



HAL
open science

Spatial patterns of soil phytoliths in a wet vs. dry neotropical forest: Implications for paleoecology

Camilla Crifò, Caroline A.E. Strömberg

► To cite this version:

Camilla Crifò, Caroline A.E. Strömberg. Spatial patterns of soil phytoliths in a wet vs. dry neotropical forest: Implications for paleoecology. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 2021, 562, pp.110100 -. 10.1016/j.palaeo.2020.110100 . hal-03493146

HAL Id: hal-03493146

<https://hal.science/hal-03493146>

Submitted on 2 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Spatial patterns of soil phytoliths in a wet vs. dry Neotropical forest: implications for paleoecology**

2

3 Camilla Crifo^{a,b,c*} & Caroline A. E. Strömberg^{a**}

4 ^aDepartment of Biology and Burke Museum of Natural History & Culture, University of

5 Washington, Seattle, WA 98195, USA

6 ^bÉcole Pratique des Hautes Études, PSL Research University, Les Patios Saint-Jacques 4-14

7 rue Ferrus, 75014, Paris, France

8 ^cISEM, Université Montpellier, UMR 5554, C/C 065, Place Eugène Bataillon, 34095, Montpellier

9 Cedex 05, France

10

11

12 *Correspondence author. E-mail: camilla.crifo@ephe.sorbonne.fr

13 **Email: caestrom@uw.edu

14

15 **Abstract**

16 To improve paleoenvironmental reconstruction based on fossil phytoliths, it is vital to establish
17 modern calibrations that explicitly consider important functional aspects of vegetation, such as
18 structure, composition, diversity, and spatial heterogeneity. No such studies currently exist for Central-
19 and South America. To begin to solve this problem, we address two questions: 1) do phytolith
20 assemblages reflect the overall differences in structure, composition, and diversity between rainforest
21 and dry forest? and 2) do phytolith assemblages from lateral transects capture variation in structure
22 and composition typical of these vegetation types? At La Selva (rainforest) and Palo Verde (dry forest)
23 biological stations (Costa Rica), we collected soil phytolith samples from, respectively, fourteen and
24 eleven quadrats along vegetation transects. Along each transect, we measured gradients in canopy
25 openness (LAI), elevation, and moisture. Statistical analyses were used to characterize and compare

26 structure, composition, and diversity to soil phytolith assemblage composition, and to test for
27 correlation with environmental gradients at each site. The two sites were then compared using
28 multivariate statistical methods. Results indicate that rainforest and dry forest vegetation can be
29 distinguished by analyzing multiple phytolith samples collected within sites. Within each vegetation
30 type, the comparison of phytolith assemblages along transects allows reconstruction of important
31 structural (canopy openness and patchiness) and compositional (dominant plant functional types)
32 aspects of the vegetation. Thus, the analysis of phytolith samples collected along temporally
33 constrained stratigraphic levels is a viable and vital tool for inferring heterogeneity in canopy cover and
34 gross vegetation composition in the deep-time fossil record.

35

36 **Keywords:** Modern analogue; vegetation structure LAI; *F/t* ratio; vegetation composition; vegetation
37 heterogeneity.

38 **1. Introduction**

39 Phytoliths are microscopic silica bodies produced in many plants by absorption of aqueous
40 monosilicic acid (H_4O_4Si) through the roots, and precipitation as opal-A in plant cells and tissues
41 (Piperno, 1988). Analysis of phytolith assemblages preserved in fossilized soil horizons or lake sediment
42 has, in the last two decades, emerged as a powerful tool in paleoecology with several advantages over
43 more traditional sources of paleobotanical information, such as macrofossils (e.g., leaves, fruits) and
44 palynomorphs (pollen and spores) (Table 1).

