

1 **TITLE: Amazon tree dominance across forest strata**

2 Frederick C. Draper<sup>1,2,3\*</sup>, Flavia R. C. Costa<sup>4</sup>, Gabriel Arellano<sup>5</sup>, Oliver L. Phillips<sup>2</sup>, Alvaro  
3 Duque<sup>6</sup>, Manuel J. Macía<sup>7,8</sup>, Hans ter Steege<sup>9,10</sup>, Gregory P. Asner<sup>3</sup>, Erika Berenguer<sup>11,12</sup>,  
4 Juliana Schietti<sup>4</sup>, Jacob B. Socolar<sup>13</sup>, Fernanda Coelho de Souza<sup>4</sup>, Kyle G. Dexter<sup>14</sup>, Peter M.  
5 Jørgensen<sup>15</sup>, J. Sebastian Tello<sup>15</sup>, William E. Magnusson<sup>4</sup>, Timothy R. Baker<sup>2</sup>, Carolina V.  
6 Castilho<sup>16</sup>, Abel Monteagudo-Mendoza<sup>17</sup>, Paul V. A. Fine<sup>18</sup>, Kalle Ruokolainen<sup>19</sup>, Euridice  
7 N. Honorio Coronado<sup>20</sup>, Gerardo Aymard<sup>21,22</sup>, Nállarett Dávila<sup>20</sup>, Mauricio Sánchez Sáenz<sup>6</sup>,  
8 Marcos A. Rios Paredes<sup>20</sup>, Julien Engel<sup>23</sup>, Claire Fortunel<sup>23</sup>, C. E. Timothy Paine<sup>24</sup>, Jean-  
9 Yves Goret<sup>25</sup>, Aurelie Dourdain<sup>26</sup>, Pascal Petronelli<sup>26</sup>, Elodie Allie<sup>25</sup>, Juan E. Guevara  
10 Andino<sup>27</sup>, Roel J.W. Brienen<sup>2</sup>, Leslie Cayola Pérez<sup>28</sup>, Ângelo, G. Manzatto<sup>29</sup>, Narel Y.  
11 Paniagua Zambrana<sup>28</sup>, Jean-François Molino<sup>23</sup>, Daniel Sabatier<sup>23</sup>, Jérôme Chave<sup>30</sup>, Sophie  
12 Fauset<sup>31</sup>, Roosevelt Garcia Villacorta<sup>32</sup>, Maxime Réjou-Méchain<sup>23</sup>, Paul E. Berry<sup>5</sup>, Karina  
13 Melgaço<sup>4</sup>, Ted R. Feldpausch<sup>33</sup>, Elvis Valderamma Sandoval<sup>34</sup>, Rodolfo Vasquez Martinez<sup>35</sup>,  
14 Italo Mesones<sup>18</sup>, André B. Junqueira<sup>4,36</sup>, Katherine H. Roucoux<sup>37</sup>, José J. de Toledo<sup>38</sup>, Ana C.  
15 Andrade<sup>4</sup>, José Luís Camargo<sup>4</sup>, Jhon del Aguila Pasquel<sup>20</sup>, Flávia D. Santana<sup>4</sup>, William F.  
16 Laurance<sup>39</sup>, Susan G. Laurance<sup>39</sup>, Thomas E. Lovejoy<sup>40</sup>, James A. Comiskey<sup>41,42</sup>, David R.  
17 Galbraith<sup>2</sup>, Michelle Kalamandeen<sup>43,44</sup>, Gilberto E. Navarro Aguilar<sup>34</sup>, Jim Vega Arenas<sup>34</sup>,  
18 Carlos A. Amasifuen Guerra<sup>45</sup>, Manuel Flores<sup>34</sup>, Gerardo Flores Llampazo<sup>20</sup>, Luis A. Torres  
19 Montenegro<sup>46</sup>, Ricardo Zarate Gomez<sup>20</sup>, Marcelo P. Pansonato<sup>47</sup>, Victor Chama Moscoso<sup>17,35</sup>,  
20 Jason Vleminckx<sup>3</sup>, Oscar J Valverde Barrantes<sup>3</sup>, Joost F. Duivenvoorden<sup>48</sup>, Sidney Araújo de  
21 Sousa<sup>49</sup>, Luzmila Arroyo<sup>50</sup>, Ricardo O. Perdiz<sup>4</sup>, Jessica Soares Cravo<sup>4</sup>, Beatriz S. Marimon<sup>51</sup>,  
22 Ben Hur Marimon Junior<sup>51</sup>, Fernanda Antunes Carvalho<sup>52</sup>, Gabriel Damasco<sup>53</sup>, Mathias  
23 Disney<sup>54</sup>, Marcos Salgado Vital<sup>49</sup>, Pablo R. Stevenson Diaz<sup>55</sup>, Alberto Vicentini<sup>4</sup>, Henrique  
24 Nascimento<sup>4</sup>, Niro Higuchi<sup>4</sup>, Tinde Van Andel<sup>9</sup>, Yadvinder Malhi<sup>12</sup>, Sabina Cerruto  
25 Ribeiro<sup>56</sup>, John W. Terborgh<sup>57</sup>, Raquel S. Thomas<sup>58</sup>, Francisco Dallmeier<sup>59</sup>, Adriana Prieto<sup>60</sup>,

26 Renato R. Hilário<sup>38</sup>, Rafael P. Salomão<sup>61,62</sup>, Richarlly da Costa Silva<sup>56</sup>, Luisa F. Casas<sup>63</sup>, Ima  
27 C. Guimarães Vieira<sup>62</sup>, Alejandro Araujo-Murakami<sup>50</sup>, Fredy Ramirez Arevalo<sup>64</sup>, Hirma  
28 Ramírez-Angulo<sup>65</sup>, Emilio Vilanova Torre<sup>65,66</sup> Maria C. Peñuela<sup>67</sup> Timothy J. Killeen<sup>68</sup>,  
29 Guido Pardo<sup>69</sup>, Eliana Jimenez-Rojas<sup>70</sup>, Wenderson Castro<sup>56</sup>, Darcy Galiano Cabrera<sup>17</sup>, John  
30 Pipoly<sup>71,72</sup>, Thaianie Rodrigues de Sousa<sup>4</sup>, Marcos Silvera<sup>73</sup>, Vincent Vos<sup>69</sup>, David Neill<sup>74</sup>,  
31 Percy Núñez Vargas<sup>17</sup>, Dilys M. Vela<sup>75</sup>, Luiz E.O.C Aragão<sup>76</sup>, Ricardo Keichi Umetsu<sup>51</sup>,  
32 Rodrigo Sierra<sup>77</sup>, Ophelia Wang<sup>78</sup>, Kenneth R. Young<sup>79</sup>, Nayane C.C.S. Prestes<sup>51</sup>, Klécia G.  
33 Massi<sup>80</sup>, José Reyna Huaymacari<sup>20</sup>, Germaine A. Parada Gutierrez<sup>50</sup>, Ana M. Aldana<sup>55</sup>,  
34 Miguel N. Alexiades<sup>81</sup>, Fabrício Baccaro<sup>82</sup>, Carlos Céron<sup>83</sup>, Adriane Esquivel Muelbert<sup>84</sup>,  
35 Julio M. Grandez Rios<sup>20</sup>, Antonio S. Lima<sup>62</sup>, Nigel C.A. Pitman<sup>85</sup>, Luis Valenzuela  
36 Gamarra<sup>35</sup>, Cesar J. Cordova Oroche<sup>20</sup>, Alfredo F. Fuentes<sup>28</sup>, Walter Palacios<sup>86</sup>, Sandra  
37 Patiño<sup>87</sup>, Armando Torres-Lezama<sup>65</sup> and Christopher Baraloto<sup>1</sup>

38

39 Affiliations

- 40 1. Institute of Environment, Department of Biological Sciences, Florida International  
41 University, Miami, FL, USA
- 42 2. School of Geography, University of Leeds, Leeds, UK
- 43 3. Center for Global Discovery and Conservation Science, Arizona State University,  
44 Tempe, AZ, USA
- 45 4. Instituto Nacional de Pesquisas da Amazônia – INPA, Manaus, AM, Brazil
- 46 5. Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA
- 47 6. Departamento de Ciencias Forestales, Universidad Nacional de Colombia, Medellín,  
48 Antioquia, Colombia
- 49 7. Departamento de Biología, Universidad Autónoma de Madrid, Spain
- 50 8. Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM),  
51 Universidad Autónoma de Madrid, Madrid, Spain
- 52 9. Naturalis Biodiversity Center, Leiden, The Netherlands
- 53 10. Systems Ecology, Vrije Universiteit, Amsterdam, The Netherlands
- 54 11. Lancaster Environment Centre, Lancaster University, Lancaster, UK
- 55 12. Environmental Change Institute, University of Oxford, Oxford, UK
- 56 13. Faculty of Environmental Sciences and Natural Resource Management, Norwegian  
57 University of Life Sciences, Ås, Akershus, Norway
- 58 14. School of Geosciences, University of Edinburgh, Edinburgh, UK
- 59 15. Missouri Botanical Garden, St Louis, MO, USA
- 60 16. Brazilian Agricultural Research Corporation - Embrapa, Embrapa Roraima, Roraima,  
61 Brazil
- 62 17. Universidad Nacional de San Antonio Abad del Cusco, Cusco, Peru
- 63 18. Department of Integrative Biology, University of California Berkeley, Berkeley, CA,  
64 USA
- 65 19. Department of Biology, University of Turku, Turku, Finland
- 66 20. Instituto de Investigaciones de la Amazonía Peruana, Iquitos, Peru
- 67 21. UNELLEZ-Guanare, Programa de Ciencias del Agro y el Mar, Herbario Universitario  
68 (PORT), Mesa de Cavacas, estado Portuguesa, Venezuela
- 69 22. Compensation International Progress S. A. – Ciprogress Greenlife, Bogotá, D. C.  
70 Colombia
- 71 23. AMAP, Université de Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier,  
72 France
- 73 24. Environmental and Rural Science, University of New England, Armidale, New South  
74 Wales 2351 Australia
- 75 25. INRA, UMR EcoFoG, AgroParisTech, CNRS, CIRAD, Université des Antilles,  
76 Université de Guyane, Kourou, France
- 77 26. CIRAD, UMR EcoFoG, Kourou, France
- 78 27. BIOMAS, Universidad de Las Américas, Quito, Ecuador
- 79 28. Instituto de Ecología, Herbario Nacional de Bolivia, La Paz, Bolivia
- 80 29. Departamento de Biologia, Universidade Federal de Rondônia, Porto Velho, Brazil
- 81 30. Laboratoire Evolution et Diversité Biologique (EDB) CNRS/UPS, Toulouse, France

