic composition across the South

705   American Dry Diagonal. *Journal of Biogeography*, **42**, 1566-1576.

706   Neves, D. M., Dexter, K. G., Pennington, R. T., Valente, A. M., Bueno, M. L., Eisenlohr, P.

707   V., Fontes, M. A. L., Miranda, P. L. S., Moreira, S. N., Rezende, V. L., Saiter, F. Z., &

708   Oliveira-Filho, A. T. (2017). Dissecting a biodiversity hotspot: The importance of

709   environmentally marginal habitats in the Atlantic Forest Domain of South America.

710   *Diversity and Distributions*, 23, 898-909.

711   Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B.,

712   Simpson, G. L., Solymos, P., Stevens, M. H. H. & Wagner, H. (2016). *vegan:*

713   *community ecology package.* R package version 2.0–3. Available at: http://CRAN.R-

714   project.org/package=vegan.

715   Oliveira-Filho, A. T. (2015). Um Sistema de classificação fisionômico-ecológica da

716   vegetação Neotropical. In P. V. Eisenlohr, J. M. Felfili, M. M. R. F. Melo, L. A.

717   Andrade, & J. A. A. Meira-Neto (Eds.), *Fitossociologia no Brasil: métodos e estudos*

718   *de casos* (Vol. 2, pp. 452-473). Viçosa, Brazil: Editora UFV.

719   Pennington, R. T., Lewis, G., & Ratter, J. A. (2006). *Neotropical savannas and dry forests:*

720   *plant diversity, biogeography and conservation.* Florida: CRC Press.

721   Pitman, N. C. A., Mogollon, H., Davila, N., Ríos, M., Garcia-Villacorta, R., Guevara, J.,

722   Baker, T. R., Monteagudo, A., Phillips, O. L., & Vasquez-Martinez, R. (2008). Tree

723   community change across 700 km of lowland Amazonian forest from the Andean

724   foothills to Brazil. *Biotropica*, **40**, 525-535.

725 Prance, G. T. (1979). Notes on the Vegetation of Amazonia III. The terminology of

726   Amazonian forest types subject to inundation. *Brittonia*, **31**, 26-38.

727 Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*.

728   Cambridge, NY: Cambridge University Press.

729 Raghoenandan, U. P. D. (2000). The Guianas (Guyana, Suriname, French Guiana). In S.

730   Porembski, & W. Barthlott (Eds), *Inselbergs - biotic diversity of isolated rock outcrops*

731   *in tropical and temperate regions* (pp. 315-338). Berlin: Springer-Verlag, Berlin,

732   Ecological Studies.

733 Ratter, J. A., Bridgewater, S. & Ribeiro, F. (2006). Biodiversity patterns of the woody

734   vegetation of the Brazilian Cerrado. In R. T. Pennington, G. P. Lewis, & J. A. Ratter

735   (Eds.), *Neotropical savannas and seasonally dry forests: plant diversity, biogeography*

736   *and conservation* (pp.31-65). Florida: CRC Press.

737 Rossetti, D. F., Bertani, T. C., Zani, H., Cremon, E. H., & Hayakawa, E. H. (2012). Late

738   Quaternary sedimentary dynamics in Western Amazonia: implications for the origin of

739   savanna/forest contrasts. *Geomorphology*, **177**, 74-92.

740 Salo, J., Kalliola, R., Häkkinen, I., Mäkinen, Y., Niemelä, P., Puhakka, M. & Coley, P. D.

741   (1986). River dynamics and the diversity of Amazon lowland forest. *Nature*, **322**, 254-

742   258.

743 Salovaaraa, K. J., Thessler, S., Malik, R. N. & Tuomisto, H. (2005). Classification of

744   Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery.

745   *Remote Sensing of Environment*, **97**, 39-51.

746 Sarmiento, G. (1983). The savannas of tropical America. In D. W. Goodall (Ed.), *Ecosystems*

747   *of the world – tropical savannas* (pp. 245-288). Amsterdam, The Netherlands: Elsevier.

748 Sarmiento, G. (1984). *The ecology of neotropical savannas.* Cambridge, MA: Harvard

749   University Press.

750   Scudeller, V. (2018). Do the igapó trees species are exclusive to this phytophysiognomy? Or

751   geographic patterns of tree taxa in the igapó forest – Negro River – Brazilian Amazon.

752   In R. W. Myster (Ed.), *Igapó (black-water flooded forests) of the Amazon Basin* (pp.

753   185-207). New York: Springer.

754   Simpson, G. G. (1960). Notes on the measurement of faunal resemblance. *American Journal*

755   *of Science*, **258**, 300-311.