45 First, because of their relatively inert composition, phytoliths expand the range of depositional
46 environments that can be sampled to well-oxidized 'uplands,' which do not typically preserve organic
47 plant fossils (Strömberg et al., 2018). Second, despite the polymorphism ('multiplicity') and homoplasy
48 or plesiomorphy ('redundancy') inherent in phytolith production (Piperno, 1988; Rovner, 1971), limiting
49 the taxonomic information they offer, phytolith assemblages capture key vegetation information for
50 paleoenvironmental reconstruction. For example, the abundances of open-habitat grasses vs. woody
51 plants is often accurately reflected in soil phytolith assemblages across many ecosystems, forming a
52 proxy for habitat openness (Barboni et al., 2007; Bremond et al., 2005a; but see Dunn et al., 2015;
53 Hyland et al., 2013; Strömberg, 2004). Third, phytoliths are thought to commonly reflect relatively local
54 vegetation compared to palynomorphs, allowing for reconstruction of variation in vegetation structure
55 and composition across the landscape (e.g., Hyland et al., 2018; Strömberg et al., 2013). Finally, the
56 number of unique phytolith morphologies (morphotypes) in an assemblage has been used to estimate
57 past plant diversity (e.g., Hyland et al., 2015). Because of these advantages, phytolith analysis has shed
58 vital light on regional patterns of Cenozoic grass evolution and expansion in the Americas and eastern
59 Mediterranean (e.g., Dunn et al., 2015; Harris et al., 2017; Loughney et al., 2020; Strömberg, 2005;
60 Strömberg et al., 2007b).

61 Despite the recent surge in phytolith-based paleoecological studies, much work remains to
62 standardize the deep-time use of phytoliths. For example, modern analogue studies of soil phytoliths
63 suffer from inconsistent sampling schemes that cannot be equally applied in the fossil record (Crifò and
64 Strömberg, 2020). Moreover, too few modern analogue studies focus on how to reconstruct functional
65 characteristics of the vegetation (structure, composition, diversity, spatial heterogeneity) that are key
66 to understanding plant community responses to past environmental perturbations, impacts of
67 vegetation change on faunal communities, and the links between ancient and modern vegetation.
68 Substantial work has been done on calibrating phytolith proxies for tree cover in Africa (e.g., Aleman et
69 al., 2012; Barboni et al., 2007; Bremond et al., 2005a), but other aspects of vegetation type, and other
70 continents have received less attention (but see Fredlund and Tieszen, 1994; Hyland et al., 2013; Kerns
71 et al., 2001). For example, relative proportions of different phytolith morphotypes are rarely used
72 quantitatively in vegetation inference, even though their plant functional type (PFT) affinity is often
73 known based on modern reference collections (e.g., Piperno, 1985), and could enrich interpretations of
74 gross plant community composition. Furthermore, it remains largely untested whether phytolith
75 assemblages capture spatial heterogeneity in vegetation structure and composition, and in particular at
76 scales relevant to paleontology (with outcrops typically less than a few kilometers wide). The few
77 published modern analogue studies investigating the spatial resolution of the phytolith record suggest
78 that sampling along horizontal transects (the equivalent of sampling along stratigraphic levels) records
79 local-regional variation in canopy cover (e.g., Aleman et al., 2012; Fredlund and Tieszen, 1994;
80 Mercader et al., 2011; Piperno, 1988). However, the inconsistency in spatial scales examined in these
81 studies (10^1 – 10^5 meters), as well as in sampling and analytical approaches (see Crifò and Strömberg,
82 2020) make it difficult to generalize from these results.

83 A recent study of soil phytolith assemblages from a Costa Rican rainforest and dry forest
84 proposed guidelines for how to sample soil horizons within (10 x10 meter) quadrats (Crifò and

85 Strömberg, 2020), laying the groundwork for modern analog studies that are maximally comparable to
86 deep-time phytolith work. Here, we use this framework to investigate larger-scale spatial patterns of
87 soil phytoliths, with the goal of improving our ability to reconstruct ecologically meaningful parameters
88 of vegetation types in these two functionally different habitats. The tropical rainforest at La Selva
89 Biological Station (hereafter referred to as La Selva [LS]), and the tropical dry forest at Palo Verde
90 National Park (hereafter Palo Verde [PV]) experience different climates, in large part driving differences
91 in overall vegetation diversity, structure, and composition, and in spatial heterogeneity in these
92 characteristics. Thus, tropical, wet forests tend to be more species rich (2,077 species in LS vs. 619
93 species in PV; OTS, 2019), have taller, denser, more stratified, and homogeneous canopy cover, and
94 less ground cover compared to dry forests. They are dominated by woody evergreen taxa with
95 aseasonal phenology and adapted to mesic conditions; lianas, epiphytes (e.g., bromeliads, orchids),
96 palms, and gingers are abundant, and occupy a variety of forest strata. Dry forests tend to have smaller
97 and patchier canopy and denser ground cover (grasses and xerophytic herbaceous taxa, e.g.,
98 bromeliads) than rainforests. Deciduous, dry-adapted taxa with seasonal phenology dominate, with
99 evergreen taxa forming patches in moist areas (Murphy and Lugo, 1986; Quirico Jiménez et al., 2016).