- 82 31. School of Geography, Earth and Environmental Sciences, University of Plymouth,  
83 UK
- 84 32. Department of Ecology and Evolutionary Biology, Cornell University, Ithica, NY,  
85 USA
- 86 33. Department of Geography, University of Exeter, Exter, UK
- 87 34. Facultad de Ciencias Biológicas, Universidad Nacional de la Amazonía Peruana,  
88 Iquito, Peru
- 89 35. Estación Biológica del Jardín Botánico de Missouri, Oxapampa, Peru
- 90 36. Institut de Ciència i Tecnologia Ambientals, Universitat Autònoma de Barcelona,  
91 Barcelona, Spain
- 92 37. School of Geography & Sustainable Development, University of St Andrews, St  
93 Andrews, UK
- 94 38. Department of Environment and Development, Federal University of Amapá, Brazil
- 95 39. Centre for Tropical Environmental and Sustainability Science (TESS) and College of  
96 Marine and Environmental Sciences, James Cook University
- 97 40. Department of Environmental Science and Policy, George Mason University, Fairfax,  
98 VA, USA
- 99 41. Inventory and Monitoring Program, National Park Service, Fredericksburg, VA, USA
- 100 42. Smithsonian Institution, Washington, DC, USA
- 101 43. Department of Plant Sciences, University of Cambridge, UK.
- 102 44. Living with Lakes Centre, Laurentian University, Canada.
- 103 45. DRGB, Instituto Nacional de Innovación Agraria (INIA), Av. La Molina 1981, Lima,  
104 Peru
- 105 46. Herbarium Amazonense (AMAZ). Universidad Nacional de la Amazonia Peruana,  
106 Loreto, Perú
- 107 47. Department of Ecology, Universidade de São Paulo, São Paulo, Brazil
- 108 48. Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, The  
109 Netherlands
- 110 49. Centro de Estudos da Biodiversidade, Universidade Federal de Roraima, Boa Vista,  
111 RR, Brazil
- 112 50. Museo de Historia Natural Noel Kempff Mercado, Universidad Autónoma Gabriel  
113 Rene Moreno, Santa cruz, Bolivia
- 114 51. Faculdade de Ciências Agrárias, Biológicas e Sociais Aplicadas, Universidad do  
115 Estado de Mato Grosso, Nova Xavantina, MT, Brazil
- 116 52. Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, Brazil
- 117 53. Department of Ecology, Evolution and Behaviour, University of Minnesota,  
118 Minneapolis, MN, USA
- 119 54. Department of Geography, University College London, London, UK
- 120 55. Departamento de Ciencias Biológicas, Universidad de Los Andes (Colombia),  
121 Bogotá, Colombia
- 122 56. Centro de Ciências Biológicas e da Natureza, Universidade Federal do Acre, Rio  
123 Branco, Brazil
- 124 57. Nicholas School of the Environment, Duke University, Durham, NC, USA

- 125 58. Iwokrama International Centre for Rainforest Conservation and Development,  
126 Georgetown, Guyana
- 127 59. Smithsonian's National Zoo & Conservation Biology Institute, Washington, DC,  
128 USA
- 129 60. Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá,  
130 Colombia
- 131 61. Universidade Federal Rural da Amazônia - UFRA/CAPES, Belém, Pará, Brazil
- 132 62. Museu Paraense Emílio Goeldi, Belém, Pará, Brasil
- 133 63. Laboratorio de Ecología de Bosques Tropicales y Primatología, Fundación Natura  
134 Colombia, Universidad de Los Andes, Bogotá, Colombia
- 135 64. Facultad de Forestales, Universidad Nacional de la Amazonía Peruana, Iquito, Peru
- 136 65. Institute of Research for Forestry Development, Universidad de los Andes, Merida,  
137 Venezuela
- 138 66. School of Environmental and Forest Sciences (SEFS), University of Washington,  
139 Seattle, WA, USA
- 140 67. Universidad Regional Amazónica Ikiam, Tena, Ecuador
- 141 68. Agteca-Amazonica, Santa Cruz, Bolivia
- 142 69. Universidad Autónoma del Beni, Riberalta, Bolivia
- 143 70. Instituto Amazónico de Investigaciones IMANI, Universidad Nacional de Colombia  
144 Sede Amazonia
- 145 71. Broward County Parks and Recreation, Miami, FL, USA
- 146 72. Biological Sciences, Florida Atlantic University-Davie, Miami, FL, USA
- 147 73. Museu Universitário, Universidade Federal do Acre, Brazil
- 148 74. Facultad de Ingeniería Ambiental, Universidad Estatal Amazónica, Ecuador
- 149 75. Department of Biology, Washington University in St. Louis, St Louis, MO, USA
- 150 76. National Institute for Space Research (INPE), São José dos Campos-SP, Brazil
- 151 77. Geoinformática & Sistemas (GeoIS), Quito, Ecuador
- 152 78. School of Earth Sciences and Environmental Sustainability, Northern Arizona  
153 University, AZ, USA
- 154 79. Department of Geography and the Environment, University of Texas at Austin, TX,  
155 USA
- 156 80. Instituto de Ciência e Tecnologia, São Paulo State University (UNESP), São José dos  
157 Campos, Brazil
- 158 81. School of Anthropology and Conservation, University of Kent, Canterbury, UK
- 159 82. Universidade Federal do Amazonas, Manaus, Brazil
- 160 83. Herbario Alfredo Paredes (QAP), Universidad Central del Ecuador, Quito, Ecuador
- 161 84. School of Geography, Earth and Environmental Sciences, University of Birmingham,  
162 UK
- 163 85. Science and Education, The Field Museum, Chicago, IL, USA
- 164 86. Universidad Tecnica del Norte, Herbario Nacional del Ecuador, Quito, Ecuador
- 165 87. Research Institute Alexander von Humboldt, Bogotá, Colombia
- 166
- 167

168 **Abstract**

169 The forests of Amazonia are among the most biodiverse plant communities on Earth. Given  
170 the immediate threats posed by climate and land-use change, an improved understanding of  
171 how this extraordinary biodiversity is spatially organized is urgently required to develop  
172 effective conservation strategies. Most Amazonian tree species are extremely rare, but a small  
173 number are common across the region. Indeed, just 227 “hyperdominant” species account for  
174 more than 50% of all individuals > 10 cm dbh. Yet, the degree to which the phenomenon of  
175 hyperdominance is sensitive to tree size, the extent to which the composition of dominant  
176 species changes with size-class, and how evolutionary history constrains tree  
177 hyperdominance, all remain unknown. Here, we use a unique floristic dataset to show that,  
178 while hyperdominance is a universal phenomenon across forest strata, different species  
179 dominate the forest understory, midstory and canopy. We further find that although species  
180 belonging to a range of phylogenetically dispersed lineages have become hyperdominant in  
181 small size-classes, hyperdominants in large size-classes are restricted to a few lineages. These  
182 results suggest that achieving hyperdominance over large geographic regions has been much  
183 more challenging for canopy and emergent tree species than for understorey species. Our  
184 results demonstrate that it is essential to consider all forest strata in order to understand  
185 regional patterns of dominance and composition in Amazonia. More generally, through the  
186 lens of 654 hyperdominant species, we outline a tractable pathway for understanding the  
187 functioning of half of Amazonian forests across vertical strata and geographical locations.

188

189 **Main text**

190 The immense diversity of Amazonian forests is one of Earth's great natural wonders, and  
191 underpins the functioning and resilience of ecosystems<sup>1,2</sup> that play a crucial role in the global  
192 carbon and water cycles<sup>3-5</sup>. Despite three centuries of investigation, however, our collective  
193 understanding of how this diversity is organized at regional scales remains limited<sup>6</sup>.

194 Confronted with such overwhelming diversity, the challenge of monitoring a few hundred  
195 hyperdominant species (i.e. those species that together account for 50% of individuals across  
196 Amazonia<sup>7</sup>) becomes more tractable than monitoring the many thousands of rare species,  
197 particularly given the pace of action required for contemporary management decisions<sup>8,9</sup>.

198 Understanding the ecology and distribution of hyperdominant species is essential because  
199 these species dominate key ecosystem processes (e.g. carbon storage and cycling<sup>10</sup>) and may  
200 serve as an effective proxy for general biodiversity patterns<sup>11</sup>.

201 Existing studies of Amazonian hyperdominance and regional-scale dominance have been  
202 limited by excluding small-stemmed individuals (<10 cm diameter) and by considering all  
203 individuals as equivalent regardless of diameter size-class<sup>7,11-15</sup>. Excluding small-stemmed  
204 species represents an important oversight because several thousand Amazonian tree species  
205 rarely or never reach 10 cm in diameter<sup>16-20</sup>. While local-scale and taxa-specific dominance  
206 has been documented in small size classes<sup>21-24</sup>, basin-wide hyperdominance in small size  
207 classes has not been confirmed. Consequently, species dominating the understory of  
208 Amazonian forests at a whole-Amazon scale are not yet identified. Treating all stems > 10 cm  
209 in diameter as equivalent is also likely to over emphasize the dominance of mid-statured tree  
210 species (e.g. 10-20 cm diameter). The power-law relationship between stem density and  
211 diameter means that small-stemmed individuals (e.g. < 20 cm) are at least an order of  
212 magnitude more abundant than larger individuals (e.g. > 50 cm)<sup>25</sup>. This skewed

213 understanding of dominance is highlighted by the difference between lists of hyperdominant  
214 species assembled using biomass rather than abundance<sup>10</sup>. This disparity suggests that a more  
215 nuanced approach that measures dominance separately across forest strata is required.

216 Based on existing local-scale studies and field observations, we expect the composition of  
217 hyperdominant species to vary substantially across forest strata due to different  
218 environmental filters (e.g. variable light profiles) and different species pools. Existing studies  
219 also suggest that compositional similarity between understory and canopy hyperdominant  
220 species may vary regionally<sup>26,27</sup>, perhaps due to regional variation in forest structure and rates  
221 of turnover<sup>28</sup>. For example, western Amazonia is known to have a floristically distinct  
222 understory, whereas understory communities in central and eastern Amazonia are thought to  
223 be comprised primarily by juveniles of larger tree species<sup>26</sup>.

224 We know that some taxonomic tree clades contain many hyperdominant species<sup>7</sup>, and that  
225 genus-level abundance has a significant phylogenetic signal<sup>29</sup>, yet no formal analysis of the  
226 phylogenetic structure of hyperdominance has been undertaken. Moreover, we may expect  
227 that hyperdominant species in different strata will display different phylogenetic patterns.  
228 Specifically, we hypothesize that hyperdominant species in large size classes from across  
229 Amazonia will be phylogenetically clustered for several reasons. First, maximum potential  
230 tree size has a significant phylogenetic signal in Amazonia<sup>30</sup>, and those genera able to occupy  
231 canopy and emergent strata are concentrated in specific lineages (e.g. families or orders) that  
232 are primarily located within a few deep clades (e.g. Fabids and Ericales)<sup>30</sup>. Second, while  
233 there is climatic variation across Amazonia, the above-canopy environment consists of high  
234 solar radiation, high temperatures, low humidity, and high diurnal variability irrespective of  
235 location<sup>31</sup>. These harsh but spatially consistent environmental conditions provide limited  
236 niche space, and are likely to filter for a distinct suite of functional characteristics that may



237 only have arisen in species belonging to a few specific lineages. Third, tall trees tend to  
238 disperse better than smaller trees and shrubs<sup>32,33</sup>, and at least some common large tree  
239 lineages have been well dispersed throughout Amazonia across evolutionary timescales<sup>34</sup>.  
240 This greater dispersal ability may mean that the strongest competitors for the canopy strata  
241 have been able to disperse consistently throughout Amazonia for millennia, thereby  
242 becoming hyperdominant across regions.

243 On the other hand, we may expect hyperdominant species in small understory strata may be  
244 more phylogenetically dispersed. First, because small trees and shrubs typically are more  
245 likely to be dispersal limited<sup>32,33</sup>, the strongest understory competitors may be less likely to  
246 disperse across regions and outcompete functionally equivalent species in other locations.  
247 Secondly, understory species are often locally abundant and frequently have fast generation  
248 times<sup>35</sup>. Over evolutionary timescales these high abundances and fast generation times may  
249 be likely to increase diversification among locally-restricted understory species<sup>36</sup>. Third, the  
250 below canopy environment is more spatially heterogeneous, due to variation in forest  
251 structure, and the frequency and size of forest gaps, potentially leading to increased niche  
252 partitioning in smaller size classes. Moreover, because forest structure varies across  
253 Amazonia (e.g. taller denser canopy in Guiana shield vs shorter more dynamic canopy in  
254 western Amazonia)<sup>28</sup>, smaller-statured species may be exposed to different abiotic and biotic  
255 filters across large spatial scales, and develop greater local specialization associated with  
256 distinct functional characteristics. Therefore, we further predict that understory  
257 hyperdominants from different regions should be more distantly related than hyperdominants  
258 in larger size classes.