756   Silva, R. M., Mehlig, U., Santos, J. U. M., & Menezes, M. P. M. (2010). The coastal restinga

757   vegetation of Pará, Brazilian Amazon: a synthesis. *Revista Brasileira Botânica*, **33**,

758   563-573.

759   Silva-Souza, K. J. P., & Souza, A. F. (2020). Woody plant subregions of the Amazon forest.

760   *Journal of Ecology* **108**, 1–15.

761   Soares-Filho, B., Nepstad, D., Curran, L. Cerqueira, G. C., Garcia, R. A., Ramos, C. A., Voll,

762   E., McDonald, A., Lefebvre, P. & Schlesinger, P. (2006). Modelling conservation in

763   the Amazon basin. *Nature* **440**, 520–523.

764   Steyermark, J., Berry, P., & Holst, B (1995-2005). *Flora of the Venezuelan Guayana - 9 vols.*

765   St. Louis, MI: Missouri Botanical Gardens Press.

766   Stropp, J., Van der SleenI, P., Assunção, P. A., Silva, A. L., & ter Steege, H. (2011). Tree

767   communities of white-sand and terra-firme forests of the upper Rio Negro. *Acta*

768   *Amazonica*, **41**, 521-544.

769   Terborgh, J., & Andresen, E. (1998). The composition of Amazonian forests: patterns at local

770   and regional scales. *Journal of Tropical Ecology*, **14**, 645-664.

771   ter Steege, H., Lilwah, R., Ek, R. C., van der Hout, P., Thomas, R., van Essen, J, & Jetten,

772   V.G. (2000). *Composition and diversity of the rain forest in central Guyana - an*

773   *addendum to 'Soils of the rainforest in Central Guyana.* Utrecht, The Netherlands:

774   Tropenbos Guyana Program, Utrecht University.

775   ter Steege, H., Pitman, N.C., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.F.,

776   Prévost, M.F., Spichiger, R., Castellanos, H., and von Hildebrand, P. (2006).

777   Continental-scale patterns of canopy tree composition and function across

778   Amazonia. *Nature*, **443**, 444-447.

779   ter Steege, H., Vaessen, R. W., Cárdenas-López, D., Sabatier, D., Antonelli, A., de Oliveira,

780   S. M., Pitman, N. C. A., Jørgensen, P. M., & Salomão, R. P. (2016). The discovery of

781   the Amazonian tree flora with an updated checklist of all known tree taxa. *Scientific*

782   *Reports*, **6**, 29549.

783   Tichý, L., & Chytrý, M. (2006). Statistical determination of diagnostic species for site groups

784   of unequal size. *Journal of Vegetation Science*, **17**, 809-818.

785   Webster, G. L. (1995). The panorama of neotropical cloud forests. In S. P. Churchill, H.

786   Balslev, E. Forero, & J. L. Luteyn (Eds.), *Biodiversity and conservation of neotropical*

787   *montane forests* (pp. 53-77). New York, NY: The New York Botanical Garden.

788   Whitmore, T. C. (1990). *An introduction to tropical rain forests.* Oxford, UK: Clarendon

789   Press.

790   Wittmann, F., Anhuf, D., & Junk, W. J. (2002). Tree species distribution and community

791   structure of central Amazonian várzea forests by remote-sensing techniques. *Journal of*

792   *Tropical Ecology*, **18**, 805-820

793   Wittmann, F., Schöngart J., & Junk, W. J. (2010). Phytogeography, species diversity,

794   community structure and dynamics of Amazonian floodplain forests. In W. J. Junk, M.

795   T. F. Piedade, F. Wittmann, J. Schöngart, & P. Parolin (Eds.), *Amazonian floodplain*

796   *forests: ecophysiology, biodiversity and sustainable management.* (pp. 61-102).

797   Amsterdam, The Netherlands: Springer, Ecological Studies.

798   Wittmann, F., Householder, E., Piedade, M. T. F., Assis, R. L., Schöngart, J., Parolin, P., &

799   Junk W. J. (2012). Habitat specificity, endemism and the neotropical distribution of

800   Amazonian white-water floodplain trees. *Ecography*, **36**, 690-707

801   Worbes, M., Klinge, H., Revilla, J. D., & Martins, C. (1992). On the dynamics, floristic

802   subdivision and geographical distribution of várzea forest in Central Amazonia. *Journal*

803   *of Vegetation Science*, **3**, 553-564.

804 **SUPPLEMENTARY INFORMATION**

805 Additional supporting information may be found in the online version of this article:

806

807 **Table S1.** List of 1,584 tree communities used in this study with their respective metadata,

808 including latitudes and longitudes, vegetation types, and sources.

809

810 **Table S2.** Diagnostic species of the nine main vegetation types of the Amazon Domain by

811 applying the Tichý and Chytrý procedure to the species matrix.