100 We analyze the relationship between soil phytolith assemblages and vegetation structure,
101 composition, diversity, and spatial heterogeneity in structure and composition of these two forest
102 types. To do so, we first sampled along transects over an area comparable to a fossil locality (1-2 km)
103 using lower A-horizon soil samples, shown to be optimally comparable to paleosols (see Crifò and
104 Strömberg, 2020). Second, we quantified functional aspects of vegetation, namely degree of tree
105 cover, composition, and species richness. For our phytolith assemblage analysis, we chose the so-called
106 “general approach”, which was designed with the deep-time record in mind (Strömberg, 2003, 2004;
107 see Strömberg et al., 2018), as opposed to for a specific geographic region. Third, to investigate
108 potential drivers of differences within and between sites, we measured environmental variables such as

109 altitude and proximity to water of the quadrats. We also considered parameters that can bias phytolith
110 assemblages, thus distorting the vegetation signal, by estimating canopy cover and phytolith
111 productivity of dominant species. Given such production biases and known phytolith preservation
112 biases (reviewed in Crifò and Strömberg, 2020; Strömberg et al., 2018), we then tested for covariation
113 between observed and inferred patterns of vegetation structure (canopy openness), species
114 composition, and diversity. We address two central questions: (1) do phytolith assemblages reflect the
115 overall differences in structure, composition, and diversity between rainforest and dry forest
116 vegetation? and (2) do phytolith assemblages from lateral transects capture variation in structure and
117 composition typical of these vegetation types?

118 Although climatic and functional differences between the two vegetation types might
119 themselves influence soil phytolith assemblage formation and preservation in several ways, including
120 their spatiotemporal resolution (Crifò and Strömberg, 2020), we hypothesize that (a) phytolith
121 assemblages will reflect structural and compositional differences between the rainforest and dry forest,
122 allowing for them to be distinguished in the fossil record; (b) because of multiplicity and redundancy in
123 phytolith production, phytolith assemblages do not record differences in species richness; and (c)
124 phytolith analysis of samples collected along transects within each vegetation type can reconstruct
125 landscape-scale (over 10^2 – 10^3 meters) variation in forest structure and composition.

126

127 **2. Material and methods**

128 2.1 *Study sites*

129 Soil and plant materials were collected in Costa Rica at the Organization for Tropical Studies (OTS) La
130 Selva ($10^{\circ} 25' 46.3116''$ N, $84^{\circ} 0' 24.8328''$ W) and Palo Verde ($10^{\circ} 20' 44.682''$ N, $85^{\circ} 20' 29.5656''$ W)
131 biological stations (Fig. 1), which are protected areas with known disturbance history, and therefore
132 ideal for our study. LS includes a large area with primary neotropical rainforest. PV includes areas with