259 Here we assemble a unique dataset of 1240 floristic inventory plots distributed across lowland  
260 Amazonia, which include stems as small as 2.5 cm (Figure 1). Based on individual diameter

261 measurements and species level identifications, we implement a spatially-stratified  
262 resampling approach to estimate basin-wide relative abundances for all tree species across six  
263 diameter size classes from the understory (2.5 – 5 cm) to the forest canopy (> 50 cm). Using  
264 this dataset, we identify those species dominating different strata of Amazonian forests and  
265 ask: Q1. Is hyperdominance a constant phenomenon across Amazonian tree strata? and Q2.  
266 Does the identity of hyperdominant species differ across Amazonian tree strata, and how does  
267 this vary regionally? We also used a recently developed genus-level molecular  
268 phylogeny<sup>37,38</sup> to ask Q.3 Do patterns of phylogenetic clustering in hyperdominant species  
269 vary across forest strata? And does this correspond with our expectations of increased  
270 clustering in large-stemmed canopy strata and increased phylogenetic dispersal in small-  
271 stemmed understory strata?

272 **Results and Discussion**

273 *Consistent hyperdominance across strata*

274 We find that hyperdominance occurs throughout the Amazonian flora across forest strata, but  
275 the proportion of species that qualify as hyperdominant (i.e. together account for 50% of  
276 individuals) varies across size classes and regions from 3 – 12 % (Figure 2A). At the basin-  
277 wide Amazonian scale, the proportion of species that qualify as hyperdominant in our dataset  
278 (~ 4%) is broadly consistent with empirically derived species counts from previous studies<sup>7,10</sup>.  
279 This consistency of hyperdominance across size classes suggests that regional dominance of  
280 tree communities is a feature shared across Amazonian forest strata.

281 A larger species pool will necessarily result in stronger patterns of dominance because it  
282 contains more rare species, which will decrease the proportion of species that qualify as  
283 hyperdominant, even if the abundance of the most dominant species remains constant.  
284 However, the relationship between species richness and the level of dominance we observed  
285 in a given size class or region is weak and primarily driven by the basin-wide data (Figure  
286 2B). Therefore, our results suggest that variation in dominance among size classes and  
287 regions is not an artefact reflecting the variable sampling intensity among regions and size  
288 classes.

289 Some size classes are consistently more ‘dominated’ than others (i.e. a lower proportion of  
290 species are required to account for 50% of individuals). In particular, the 10 – 20 cm size  
291 class consistently displays the strongest dominance patterns (Figure 2). The two smallest size  
292 classes have weaker dominance patterns, perhaps because smaller-stemmed species are more  
293 dispersal limited than larger individuals<sup>32</sup>, and therefore less likely to be dominant over large  
294 areas and more likely to locally diversify. A clear exception to this occurs in forests on the  
295 Guiana Shield, where patterns of dominance are stronger in larger size classes than smaller

296 ones. This may be partially explained by the relatively low diversity in the understory of  
297 these forests, due perhaps to greater resource limitation imposed by extreme shade from the  
298 more structured canopy in addition to low fertility associated with oligotrophic soils in this  
299 region<sup>16</sup>.

300 The proportion of species that qualified as hyperdominant at the regional scale was generally  
301 higher than in basin-scale analyses, i.e. dominance patterns are weaker at the regional scale  
302 (Figure 2). This pattern is primarily driven by those exceptionally common and widespread  
303 species that achieve hyperdominance in two or more regions. However, several of these  
304 widespread hyperdominant species may be species complexes, as recently shown for *Protium*  
305 *heptaphyllum* and *Astrocaryum murumuru*<sup>39,40</sup>. Solving these issues will require more  
306 integrative taxonomic studies (e.g. incorporating DNA analyses alongside spectroscopy<sup>39,41</sup>)  
307 of other widespread hyperdominant species, which would help to further assess the validity of  
308 hyperdominant species identifications.

309 Southwest Amazonia exhibits stronger patterns of dominance than all other Amazonian  
310 regions in all but the largest size class (Figure 2). It is not immediately clear why this region  
311 has such strong patterns of dominance. However, it may be due in part to less environmental  
312 heterogeneity in this region, which contains relatively few areas of white-sand forest, swamp  
313 forests or seasonally-inundated forests<sup>42,43</sup>. Although we do not explicitly consider habitat  
314 type in this study, many hyperdominant species are known to be dominant only in a single  
315 habitat type<sup>7</sup>. Therefore, less environmental heterogeneity should lead to fewer  
316 hyperdominant species. The strong dominance patterns in southwestern Amazonia matter  
317 because several landmark studies have focussed on patterns of dominance in this  
318 region<sup>12,13,21</sup>, and these patterns may not be representative of Amazonia more generally<sup>44</sup>.

319 *Identity of hyperdominant species across strata and region*

320 The identity of hyperdominant species varies substantially across forest strata and region.  
321 Over a third (38 %) of hyperdominant species are only dominant in a single size class within  
322 a single region, and nearly two thirds (62 %) are dominant in two or fewer size classes and  
323 two or fewer regions (Figure 3). Only one species, *Eschweilera coriacea*, achieves  
324 hyperdominance across all six size classes and all five regions. These results provide clear  
325 evidence that hyperdominant tree species composition is vertically stratified throughout  
326 Amazonia. Therefore, considering all individuals greater than 10 cm in diameter as  
327 equivalent completely overlooks the nuanced vertical stratification of tropical forests.  
328 Moreover, even though 10 cm diameter cutoff protocols are well-suited to monitor carbon  
329 fluxes<sup>3</sup>, alternative plot designs or data treatments may be better suited to monitor spatial  
330 variation in floristic diversity and composition<sup>45-47</sup>.

331 Our multivariate analysis illustrates two strong axes of compositional variation among  
332 hyperdominant tree species (Figure 4 panel A). The first axis differentiates the five regions,  
333 while the second represents a gradient across six tree size classes. This compositional  
334 variation across strata is important because our best current methods of observing forests at  
335 large scales are through either: 1. Remote sensing approaches, which detect only those trees  
336 that reach sky-facing canopy positions; or 2. Plot networks, which are heavily influenced by  
337 species dominant in smaller or intermediate size-classes. Our results demonstrate that species  
338 dominating the view from above the canopy are different from those that dominate the view  
339 from below, thereby emphasising the mismatch between remotely sensed and plot-based  
340 studies. Addressing this mismatch will be essential to successfully integrating field and  
341 remote sensing data at large scales in Amazonia.

342 Despite this compositional mismatch, our data also suggest that while canopy  
343 hyperdominants comprise different species from those that dominate the understory, there is

344 an important association in hyperdominant species composition between size classes within  
345 regions, i.e. hyperdominant species clearly form distinct regional groups across the first  
346 NMDS axis. Therefore, remotely sensed data from forest canopies may serve as an effective  
347 proxy for compositional patterns in smaller size classes, as has been found recently in  
348 understory tree, fern and lycophyte communities<sup>48,49</sup>.

349 Our results contrast with previous observations, which suggest that the understories of  
350 Eastern Amazonia are primarily composed of juvenile individuals of large-stemmed species  
351 whereas western Amazonia has a more specialist and compositionally distinct understory<sup>26</sup>.  
352 Instead, we find that across all regions, tree species that dominate forest understory tree  
353 communities are compositionally distinct from those that dominate the canopy, i.e. in all  
354 regions hyperdominant species form a distinct compositional gradient across strata, as  
355 reflected by the second NMDS axis (Figure 4). Indeed, there is no difference in potential  
356 maximum size among understory dominant species from different regions (Figure S.2).

357 At the family level, there is a clear positive relationship between the number of  
358 hyperdominant species and total species richness per family (Figure S3). However, our  
359 statistical null modelling approach shows that at a basin-wide scale several plant families  
360 have significantly more or fewer hyperdominant species than would be expected based on  
361 their species richness. Moreover, some families have more hyperdominant species than  
362 expected across several size classes; for example, *Arecaceae*, *Burseraceae* and *Myristicaceae*  
363 have more hyperdominant species than expected across all but the largest size class. Other  
364 families are overrepresented in terms of hyperdominant species in only smaller (e.g.  
365 *Violaceae* and *Siparunaceae*) or larger size classes (e.g. *Moraceae*). Alternatively, commonly  
366 occurring tree families including *Rubiaceae* and *Lauraceae* have consistently fewer  
367 hyperdominant species than we would expect based on their species richness.

368 Our results further reinforce the importance of the legume family Fabaceae in dominating  
369 Amazonian forests. At a basin-wide scale Fabaceae is the family with the greatest number of  
370 hyperdominant species across all size-classes, largely because Fabaceae is by far the most  
371 species rich family. While Fabaceae species are less common than would be expected by  
372 chance given their high species richness in small size classes, in the largest size-class  
373 Fabaceae are significantly overrepresented, and account for more than 30% of hyperdominant  
374 species.

375 *Phylogenetic structure of hyperdominance across Amazonian tree strata*

376 Our phylogenetic analyses demonstrate that while many lineages contain hyperdominant  
377 species (Figure 5), those species that are hyperdominant in the canopy of Amazonian forests  
378 show contrasting phylogenetic patterns to those that are hyperdominant in small understory  
379 strata (Figure 6).

380 We find overall support for our prediction that hyperdominant in larger size classes tend to be  
381 concentrated in a few closely related lineages, for example in Fabaceae and Moraceae as well  
382 as Lecythidaceae and Sapotaceae. This phylogenetic clustering of canopy hyperdominant  
383 species is highlighted by our mean pairwise phylogenetic distance (MPD) null modelling  
384 analysis (Figures 6 and S4), which shows that hyperdominant species in the largest size  
385 classes are consistently more closely related than would be expected by chance. Our  
386 phylogenetic composition results reveal that canopy strata across the basin are dominated by  
387 species belonging to closely related lineages (Figure 4 panel B). The close phylogenetic  
388 relationship among large-stemmed regionally dominant tree species across the basin suggests  
389 that these species have been well dispersed across the basin through evolutionary time,  
390 supporting previous studies that found evidence for widespread dispersal in several common  
391 Amazonian tree lineages<sup>34</sup>.

392 These findings have important implications: If we accept the premise that phylogenetic  
393 diversity is an effective proxy for ecological or indeed functional diversity<sup>50-52</sup>, the high  
394 phylogenetic similarity among canopy species suggests there is lower functional diversity  
395 within a large proportion of the canopy strata. This low functional diversity may in turn  
396 reduce the resilience of these canopy communities to climate change. Previous studies have  
397 shown that large canopy trees in Amazonia have distinct trait profiles (e.g. hydraulic traits)<sup>53</sup>,  
398 appear to be particularly affected by drought<sup>54</sup>, and play a crucial role in Amazonian forest  
399 carbon storage and cycling<sup>10</sup>. We propose that future research should continue to uncover the  
400 functional diversity and potential vulnerability both within and among lineages of these  
401 canopy hyperdominant species.