133 remnants of neotropical seasonal and transitional dry forests (OTS, 2019). At LS, mean annual
134 precipitation (MAP) is 4,000 mm with no pronounced seasonality, and mean annual temperature (MAT)
135 is 25.8 °C (OTS, 2019). At Palo Verde MAP is 1,500 mm (OTS, 2019), and MAT is 25 °C (Gillespie et al.,
136 2000) with 95% of precipitations falling during the rainy season (May to December), (Quirico Jiménez et
137 al., 2016). The most abundant families at LS are Arecaceae (e.g., *Welfia regia*, *Socratea exorrhiza*),
138 Bromeliaceae, and Orchidaceae. Among woody plants, LS flora is dominated by the families
139 Annonaceae, Burseraceae (e.g., *Protium pittieri*), Fabaceae (e.g., *Inga pezizifera*, *Pentaclethra*
140 *macroloba*), as well as Euphorbiaceae, Lauraceae (e.g., *Ocotea* sp.), Malvaceae, Melastomataceae (e.g.,
141 *Clidemia* sp., *Miconia* sp.), Meliaceae, Sapotaceae, Moraceae (e.g. *Ficus* sp.), Piperaceae (e.g. *Piper* sp.),
142 and Rubiaceae (e.g., *Warszewiczia coccinea*, *Psycothria* sp.) (Gentry, 1993; OTS 2020). The herbaceous
143 vegetation is dominated by species in the family Bromeliaceae, Cyperaceae Orchideaceae, Poaceae,
144 and in families of the order Zingiberales (e.g., Heliconiaceae, Marantaceae, Zingiberaceae). Within
145 Poaceae, most genera belong to the subfamilies Bambusoideae (e.g., *Bambusa*, *Chusquea*),
146 Chloridoideae (e.g., *Chloris*, *Eragrostis*), and Panicoideae (e.g., *Panicum*, *Paspalum*, *Hymenachne*). PV
147 vegetation is dominated mainly by deciduous species with some semi-deciduous patches. The PV dry
148 forest is dominated by deciduous species belonging to the families Acanthaceae (e.g., *Aphelandra*
149 *scabra*), Anacardiaceae (e.g., *Astronium graveolens*, *Spondias mombin*), Bignoniaceae (e.g., *Tabebuia*
150 *ochracea*), Burseraceae (e.g., *Boursera simarouba*), Erythroxylaceae (e.g., *Erythroxylum havanense*),
151 Fabaceae (e.g., *Acacia collinsii*, *Dalbergia retusa*, *Lonchocarpus minimiflorus*, *Lysiloma divaricatum*),
152 Malvaceae (e.g., *Apeiba tribourbou*, *Pachira quinata*), Primulaceae (e.g., *Bonellia nervosa*), Rubiaceae
153 (e.g., *Calycophyllum candidissimum*), Verbenaceae (e.g., *Rehdera trinervis*), and Zygophyllaceae (e.g.,
154 *Guaiacum sanctum*). Evergreen species occupy small vegetation patches and belong predominantly to
155 the families Anacardiaceae (e.g., *Anacardium excelsum*), Chrysobalanaceae (e.g., *Licania arborea*),
156 Elaeocarpaceae (e.g., *Sloanea terniflora*), Fabaceae (e.g., *Andira inermis*, *Zygia longifolia*), Moraceae

157 (e.g., *Brosimum alicastrum*, *Trophis racemosa*), Polygonaceae (e.g., *Coccoloba caracasana*), Sapindaceae
158 (e.g., *Thouinidium decandrum*), and Sapotaceae (e.g., *Manilkara chicle*, *Sideroxylon capiri*). The
159 herbaceous vegetation is dominated by species in the family Poaceae (e.g., *Paspalum*) and Cyperaceae
160 (e.g., *Cyperus*). Within Poaceae, most genera belong to the subfamilies Panicoideae (e.g., *Panicum*,
161 *Paspalum*, *Urochloa*) and Chloridoideae (e.g., *Bouteloua*, *Eragrostis*) (OTS, 2020; Quirico Jiménez et al.,
162 2016).

163 2.2 Sample and data collection

164 2.2.1 LAI and floristic data

165 At LS and PV, respectively, we established, as randomly as possible, 14 and eleven 10x10-meter
166 quadrats along vegetation transects (see Crifò and Strömberg, 2020), but avoiding disturbed areas. In
167 LS we traced a single, ~1.5 km long transect; in PV we traced two transects of ~1 km and 0.5 km length.
168 We chose two lowland sites, and relatively short transects because these settings are analogous to a
169 typical fossil locality (i.e., usually representing a lowland landscape, and with continuous transects
170 typically spanning from a few meters to several hundreds of meters).

171 To estimate (observed) canopy openness in each quadrat, we measured LAI from hemispherical
172 photographs, averaging LAI values from 10 photos/quadrat (during the dry season, i.e., February 2016)
173 following Dunn et al. (2015) (see Table S1, in Supplementary Information SI 1 for LAI raw data). To
174 document within-site and within-quadrat species composition, species richness, and tree
175 biomass, we identified to species all trees > 1 cm dbh (diameter at breast height, i.e., at 1.4
176 meters), recorded the number of individuals, and measured the dbh, with the exception of
177 leafless, deciduous species (≤ 1 /quadrat). The contribution of each species to the quadrat tree
178 biomass (hence to the phytolith pool) was estimated based on their relative basal area (rBA_{sp})

179 (raw data in Table S2 in Supplementary Information SI 2). A positive relationship between tree basal
180 area and LAI was previously documented in rainforests (Clark et al., 2008) and dry forests (e.g.,
181 Kalacska et al., 2005).

182 Using dbh measurements, the basal area of each species (BA_{sp}) was calculated as:

183
$$BA_{sp} = \sum_{i=1}^n 2\pi \left(\frac{dbh_i}{2}\right)^2$$

184 Then, in each quadrat, species rBA_{sp} was calculated as:

185
$$rBA_{sp} = \frac{BA_{sp}}{A} \times 100$$

186 where A is the sum of the basal areas of all trees in a quadrat. We consider all individuals with $dbh \geq 1$
187 cm, because even small plants might contribute considerably to the soil phytolith pool (e.g.,
188 multiple individuals per species and/or high phytolith producers). rBA_{sp} was not estimated for
189 herbs.

190 Plant community composition for each quadrat was assessed by compiling a matrix of tree
191 species presence/absence data from species counts in the field (palm species were tallied, herbs were
192 not). We used this matrix to calculate α diversity as the total number of tree and palm species identified
193 in a quadrat, and γ diversity as the sum of all species at each site. We used species count data to
194 generate species relative abundance matrices for further statistical tests (see Table S3, Supplementary
195 Information SI 3, and Table S4, Supplementary Information SI 4, for raw species count data).

196 2.2.2 *Environmental gradients*

197 Although the choice of the study settings did not allow us to trace transects that capture large
198 environmental gradients, likely to result in major vegetation turnover, we expect soil phytolith
199 assemblages to reflect at least some gradients. At both La Selva and Palo Verde, the quadrats located
200 closer to the river tend to be at lower elevation and, thus, are seasonally inundated, whereas quadrants

201 located at higher elevation are not. Therefore, to account for the potential influence of elevation and
202 moisture on the vegetation and the phytolith record, we assessed elevation and moisture gradients by
203 measuring the elevation and distance from the closest water basin (Río Puerto Viejo at LS, and Río
204 Tempisque at PV) of each quadrat, using Google Earth Pro 7.3.4.5776 (see Table S5, in Supplementary
205 Information SI 5).

206 2.2.3 *Herbarium and soil samples*

207 At LS and PV, we collected one soil sample per quadrat from the lower portion (5-10 cm depth)
208 of the soil A-horizon after removing the leaf litter and organic material (O horizon). Soil samples from
209 this portion of the soil are a better analogue to paleosol samples (in which the upper portion is rarely
210 preserved due to erosion) than surface samples (see Crifò and Strömberg, 2020). This resulted in 14
211 (Q1-14) and 11 (Q15-25) samples from, respectively, LS and PV. Additionally, we collected herbarium
212 specimens (i.e., leaves, and small branches when compound leaves were present, as well as thorns
213 when present) for each newly observed plant species (unless leaves were absent), in total 115 and 60
214 plant samples (including trees, shrubs, palms, lianas, and herbs) from, respectively, LS and PV (see
215 Table S6 in Supplementary Information SI 6 for a list of herbarium specimens).

216 We repositied soil samples in the Burke Museum of Natural History and Culture (UWBM), and
217 two sets of herbarium vouchers at the UWBM Botany Collections (WTU herbarium), University of
218 Washington, Seattle, and at the Herbario Nacional de Costa Rica, San José.