402 We find contrasting phylogenetic patterns in smaller, understory size-classes, which are  
403 widespread across the phylogeny as we predicted. Indeed, at the basin-wide scale,  
404 hyperdominant species in understory size classes are less closely related than expected by  
405 chance (Figure 6). This dispersed phylogenetic pattern is largely due to understory  
406 hyperdominants occurring across the major angiosperm clades (Figure S.5), but may also be  
407 because our list of understory hyperdominant species is composed of both understory  
408 specialist taxa as well as larger-statured species that achieve dominance as juveniles. Hence,  
409 this mixture of life stages and functional strategies across distinct clades is more likely to lead  
410 to a more phylogenetically dispersed assemblage. Nevertheless, our findings highlight that  
411 several characteristic understory genera such as *Piper* (Piperaceae), *Rinorea* (Violaceae), and  
412 *Miconia* (Melastomataceae) contain numerous hyperdominant species, which have not been  
413 recognised by previous studies of large stem (>10 cm) dominance<sup>7,12,13</sup>.

414 Our phylogenetic compositional analysis also supports our hypothesis that within understory  
415 strata, hyperdominant species from different regions are distantly related (Figure 4 panel B).



416 These results are consistent with limited dispersal and diversification of understory  
417 hyperdominant species at a basin-wide scale over evolutionary timescales, as has been  
418 suggested by others<sup>55</sup>. Furthermore, the ability to become dominant in the understory of  
419 Amazonian forests is found across a diverse range of lineages, and therefore is relatively  
420 common. Because many of these lineages are distantly related, this suggests that a range of  
421 functional strategies has evolved to achieve hyperdominance in Amazonian understories.  
422 Furthermore, the high phylogenetic distance among understory hyperdominant species is  
423 consistent with the hypothesis that greater environmental niche space in the forest understory  
424 has contributed to higher phylogenetic diversity.

#### 425 *Future Directions*

426 The mechanisms that allow certain species to become hyperdominant remain elusive.  
427 However, our results provide a basis for testing hypotheses related to specific ecological  
428 mechanisms. Future analyses should capitalise on increasingly available functional trait data  
429 to tackle these issues. We expect species that dominate the canopy to be functionally distinct  
430 from those that dominate the understory; therefore, a size-class constrained framework may  
431 help to illuminate the mechanisms that underpin hyperdominance. In particular, a large-scale  
432 assessment of quantitative dispersal traits across a range of species may help to unravel why  
433 hyperdominant species in understory size classes display such different phylogenetic patterns  
434 to those in the canopy.

435 Previous studies have presented a compelling case for pre- or post- Columbian peoples  
436 increasing the abundance of many hyperdominant species in order to extract products such as  
437 fruits, nuts or building materials<sup>56,57</sup>. Here, we show that many of these “domesticated”  
438 hyperdominants (e.g. *Euterpe precatoria* and *Theobroma cacao*) are in fact only dominant in  
439 smaller size classes. One possible explanation is that it is easier to harvest and manage small

440 understory trees and shrubs than large canopy trees; therefore, species that dominate larger  
441 size classes may have been less influenced by human activity than species that dominate  
442 smaller size classes. It is important to note that this is not the case in all instances, and there is  
443 substantial evidence that some large-statured species (e.g. *Bertholletia excelsa*) were also  
444 managed during pre-Columbian times<sup>58</sup>. Further investigation into the role of humans in  
445 shaping the composition of Amazon understories may help explain why such distantly related  
446 species have become dominant in different Amazonia regions. For example, paleoecological  
447 records may reveal if different groups of indigenous peoples have propagated different tree  
448 species in different regions.

#### 449 **Conclusions**

450 There is a pressing demand to quantify and monitor the biodiversity of Amazonia in the  
451 coming decades, however, we currently lack the resources necessary to undertake the  
452 ‘Linnaean renaissance’ required to fully document the biota of arguably Earth’s most diverse  
453 forests. By identifying those species that are hyperdominant across forest strata, we outline a  
454 size-class based framework for understanding Amazonian forests, irrespective of strata or  
455 location. This framework has revealed that species dominating either the canopy or  
456 understories of Amazonian forests not only are taxonomically distinct but also represent  
457 different phylogenetic patterns. Species belonging to a range of phylogenetically dispersed  
458 lineages have become hyperdominant in small size classes, whereas species that are  
459 hyperdominant in large size-classes belong to a few specific lineages.

#### 460 **Acknowledgements**

461 We dedicate this study to the late Alwyn Gentry, who not only established 41 of the plots that  
462 form the foundation of our analyses, but also pioneered the synthetic approach that underpins  
463 our study. This paper is a product of the RedGentry, RAINFOR, PPBio and ATDN networks,

464 Data from many of these networks are curated by ForestPlots.net, a cyber-infrastructure  
465 initiative that unites plot records and their contributing scientists from the world's tropical  
466 forests. These initiatives have been supported by numerous people and grants, but we are  
467 indebted to hundreds of institutions, field assistants, botanists, and local communities for help  
468 in establishing plots and identifying their >4600 species. We would especially like to thank  
469 the following for their important role: Eduardo Hase, Raimunda Nazaré Oliveira de Araújo,  
470 Samuel Almeida, Julio Serrano, João Batista de Silva, Katia Cangani, Ocírio Souza Pereira,  
471 Julio do Vale, Maria Carmozina, Everaldo da Costa Pereira, Sebastiao Salvino de Souza,  
472 Cecilia Ballón Falcón, Massiel Corrales Medina, Antônio Magalhães da Silva, Jose Farreras,  
473 Franklin Molina and Jon Lloyd.

474 FCD was funded by an EU MSC global fellowship 794973 'E-FUNDIA'; and FCD and CB  
475 supported the collaborative network with funds from l'Institut national de recherche pour  
476 l'agriculture, l'alimentation et l'environnement (INRAE), Florida International University  
477 and the William R. Kenan, Jr. Charitable Trust. Funding for field plot data collection came  
478 from a wide range of sources but particularly the following grants: *Conselho Nacional de*  
479 *Desenvolvimento Científico e Tecnológico/Projetos Ecológicos de Longa Duração-*  
480 *CNPq/PELD (Nr. 441244/2016-5), Agence Nationale de la Recherche Blanc projet*  
481 *NEBEDIV (ANR- 13-BSV7-009), an "Investissement d'avenir" grant from the Agence*  
482 *Nationale de la Recherche (CEBA, ref. ANR-10- LABX-25-01) and Spanish Ministry of*  
483 *Economy and Competitiveness, grant numbers CGL2015-72431-EXP and CGL2016-75414-*  
484 *P.* Many bodies funded the development of RAINFOR and ForestPlots.net, with key support  
485 including from the Natural Environment Research Council (NERC) grants NE/F005806/1,  
486 NE/D005590/1, NE/N012542/1, NE/N011570/1, as well as a European Research Council  
487 (ERC) grant T-FORCES (291585) and Gordon and Betty Moore Foundation grant (#1656) to  
488 O.L.P. This study is number XXX of the Technical Series of the Biological Dynamics of

489 Forest Fragments (BDFFP-INPA). Finally, we would like to thank Frans Bongers, Julissa  
490 Roncal and one anonymous reviewer for their valuable comments on the manuscript.

491

#### 492 **Author Contribution**

493 FCD and CB conceived the study. FCD, GPA and CB designed the study with input from FC,  
494 GA, OLP, and HtS. FCD and JBS performed the analysis with input from CB, GPA, GA,  
495 OLP, AD, FCdS and KD. FCD wrote the manuscript with input from CB, FC, GA, OLP, AD,  
496 MJM, GPA and HtS. Initials refer to the first 14 authors and the last author. All other  
497 coauthors contributed data and had the opportunity to comment on the manuscript.

#### 498 **Data availability**

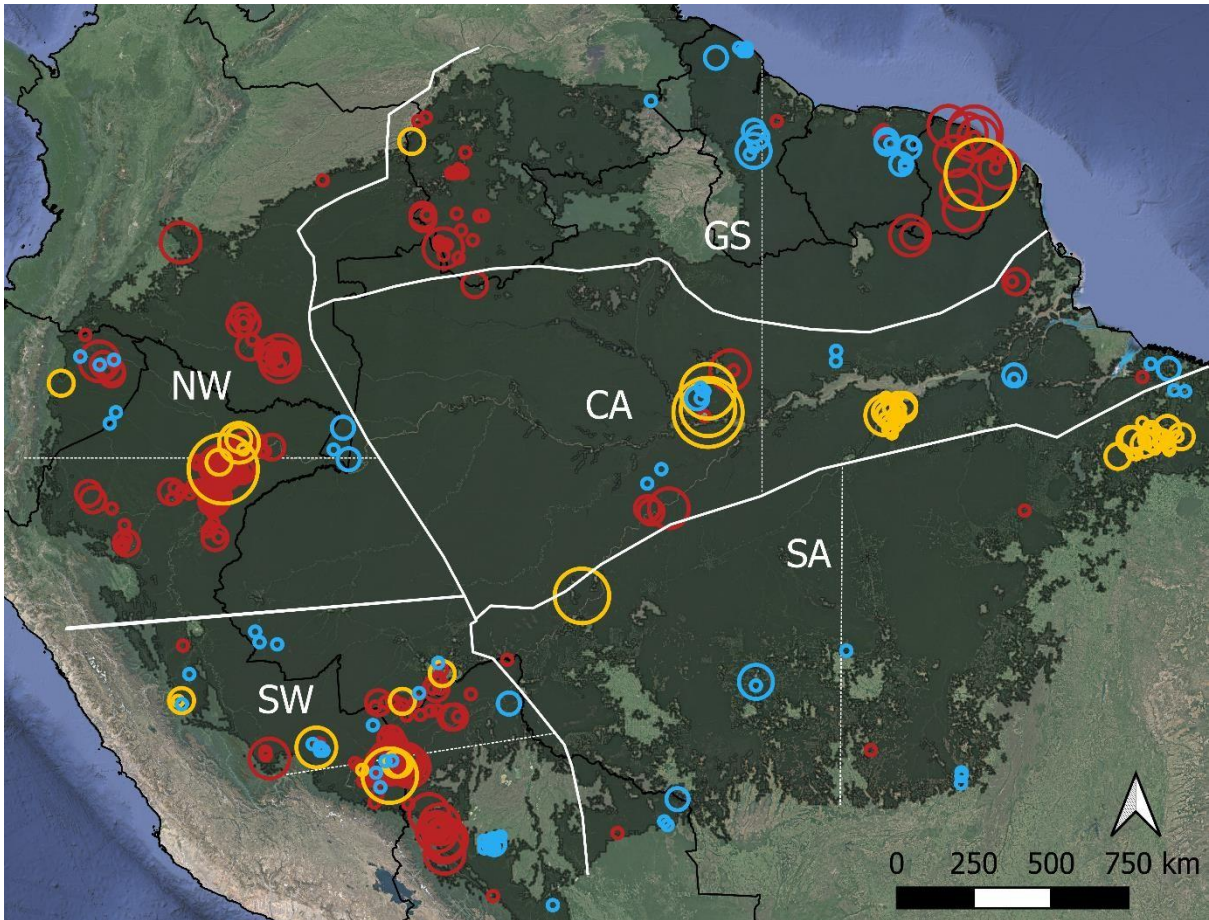
499 The datasets generated and/or analysed within this study are available from the corresponding  
500 author on reasonable request and with permission of relevant data owners.