219 2.3 *Sample processing and analysis*

220 2.3.1 *Phytolith extraction and assemblage analysis*

221 To study soil phytolith assemblage composition and estimate species-specific phytolith yield,
222 we extracted phytoliths from modern soils (1 g/sample) and dry plant material (0.025 g/species),
223 following standard extraction methods (Piperno, 1988; Zhao and Pearsall, 1998). Note that
224 fractionation of the soil phytolith samples was not conducted, because the analysis of small and large

225 fractions separately (e.g., Piperno, 2006) prevents the calculation of morphotype relative abundance
226 within whole phytolith assemblages, and is therefore not suitable for the purpose of analyzing whole
227 assemblages quantitatively. Extracted biogenic (soil/plant) silica was mounted on slides in a solid
228 medium (Meltmount®) and viewed at 1,000x magnification under a compound microscope for
229 phytolith identification and counting. A liquid mounting medium (immersion oil) allowing phytolith
230 rotation and observation of three-dimensional morphology was used for identification of grass silica
231 short cell phytoliths (hereafter GSSCP). We followed the “general approach” to phytolith assemblage
232 analysis developed by Strömberg (2003, 2004) and Strömberg et al. (2013). This approach is tailored to
233 the deep-time fossil record in that it is designed to provide taxonomic or ecological affinities for
234 phytoliths that are broadly applicable (phylogenetically, ecologically) and that could be used with fossil
235 phytolith assemblages for which the context (e.g., the biogeographic region) is incompletely known
236 (see discussion in Strömberg, 2003, 2004; Strömberg et al., 2018). To accomplish this goal, it relies on a
237 classification system that uses the relative abundance distribution of morphotypes (see list of
238 morphotypes in Table S7 in Supplementary Information SI 7) among species in a phytolith reference
239 collection combined with a comprehensive survey of the published literature to assign morphotypes to
240 classes representing broad plant functional (and sometimes taxonomic) groups (PFT) (see Table 2). By
241 analyzing our modern soil phytolith assemblages using these conservative PFT assignments, we can
242 assess to what extent the “general approach” is able to distinguish vegetation types in deep-time.
243 However, it is important to note that, because of multiplicity and redundancy in phytolith production,
244 morphotypes are seldom exclusive to a certain PFT or taxonomic group, even though they are typically,
245 or most abundantly deposited in members of those groups. For example, even though particular
246 morphotypes are by far more common in cool-temperate Pooideae grasses, they are also produced in
247 low abundances by certain PACMAD taxa (e.g., Mulholland, 1989). Therefore, certain phytoliths in our
248 Neotropical assemblages may be assigned to the Pooideae PFT (“POOID-D”) despite the low likelihood

249 of pooid grasses being present. In fact, it is an important task of our study to assess whether such
250 method-related “errors” will throw off the interpretation of our modern assemblages—and, by
251 extension, will compromise our ability to correctly interpret fossil phytolith assemblages.

252 Vegetation structure and composition are interpreted primarily from the proportions of
253 diagnostic PFT in a phytolith assemblage (see further below). Diagnostic PFT are those indicative of
254 different types of “forest indicator” (FI) taxa (i.e., palms, Zingiberales, woody dicots, and other forest
255 indicators represented by PFT#1-6; Table 2), which are often, but not always linked to forest habitats,
256 and grass clades or groupings of grasses with similar ecology (including both closed- and open habitat
257 clades corresponding to PFT# 8-14; Table 2) (Strömberg, 2003; Strömberg et al., 2018). On
258 Meltmount® slides, we counted ≥ 200 diagnostic phytoliths (PFT# 1-6 + 8-14) per slide for statistical
259 significance (Pearsall, 2000; Strömberg, 2009) and, separately, on oil slides, ≥ 200 diagnostic GSSCP
260 (PFT# 8-13). In samples where GSSCP represented $< 5\%$ of the total assemblage (primarily for LS), we
261 counted all GSSCP on one oil slide, often resulting in < 200 diagnostic GSSCP (for raw phytolith
262 morphotype data see Table S7 in Supplementary Information S 17). In addition, we carefully scanned
263 each slide to account for the presence of rare morphotypes which otherwise would go undetected.

264 Vegetation structure/tree cover was assessed using the so-called *FI-t* ratio, which estimates the
265 relative abundance of forest indicator PFT (FI) within the total count of diagnostic morphotypes
266 (Strömberg, 2003; Strömberg et al., 2007a) (see Supplementary Information SI 8). In doing so, the *FI-t*
267 ratio implicitly assumes that all FI taxa are associated with forests, making it a coarse approximation of
268 tree cover, which should not be used on its own without simultaneously studying the morphotype and
269 PFT composition of FI and GSSCP assemblages (see discussion in Strömberg, 2004; Strömberg et al.,
270 2018). In LS and PV, both forest habitats, it is reasonably safe to assume that FI derive from forest taxa,
271 but for deep-time inferences, the *FI-t* ratio must be evaluated together with independent evidence for
272 habitat openness (Dunn et al., 2015). Another widely applied metric to assess tree cover, the *D/P* index,

273 was designed with African vegetation types in mind (Alexandre et al., 1997), but has been applied to the
274 Neotropics (Dickau et al., 2013). In contrast to *FI-t*, the *D/P* index typically uses the sum of “globular
275 granulate” phytoliths (i.e., SPHEROID ORNATE sensu Neumann et al., 2019) produced by many tropical
276 ligneous dicotyledonous, over the sum of Poaceae phytoliths, that is, GSSCP plus, in some publications,
277 morphotypes such as bulliform cells (BLOCKY + BULLIFORM FLABELLATE) and trichomes (ACUTE BULBOSUS)
278 (compare Alexandre et al., 1997; Bremond et al., 2008). To compare the usefulness of the two metrics in
279 our two Neotropical vegetation types, we also calculated the *D/P* index for all our forest assemblages.
280 Specifically, we calculated two different *D/P* indices: a) D/P_1 , according to Bremond et al. (2008), and b)
281 D/P_2 , a modified version of D/P_1 which, in addition to SPHEROID ORNATE morphotypes, also takes into
282 account phytolith morphotypes produced by palms (SPHEROID ECHINATE) (see Figure S1 in
283 Supplementary Information SI 8, and Table S5).

284 To study relative differences in estimated plant community composition among quadrats and
285 sites, we used relative abundance of diagnostic phytolith (1) PFT, and (2) morphotypes (see Tables S5
286 and S7). To estimate diversity from phytolith count data, we calculated α_p diversity as the number of
287 phytolith morphotypes identified per quadrat, and γ_p diversity as the sum of all morphotypes present at
288 each of LS and PV.

289 2.3.2 *Phytolith production bias*

290 Differential phytolith yield in plants may skew soil phytolith assemblages, especially in dry
291 habitats (where low producer families, such as Fabaceae, are abundant; Hodson et al., 2005). To better
292 understand this potential bias, we assessed phytolith production semi-quantitatively for all collected PV
293 tree species. In addition, to account for the potential influence of leaf phenology on phytolith yield, we
294 classified species as evergreen, deciduous, and semi-deciduous and estimated the average phytolith
295 yield of these three groups. A negative correlation has been found between leaf longevity and silica
296 accumulation (Cooke and Leishman, 2011) suggesting that deciduous species deposit more phytoliths

297 in their leaves compared to evergreen and semi-deciduous species. In addition, deciduous trees likely
298 contribute more leaf litter, hence more silica, to the soil O horizon. First, we obtained semi-quantitative
299 measurements of diagnostic-phytolith productivity (Y_{phyto}) for each species by combining silica yield
300 estimates with counts of diagnostic phytoliths. Using this information, we ranked species from low (0)
301 to high (4) diagnostic phytolith productivity (details in SI 8). We then used the total basal area of each
302 species to evaluate its relative contribution to total quadrat biomass. By comparing the relative species
303 contribution to the quadrat total tree basal area to their phytolith yield rank, and to their leaf
304 phenological type, we flagged sites where producer bias is likely to influence phytolith assemblage
305 composition. Accordingly, quadrats with over half of the total basal area represented by non-producer,
306 very low-, or low producer species (ranks 0-2), or by evergreen species may yield phytolith assemblages
307 with lower-than-expected $FI-t$ ratios. Conversely, quadrats with > 50% of moderate or high producers
308 (ranks 3-4) or deciduous species, could yield phytolith assemblages with higher-than-expected $FI-t$
309 ratios.