#### 501 **Code availability**

502 All custom written analytical code used in this study are available online  
503 (<https://github.com/FreddieDraper/RedGentry>)

504

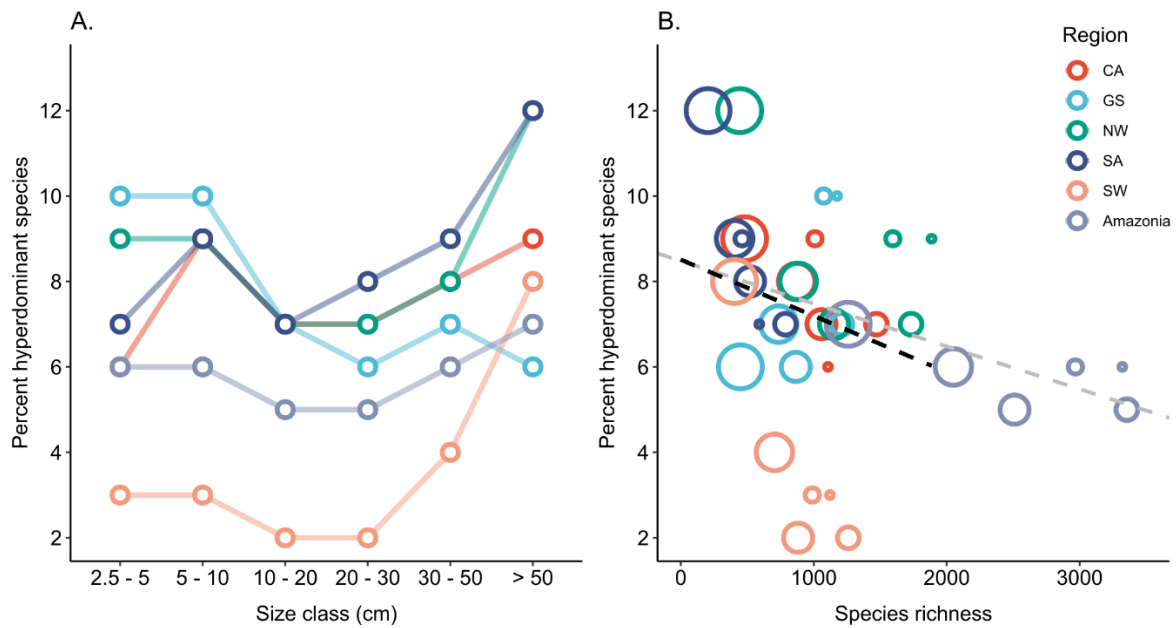
505



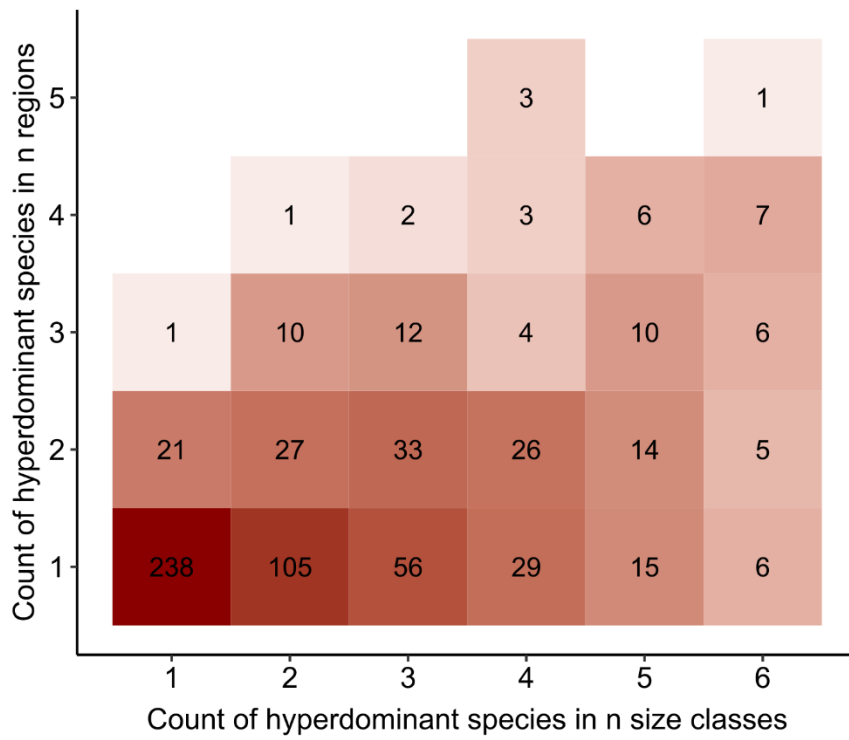
506

507 Figure 1. Map of study area and 1240 floristic inventory plots, represented by coloured  
508 points. Point size corresponds to number of plots at a given location (range 1 - 40 plots).  
509 Point colour refers to the plot size and diameter cut-off: 1. Red points: small plots < 1 ha and  
510 stems  $\geq 2.5$  cm; 2. Blue points: large plots > 1ha and stems  $\geq 10$  cm; 3. Gold points: large  
511 plots >1 ha and stems  $\geq 10$  cm with nested subplot for small stems  $\geq 2.5$  cm. Solid white lines  
512 indicate the border of the five sampling regions defined for analyses, dashed white lines show  
513 the further subdivision of sampling regions into 10 sampling zones. Sampling regions have  
514 been labelled as follows: Northwest Amazonia (NW); Southwest Amazonia (SW); Southern  
515 Amazonia (SA); Central Amazonia (CA); Guiana Shield (GS). The shaded area shows the

516 area defined as Amazonia based on: 1. Annual precipitation > 1300 mm, 2. Elevation < 1000  
517 m (above sea level), and 3. Forest cover > 70%.



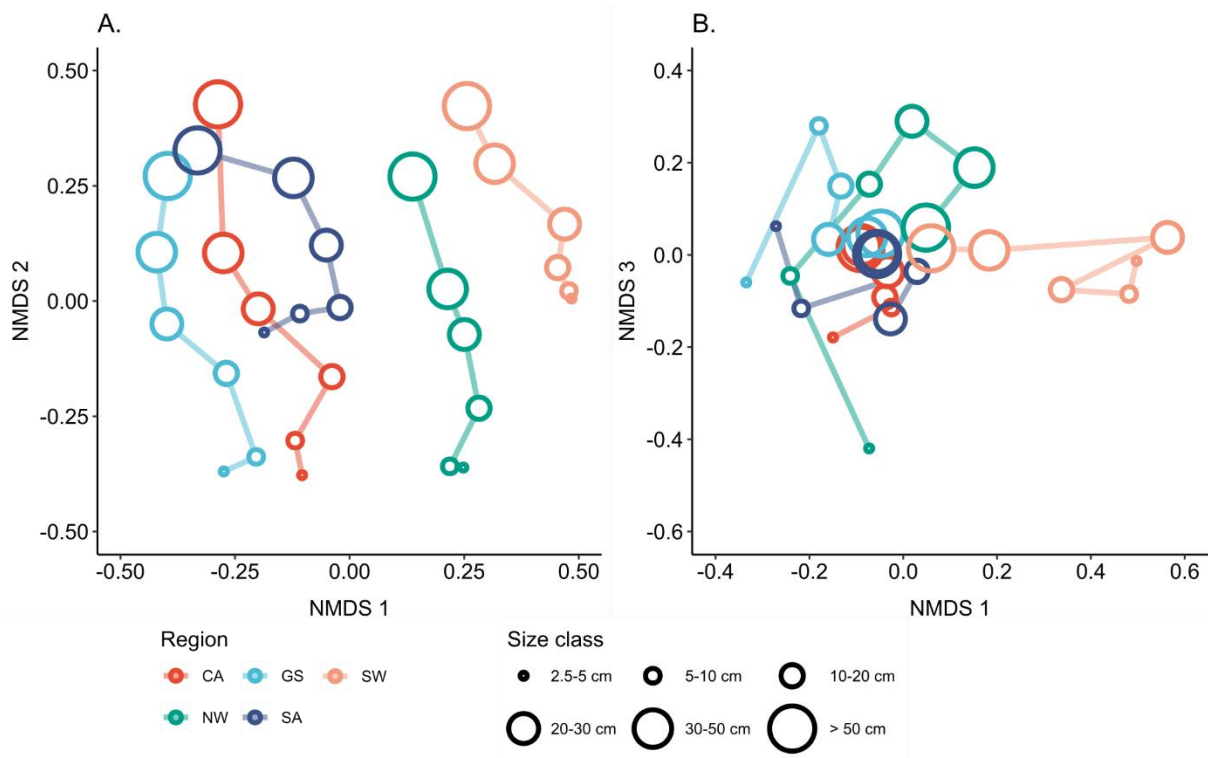
518  
 519 Figure 2. The proportion of species that are hyperdominant (i.e. together account for 50% of  
 520 individuals) within six size classes across the five Amazonian regions and the basin-wide  
 521 ‘Amazonia’ dataset (Panel A). The relationship between the proportion of species that are  
 522 hyperdominant and total species richness across six size classes (indicated by symbol size)  
 523 across the five Amazonian regions and the basin-wide ‘Amazonia’ dataset (Panel B). Dashed  
 524 lines show linear regressions based on the five regional data sets (black line;  $R^2 = 0.01$ ,  $P$   
 525  $=0.26$ ), and the five regional datasets plus the basin-wide ‘Amazonia’ dataset (grey line;  $R^2 =$   
 526  $0.08$ ,  $P = 0.05$ ). In both panels, a lower proportion of hyperdominant species indicates  
 527 stronger patterns of dominance. Sampling regions as indicated in Figure 1 are: Northwest  
 528 Amazonia (NW); Southwest Amazonia (SW); Southern Amazonia (SA); Central Amazonia  
 529 (CA); Guiana Shield (GS).



484

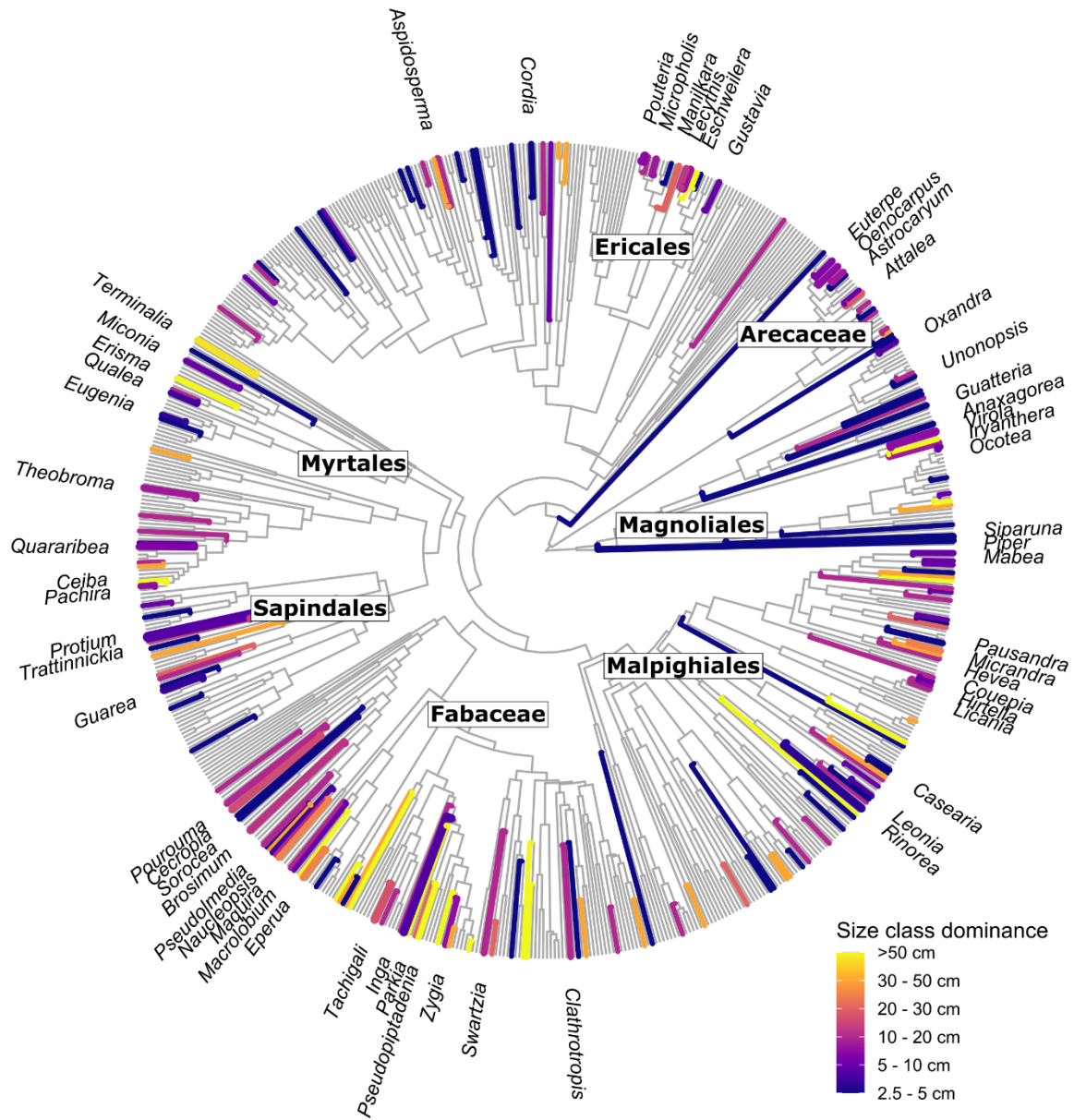
485 Figure 3. Two-dimensional histogram showing the number of species that are hyperdominant in one to six size classes and across one to five regions. Regions and size classes follow the same definitions as in figure 2.





486

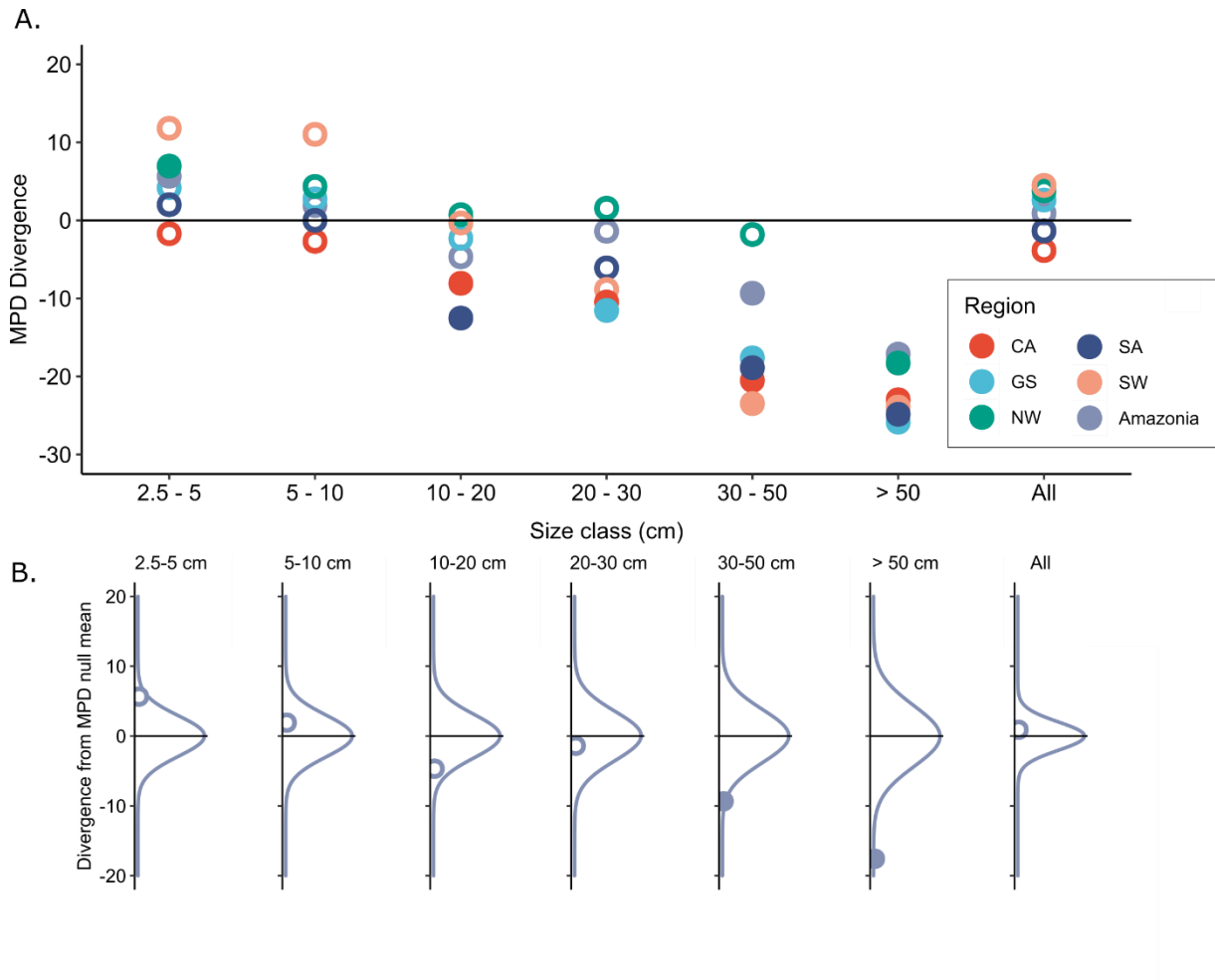
487 Figure 4. NMDS ordinations showing similarity in composition of hyperdominant species in  
 488 terms of: A.) Taxonomic species similarity (Jaccard index) and B) deep-node-weighted  
 489 phylogenetic similarity (community level mean pairwise phylogenetic distance). Sampling  
 490 regions have been labelled as follows: Northwest Amazonia (NW); Southwest Amazonia  
 491 (SW); Southern Amazonia (SA); Central Amazonia (CA); Guiana Shield (GS).



492

493 Figure 5. Hyperdominant species mapped onto a genus level Amazonian tree phylogeny. All  
 494 genera with one or more hyperdominant species have been highlighted. Genera with three or  
 495 more hyperdominant species have been labelled. Colour corresponds to the size class within  
 496 which species belonging to that genus are most frequently hyperdominant.

497



498

499 Figure 6. Divergence of mean pairwise phylogenetic distance (MPD) from null models for  
 500 each hyperdominant community (Panel A). Positive numbers indicate greater MPD than  
 501 expected by chance (i.e. species are more distantly related than expected by chance.).

502 Negative numbers indicate lower MPD than expected by chance (i.e. species are more closely  
 503 related than expected by chance). Filled symbols indicate hyperdominant communities that  
 504 were outside the 95 % confidence interval of the null distribution. Panel B shows the null  
 505 distributions and observed MPD for entire Amazonia hyperdominant communities. Regional  
 506 null distributions are provided in Figure S4.

507

## 508 **Methods**

### 509 *Floristic data*

510 Floristic data came from 1240 forest inventory plots, distributed across the Amazon basin  
511 (Fig. 1). The plot data fall into two broad categories: Firstly, the RedGentry network  
512 consisted of 1027 small forest plots (typically 0.1ha but ranging from 0.04- 0.25 ha) within  
513 which all stems with a diameter at 1.3 m in height (dbh) greater than 2.5 cm were measured  
514 and identified. Secondly, 520 larger forest plots (typically 1 ha but ranging from 0.5 to 9 ha)  
515 from the RAINFOR and ATDN networks were used. Within these larger plots all stems with  
516 a dbh greater than 10 cm were measured and identified. Many of these plots are curated and  
517 stored within ForestPlots.net, a cyber-infrastructure initiative that unites plot records and their  
518 contributing scientists from the world's tropical forests.

519 The RedGentry plot data came from a range of sources and therefore included a range of plot  
520 sizes and sampling protocols. Most plots were 0.1 ha in size and consisted of 10 transects of 2  
521 X 50 m arranged systematically around a single transect baseline following the 'Gentry  
522 protocol'<sup>46</sup>. However, 307 plots were subplots nested in within larger 1 ha inventory plots  
523 (Fig. 1). The majority of these nested 0.1 ha plots were part of the PPBio network.

### 524 *Taxonomic standardization*

525 It was not possible to standardise morphospecies across datasets as plots were installed by  
526 many different botanical teams at different times, often without accompanying herbarium  
527 vouchers. Therefore, all individuals that were not identified to species level were excluded  
528 from all subsequent analysis. These exclusions lead to a substantial loss of individuals (mean  
529 21 % of individuals per plot, Figure S.7) and were phylogenetically biased, i.e. some families  
530 had a higher degree of taxonomic uncertainty than others. Nevertheless, this approach renders

531 our analysis comparable to other landscape analyses conducted on larger stems in this and  
532 other regions.

533  
534 *Species exclusions*

535 Species names were checked for synonymy and spelling mistakes using the taxonomic names  
536 resolution service (TNRS) using the R package *taxize*<sup>59</sup>. Any species that were not recognised  
537 in the automated process were checked manually for spelling mistakes. Identifications that  
538 could not be easily assigned to a species were considered unidentified morphospecies and  
539 were removed from further analysis. Finally, our list of legitimate species names was cross-  
540 checked against the most current published checklists<sup>60,61</sup>. Species that did not occur on this  
541 Amazon checklist (887 species) were checked manually against collection records in the  
542 Tropicos database<sup>62</sup>. 39 of these were confirmed to be illegitimate Amazonian species  
543 because they have ranges either outside of our region (i.e. on another continent). A further  
544 579 species that were described as either epiphytes, lianas, herbs, or ferns were also excluded  
545 from our analysis. These lifeforms were included in some plot datasets and excluded from  
546 others. As individual datasets are normally geographically clustered, including them would  
547 likely lead to spatially biased species abundance estimates. A further 47 species were  
548 excluded because there was no recorded collection since their descriptions, we considered  
549 these individuals to be wrongly identified.

550 *Species inclusions*

551 We included 180 species in our analysis that had been excluded from previous analyses or  
552 checklists. The majority of these inclusions were small stemmed species that had previously  
553 been excluded for being shrubs or treelets. We considered these definitions to be subjective.  
554 37 of these included species have previously been considered illegitimate because they occur  
555 primarily in Savanna or seasonally dry habitats. However, because several of our plots were

556 located close to boundaries between ecosystem types, and many species are shared across  
557 these boundaries (not necessarily in their optimum habitat), we included these non-core  
558 rainforest species in our dataset.

### 559 *Defining Amazonia*

560 To ensure that our analysis included only plots located in lowland Amazonia and did not  
561 incorporate plots in marginal seasonally dry or montane environments, it was necessary to  
562 define our study area. We defined lowland Amazonia using four remotely sensed criteria: 1.  
563 Watersheds were estimated using the hydrosheds data layer<sup>63</sup>, in addition to the Amazon  
564 basin; we also included eastern branches of the Orinoco and all watersheds to the east of that  
565 mark in the Guiana Shield. 2. Elevation was measured using the global SRTM digital  
566 elevation model<sup>64</sup>, lowlands were considered to be land area below 1000m elevation  
567 following Cardoso *et al.* (2017)<sup>60</sup>. 3. Precipitation was estimated using the CHIRPS annual  
568 mean rainfall data<sup>65</sup>, and a minimum mean annual precipitation value of 1300 mm year<sup>-1</sup> was  
569 used to define moist forests following Cardoso *et al.* (2017)<sup>60</sup>. 4. Tree cover was estimated  
570 using the 2010 global forest cover map<sup>66</sup>, and all pixels with > 70 % forest cover were  
571 included. The four layers were overlaid, and the intersecting area was used to define  
572 Amazonia. This final layer was then sieved and filtered to reduce speckle, which was  
573 primarily driven by the complex patterns of deforestation along the southeastern border. All  
574 geospatial analysis was conducted using QGIS software<sup>67</sup>.

### 575 *Spatial standardisation*

576 Many species (27 %) occurred fewer than five times across the plot network. Therefore, we  
577 did not attempt to generate basin-wide population estimates for most species as other studies  
578 have done<sup>7</sup>. Instead, we used only the empirical data from plots to estimate those species

579 likely to be hyperdominant at basin-wide scales, under the assumption that this plot network  
580 is reasonably representative of abundances of the most common species.

581 Plots are not distributed evenly across Amazonia, but instead are clustered in space, for  
582 example, there are many more plots in western Amazonia than in Southern or Eastern  
583 Amazonia. Furthermore, plots varied in size and therefore so did the number of individuals  
584 per plot. To account for these biases, and to attempt to ensure the Amazonian flora was  
585 sampled as evenly as possible, we used a spatially-stratified bootstrap resampling approach.  
586 All sampling procedures were performed in the statistical language R using the tidyverse  
587 packages dplyr, tidyr, and purrr<sup>68–70</sup>.

588 This approach consisted of the following steps:

589 1.) Greater Amazonia (as defined above) was divided into 5 regions roughly following  
590 previously defined boundaries<sup>7,49</sup>. Each region was then split roughly in half to generate 10  
591 total sampling zones that were broadly similar in area (Area varied from 210,000 to 1081,000  
592 km<sup>2</sup>). Each sampling zone contained at least 40 individual plots (at least 20 small plots and at  
593 least 10 large plots).

594 2.) The entire dataset was then divided into six strata-specific datasets. This was done by  
595 stratifying the data by dbh into six size classes (2.5 – 5 cm, 5 – 10 cm, 10 – 20 cm, 20 – 30  
596 cm, 30 – 50 cm, >50 cm). Diameter was used as a proxy for tree height because tree height  
597 was not measured in most plots, and because of the strong allometric relationship between  
598 diameter and height.

599 3.) 20 small plots or large plots with nested subplots and five large plots were sampled  
600 from each sampling zone at random without replacement. This step ensured spatially even  
601 sampling across the basin, and the five additional large plots ensured a reasonable number of  
602 large individuals were sampled.

- 603 4.) From each of these 25 plots a standard number of individuals (50% of the median  
604 individuals per plot per size class) were sampled with replacement, ensuring an even number  
605 of individuals was sampled for each plot.
- 606 5.) These standardised samples from each plot were then assembled into a single species  
607 by plot matrix.
- 608 6.) Steps 2 and 3 were repeated  $10^6$  times, generating  $10^6$  estimates of abundance for each  
609 species across the basin.
- 610 7.) The mean and standard deviation of abundance for each species was calculated across  
611 the  $10^6$  estimates of abundance.
- 612 8.) Hyperdominant species were then defined as those species that together account for  
613 50 % of the mean total abundance of all individuals within each size class across all  
614 iterations.
- 615 To identify regionally dominant species, steps two-seven were repeated for each of the five  
616 predefined regions individually.

#### 617 *Phylogenetic analyses*

618 To understand where hyperdominant species are situated across the Amazonian phylogeny,  
619 we used a published genus-level molecular phylogeny for Amazonian tree species<sup>37,38</sup>. A  
620 genus-level phylogeny was used because a species-level molecular phylogeny for the full  
621 Amazonian flora does not yet exist. Genera occurring in our lists of Amazon-wide  
622 hyperdominant species were mapped onto the phylogeny, which was then pruned to remove  
623 taxa not occurring in our dataset. The final phylogenetic tree contained 646 genus tips. We  
624 then plotted the phylogeny for all genera occurring in our dataset using the R package  
625 *ggtree*<sup>71</sup>.



626 The tips of genera that contained hyperdominant species were coloured to highlight their  
627 location. Tip colours corresponded to a continuous variable that was the mean size class for  
628 hyperdominant species that were in the given genus. Genus labels were given to all genera  
629 that contained three or more hyperdominant species.

630 We used the mean pairwise phylogenetic distance (MPD) metric and a null modelling  
631 approach to test if hyperdominant species are more closely related to one another than would  
632 be expected if hyperdominance were distributed across the phylogeny at random<sup>51,72</sup>.

633 Because a species-level molecular phylogeny has not been developed across Amazonian plant  
634 taxa, we first added species tips with a uniform branch length (0.1) to all tree genera. This  
635 approach allowed us to make species-level comparisons using a genus level phylogeny, while  
636 minimizing the assumptions made about within-genus phylogenetic structure. We then  
637 calculated MPD among hyperdominant species for each hyperdominant community, and  
638 compared this observed MPD to a null distribution of expected MPD if we sampled an  
639 equivalent number of species at random across a phylogeny with an equivalent species pool<sup>72</sup>.

640 Where the observed MPD fell outside two standard deviations of the null distribution, that  
641 hyperdominant community was considered to be significantly more clustered (lower MPD) or  
642 significantly more dispersed (higher MPD) than we expect by chance. All phylogenetic  
643 analysis was conducted in R, using packages *phylomeasures*, *phytools*, and *caper*<sup>72-74</sup>.

644 MPD is known to be influenced by the extent to which species are divided among the three  
645 major angiosperm clades (Magnoliids, Monocots and Eudicots)<sup>75</sup>. Large stemmed Amazonian  
646 tree species are predominantly found within the Eudicots, while small stemmed species are  
647 found across the three clades. These deep-clade distributions are therefore likely to increase  
648 phylogenetic clustering within the large-stemmed species and increase phylogenetic  
649 overdispersion within small stemmed species. In part we account for this in measurement of  
650 MPD as we remove genera from the tree that do not occur in the size class/region for which

651 we are measuring MPD. However, to explore the effect of this deep-clade diversity further,  
652 we repeated our MPD analysis within Eudicots only (Figure S.5). This analysis demonstrates  
653 that the overall patterns of increased clustering in larger size classes is maintained within  
654 eudicots. The analysis also shows that the phylogenetic dispersion found within small-  
655 stemmed hyperdominant species is due to these understory hyperdominant species occurring  
656 across these deep phylogenetic nodes.

### 657 *Compositional analyses*

658 To understand how the composition of hyperdominant species varied across size classes and  
659 regions we used a multivariate statistical approach. Specifically, we used the Jaccard index as  
660 a metric of how similar or different the composition of hyperdominant species was among the  
661 36 communities of a given size classes within a given region, e.g. northwest amazon - 2.5-5  
662 cm size class. For clarity, these regional and size class specific groups are hereafter referred  
663 to as hyperdominant communities. The Jaccard distances were generated using with the R  
664 package *vegan*<sup>76</sup>.

665 We expanded these compositional analyses not only to consider how taxonomic composition  
666 varied among hyperdominant communities, but also, to quantify how phylogenetically similar  
667 hyperdominant communities were among size classes and regions. To do this we again used a  
668 multivariate statistical approach, however, this time using two metrics of phylogenetic beta  
669 diversity. Both phylo-beta diversity metrics were calculated at genus rather than species level  
670 as we used the genus-level phylogeny. To account for the fact that some genera contain  
671 several hyperdominant species, we used the number of hyperdominant species per genus for  
672 each hyperdominant community per size class as our input community matrix.

673 The first metric that we used was the abundance weighted MPD among hyperdominant  
674 communities, which provides a deep/basal node weighted assessment of phylogenetic beta

675 diversity<sup>50</sup>. The second metric of phylogenetic beta diversity that we used was the generalized  
676 version the unifrac method<sup>77</sup>, calculated with the R package *GUniFrac*<sup>78</sup>. We used an  $\alpha$  value  
677 of 0.5, meaning that we moderately weighted genera by the number of hyperdominant species  
678 that they contained in that site/size class. The unifrac metric provides a stable tip-weighted  
679 assessment of phylogenetic beta-diversity.

680 Because the phylogenetic analysis was conducted using a genus-level phylogeny, we do not  
681 account for any within genus phylogenetic structure that could affect these metrics. However,  
682 any within-genus structure will have little effect on patterns of MPD, as this metric is heavily  
683 weighted towards deep-node differences among communities<sup>50</sup>. The tip-weighted unifrac  
684 method is likely to be more heavily influenced by the missing within-genus structure,  
685 therefore these results are only presented in the supplementary information.

686 To reduce the dimensionality of this multivariate data and visualize the taxonomic and  
687 phylogenetic similarities among hyperdominant communities we used Non-metric  
688 multidimensional scaling (NMDS). NMDS analyses were run for at least 50 iterations and  
689 until a stable solution was reached (stress < 0.2). Each NMDS was optimized over three  
690 dimensions and displayed in an ordination plots. All NMDS ordinations were  
691 performed in the R package *vegan*<sup>76</sup>.  
692

693 **References**

- 694 1. Sakschewski, B. *et al.* Resilience of Amazon forests emerges from plant trait diversity.  
695 *Nat. Clim. Chang.* **6**, 1032–1036 (2016).
- 696 2. Poorter, L. *et al.* Diversity enhances carbon storage in tropical forests. *Glob. Ecol.*  
697 *Biogeogr* **24**, 1314–1328. (2015).
- 698 3. Brienen, R. J. W. *et al.* Long-term decline of the Amazon carbon sink. *Nature* **519**, 344–  
699 348 (2015).
- 700 4. Spracklen, D. V., Baker, J. C. A., Garcia-Carreras, L. & Marsham, J. H. The Effects of  
701 Tropical Vegetation on Rainfall. *Annu. Rev. Environ. Resour.* **43**, 193–218 (2018).
- 702 5. Staal, A. *et al.* Forest-rainfall cascades buffer against drought across the Amazon. *Nat.*  
703 *Clim. Chang.* **8**, 539–543 (2018).
- 704 6. ter Steege, H. *et al.* The discovery of the Amazonian tree flora with an updated checklist  
705 of all known tree taxa. *Sci. Rep.* **6**, 29549 (2016).
- 706 7. ter Steege, H. *et al.* Hyperdominance in the Amazonian tree flora. *Science.* **342**, 6156  
707 (2013).
- 708 8. Wilson, E. O. Biodiversity research requires more boots on the ground. *Nat. Ecol. Evol.* **1**,  
709 1590–1591 (2017).
- 710 9. Gomes, V. H. F., Vieira, I. C. G., Salomão, R. P. & ter Steege, H. Amazonian tree species  
711 threatened by deforestation and climate change. *Nat. Clim. Chang.* **9**, 547–553 (2019).
- 712 10. Fauset, S. *et al.* Hyperdominance in Amazonian forest carbon cycling. *Nat Commun* **6**, 1–  
713 9 (2015).
- 714 11. Draper, F. C. *et al.* Dominant tree species drive beta diversity patterns in western  
715 Amazonia. *Ecology* **100**, 02636 (2019).
- 716 12. Pitman, N. C. A. *et al.* Dominance and distribution of tree species in upper Amazonian  
717 terra firme forests. *Ecology* **82**, 2101–2117 (2001).

- 718 13. Pitman, N. C. A., Silman, M. R. & Terborgh, J. W. Oligarchies in Amazonian tree  
719 communities: a ten-year review. *Ecography*. **36**, 114–123 (2013).
- 720 14. Honorio Coronado, E. N. *et al.* Multi-scale comparisons of tree composition in  
721 Amazonian terra firme forests. *Biogeosciences* **6**, 2719–2731 (2009).
- 722 15. Pitman, N. C. A. *et al.* Distribution and abundance of tree species in swamp forests of  
723 Amazonian ecuador. *Ecography*. **37**, 902–915 (2014).
- 724 16. Gentry, A. H. & Emmons, L. H. Geographical Variation in Fertility, Phenology, and  
725 Composition of the Understory of Neotropical Forests. *Biotropica* **19**, 216 (1987).
- 726 17. Gentry, A. H. Patterns of Neotropical Plant Species Diversity. in *Evolutionary Biology* 1–  
727 84 (Springer US, 1982).
- 728 18. Gentry, A. H. & Dodson, C. Contribution of Nontrees to Species Richness of a Tropical  
729 Rain Forest. *Biotropica* **19**, 149 (1987).
- 730 19. Duque, A. *et al.* Insights into regional patterns of Amazonian forest structure, diversity,  
731 and dominance from three large terra-firme forest dynamics plots. *Biodivers. Conserv.* **26**,  
732 669–686 (2017).
- 733 20. Duque, A., Sánchez, M., Cavelier, J. & Duivenvoorden, J. F. Different floristic patterns of  
734 woody understorey and canopy plants in Colombian Amazonia. *J. Trop. Ecol.* **18**, 499–  
735 525 (2002).
- 736 21. Arellano, G. *et al.* Oligarchic patterns in tropical forests: role of the spatial extent,  
737 environmental heterogeneity and diversity. *J. Biogeogr.* **43**, 616–626 (2016).
- 738 22. Macía, M. J. & Svenning, J.-C. Oligarchic dominance in western Amazonian plant  
739 communities. *J. Trop. Ecol.* **21**, 613–626 (2005).
- 740 23. Vormisto, J., Svenning, J., Hall, P. & Balslev, H. Diversity and dominance in palm  
741 (Arecaceae) communities in terra firme forests in the western Amazon basin. *J. Ecol.* **92**,  
742 577–588 (2004).

- 743 24. Burnham, R. J. Dominance, diversity and distribution of lianas in Yasuní, Ecuador: who  
744 is on top? *J. Trop. Ecol.* **18**, 845–864 (2002).
- 745 25. Farrior, C. E., Bohlman, S. A., Hubbell, S. & Pacala, S. W. Dominance of the suppressed:  
746 Power-law size structure in tropical forests. *Science*. **351**, 155–157 (2016).
- 747 26. Baker, T. R. *et al.* Consistent, small effects of treefall disturbances on the composition  
748 and diversity of four Amazonian forests. *J. Ecol.* **104**, 497–506 (2016).
- 749 27. ter Steege, H. *et al.* Continental-scale patterns of canopy tree composition and function  
750 across Amazonia. *Nature* **443**, 444–447 (2006).
- 751 28. Quesada, C. A. *et al.* Basin-wide variations in Amazon forest structure and function are  
752 mediated by both soils and climate. *Biogeosciences* **9**, 2203–2246 (2012).
- 753 29. Dexter, K. & Chave, J. Evolutionary patterns of range size, abundance and species  
754 richness in Amazonian angiosperm trees. *PeerJ* **4**, e2402 (2016).
- 755 30. Coelho de Souza, F. *et al.* Evolutionary heritage influences Amazon tree ecology. *Proc.*  
756 *Royal Soc. Lond. B.* **283** (2019).
- 757 31. , Malhi, Y., & Wright J., Spatial patterns and recent trends in the climate of tropical  
758 rainforest regions. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, **359** 311–329 (2004).
- 759 32. Thomson, F. J. *et al.*. Seed dispersal distance is more strongly correlated with plant height  
760 than with seed mass. *J. Ecol.* **99**, 1299–1307 (2011).
- 761 33. Thomson, F. J. *et al.* Can dispersal investment explain why tall plant species achieve  
762 longer dispersal distances than short plant species? *New Phytol.* **217**, 407–415 (2018).
- 763 34. Dexter, K. G. *et al.* Dispersal assembly of rain forest tree communities across the Amazon  
764 basin. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2645–2650 (2017).
- 765 35. Rüger, N. *et al.* Beyond the fast-slow continuum: demographic dimensions structuring a  
766 tropical tree community. *Ecol. Lett.* **21**, 1075–1084 (2018).

- 767 36. Baker, T. R. *et al.* Fast demographic traits promote high diversification rates of  
768 Amazonian trees. *Ecol. Lett.* **17**, 527–536 (2014).
- 769 37. Coelho de Souza, F. *et al.* Evolutionary diversity is associated with wood productivity in  
770 Amazonian forests. *Nat. Ecol. Evol.* **3**, 1754–1761 (2019).
- 771 38. Neves, D. M. *et al.* Evolutionary diversity in tropical tree communities peaks at  
772 intermediate precipitation. *Sci. Rep.* **10**, 1188 (2020).
- 773 39. Damasco, G., Daly, D. C., Vicentini, A. & Fine, P. V. A. Reestablishment of *Protium*  
774 *cordatum* (Burseraceae) based on integrative taxonomy. *Taxon* **68**, 34–46 (2019).
- 775 40. Roncal, R. *et al.* Palm diversification in two geologically contrasting regions of western  
776 Amazonia. **42** 1503–1513 (2015).
- 777 41. Draper F.C. *et al.* Quantifying Tropical Plant Diversity Requires an Integrated  
778 Technological Approach. *Trends Ecol. Evol.* **35** 1100–1109 (2020).
- 779 42. Junk, W. J. *et al.* A classification of major naturally-occurring Amazonian lowland  
780 wetlands. *Wetlands* **31**, 623–640 (2011).
- 781 43. Adeney, J. M., Christensen, N. L., Vicentini, A. & Cohn-Haft, M. White-sand  
782 Ecosystems in Amazonia. *Biotropica* **48**, 7–23 (2016).
- 783 44. Tuomisto, H., Ruokolainen, K. & Yli-Halla, M. Dispersal, environment, and floristic  
784 variation of Western Amazonian forests. *Science.* **299**, 241–244 (2003).
- 785 45. Baraloto, C. *et al.* Rapid Simultaneous Estimation of Aboveground Biomass and Tree  
786 Diversity Across Neotropical Forests: A Comparison of Field Inventory Methods.  
787 *Biotropica* **45**, 288–298 (2013).
- 788 46. Phillips, O. L. *et al.* Efficient plot-based floristic assessment of tropical forests. *J. Trop.*  
789 *Ecol.* **19**, 629–645 (2003).
- 790 47. Magnusson, W. E. *et al.* RAPELD: a modification of the Gentry method for biodiversity  
791 surveys in long-term ecological research sites. *Biota Neotrop.* **5**, 19–24 (2005).

- 792 48. Draper, F. C. *et al.* Imaging spectroscopy predicts variable distance decay across  
793 contrasting Amazonian tree communities. *J. Ecol.* **107**, 696–710 (2019).
- 794 49. Tuomisto, H. *et al.* Discovering floristic and geocological gradients across Amazonia. *J.*  
795 *Biogeogr.* **46**, 1734–1748 (2019).
- 796 50. Swenson, N. G. The assembly of tropical tree communities – the advances and  
797 shortcomings of phylogenetic and functional trait analyses. *Ecography* **36**, 264–276  
798 (2013).
- 799 51. Webb, C. O. Exploring the Phylogenetic Structure of Ecological Communities: An  
800 Example for Rain Forest Trees. *Am. Nat.* **156**, 145–155 (2000).
- 801 52. Baraloto, C. *et al.* Using functional traits and phylogenetic trees to examine the assembly  
802 of tropical tree communities. *J. Ecol.* **100**, 690–701 (2012).
- 803 53. Poorter, L. *et al.* The importance of wood traits and hydraulic conductance for the  
804 performance and life history strategies of 42 rainforest tree species. *New Phytol.* **185**,  
805 481–492 (2010).
- 806 54. Rowland, L. *et al.* Death from drought in tropical forests is triggered by hydraulics not  
807 carbon starvation. *Nature* **528**, 119–122 (2015).
- 808 55. Gonzalez-Caro, S. *et al.* Scale-dependent drivers of the phylogenetic structure and  
809 similarity of tree communities in northwestern Amazonia. *J. Ecol.* (2020)
- 810 56. Levis, C. *et al.* Persistent effects of pre-Columbian plant domestication on Amazonian  
811 forest composition. *Science* **355**, 925–931 (2017).
- 812 57. McMichael, C. N. H., Matthews-Bird, F., Farfan-Rios, W. & Feeley, K. J. Ancient human  
813 disturbances may be skewing our understanding of Amazonian forests. *Proc. Natl. Acad.*  
814 *Sci. U. S. A.* **114**, 522–527 (2017).
- 815 58. Levis, C. *et al.* How people domesticated Amazonian forests. *Front. Ecol. Evol.* **5**, 171  
816 (2018).



- 817 59. Chamberlain, S. *et al.* taxize: Taxonomic information from around the web. *R package*  
818 *version 0.9.95.* (2019)
- 819 60. Cardoso, D. *et al.* Amazon plant diversity revealed by a taxonomically verified species  
820 list. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 10695–10700 (2017).
- 821 61. ter Steege, H. *et al.* Towards a dynamic list of Amazonian tree species. *Sci. Rep.* **9**, 3501  
822 (2019).
- 823 62. Tropicos.org. Missouri Botanical Garden. 26 November 2019. <http://www.tropicos.org>.
- 824 63. Lehner, B., Verdin, K. & Jarvis, A. New Global Hydrography Derived From Spaceborne  
825 Elevation Data. *Eos, Trans. Am. Geophys. Union* **89**, 93 (2008).
- 826 64. Jarvis, A., Reuter, H. I., Nelson, A. & Guevara, E. Hole-filled SRTM for the globe  
827 Version 4. *available from CGIAR-CSI SRTM 90m Database (<http://srtm.csi.cgiar.org>)*  
828 (2008).
- 829 65. Funk, C. *et al.* The climate hazards infrared precipitation with stations - A new  
830 environmental record for monitoring extremes. *Sci. Data* **2**, (2015).
- 831 66. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change.  
832 *Science*. **342**, 850–853 (2013).
- 833 67. QGIS Development Team. QGIS Geographic Information System. Open Source  
834 Geospatial Foundation Project. (2019).
- 835 68. Wickham, H., Romain, F., Henry, L. & Müller, K. dplyr: A Grammar of Data  
836 Manipulation. *R package version 0.8.3.* (2019).
- 837 69. Wickham, H. & Henry, L. tidyr: Easily Tidy Data with ‘spread()’ and ‘gather()’  
838 Functions. *R package version 0.8.3.* (2019).
- 839 70. Henry, L. & Wickham, H. purrr: Functional Programming Tools. *R package version*  
840 *0.8.3.* (2019).

- 841 71. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. ggtree : an r package for  
842 visualization and annotation of phylogenetic trees with their covariates and other  
843 associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
- 844 72. Tsirogiannis, C. & Sandel, B. PhyloMeasures: a package for computing phylogenetic  
845 biodiversity measures and their statistical moments. *Ecography* **39** 709–714
- 846 73. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other  
847 things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
- 848 74. Orme, D. *et al.* caper: Comparative Analyses of Phylogenetics and Evolution in R. *R*  
849 *package version 1.0.1.* (2018).
- 850 75. Honorio Coronado E. N. *et al.* Phylogenetic diversity of Amazonian tree communities.  
851 *Diversity and Distributions.* **21** 1295-1307 (2015)
- 852 76. Oksanen, J. *et al.* vegan: Community Ecology Package. *R package version 2.5-6.* (2019).
- 853 77. Chen, J. *et al.* Associating microbiome composition with environmental covariates using  
854 generalized UniFrac distances. *Bioinformatics* **28**, 2106–2113 (2012).
- 855 78. Chen, J. GUniFrac: Generalized UniFrac Distances. *R package version 1.1.* (2018).