310 2.4 Statistical comparisons

311 All statistical analyses were performed using R (R Development Core Team, 2019) (see SI 8).

312 2.4.1 Differences between vegetation types

313 Structure — We determined the average $FI-t$ ratio per site ($FI-t_{site}$) as the mean $FI-t$ of all
314 quadrats ($FI-t_{quadrat}$); and the average LAI per site as the mean LAI of all quadrats (see raw data in Table
315 S5). Published LAI values were used to determine and contrast the functional meaning of $FI-t$ and LAI
316 values (i.e., the vegetation type they indicate) in the two sites.

317 Composition — To determine potential floristic similarities between LS and PV floras, we
318 compiled species lists from our field observations (in Table S3 and S4). Differences in measured floristic
319 composition between LS and PV were evaluated through NMDS and cluster analysis of phytolith
320 relative abundance matrices (diagnostic phytolith PFT, “raw” morphotypes). We chose NMDS, which is

321 a preferred ordination technique in community ecology because it (1) does not assume a linear
322 relationship between variables; and (2) emphasizes relative, rather than absolute differences between
323 samples (Kruskal, 1964). Using plots of *stress* (a normalized measure of goodness of fit) versus number
324 of dimensions, and tests of stress value randomization allows determination of the minimum number of
325 dimensions required to capture the variation in the data. We also performed cluster analysis using a
326 Bray-Curtis dissimilarity matrix based on phytolith PFT and Ward's minimum variance method.

327 Diversity—To test whether phytolith assemblages reflect differences in plant species richness
328 between vegetation types, we performed linear regression analysis between floristic α and phytolith α_p
329 diversity. Note, that grasses were not included in our assessments of floristic diversity. Therefore, to
330 make our phytolith diversity estimates comparable to these floristic diversity measurements, we only
331 included non-GSSCP morphotypes.

332 2.4.2 *Variation within vegetation types*

333 For both vegetation structure and composition, we evaluated whether (1) observed and
334 inferred structure and composition display spatial patterns that are correlated with environmental
335 gradients (elevation and distance from water; for composition also canopy openness); and (2) observed
336 and inferred patterns of structure and composition coincide.

337 Structure—Within-quadrat tree cover was calculated from single assemblages as $FI-t_{quadrat_i}$
338 95% confidence intervals (CI) were constructed using bootstrap analysis (10,000 replicates) (Strömberg,
339 2009) (see SI 8). To evaluate the similarity between each $FI-t_{quadrat}$ and the $FI-t_{site_i}$ and between all $FI-$
340 $t_{quadrat}$, we conducted pairwise comparisons between $FI-t_{quadrat}$ values and the $FI-t_{site}$ CI, and between all
341 $FI-t_{quadrat}$ values and CI. We considered $FI-t_{quadrat}$ being significantly different from $FI-t_{site}$ and from other
342 $FI-t_{quadrat}$ values, if it did not overlap with their CI (see Strömberg, 2009). To test for significant
343 relationships between environmental gradients (elevation, distance from water) and vegetation
344 structure (observed [LAI], inferred [$FI-t$ ratios]), we performed multiple Spearman rank-order tests of

345 correlation within each site (see SI 8). Linear regression analysis was used to test for a significant
346 relationship between LAI and *FI-t* (as well as between LAI and the *D/P* index; see Figure S1 in SI 8).

347 Composition —To evaluate the relationship between environmental gradients and plant
348 community composition (observed, inferred) along transects at each site, we first performed NMDS on,
349 respectively, the species relative abundance data and the phytolith relative abundance data (PFT, “raw”
350 morphotypes). To study within-site differences in inferred composition we also performed cluster
351 analysis using the Bray-Curtis dissimilarity matrix with phytolith PFT and Ward’s minimum variance
352 method. We investigated spatial autocorrelation in observed and inferred plant community
353 composition along the transects by performing Mantel tests (see SI 8). To compare plant species
354 composition and phytolith PFT composition with environmental gradients in elevation and moisture,
355 we used linear regression analysis of the quadrat scores of the first two NMDS axes against these
356 environmental variables.

357

358 3. Results

359 3.1 *Soil and plant phytolith yield*

360 LS soils samples yielded well preserved phytoliths; PV soil samples yielded variably preserved
361 phytoliths, including one unproductive sample (in quadrat 18, excluded from this study) and especially
362 poor preservation in quadrats 19, 22, and 23; (phytoliths images in Figures S2 and S3, in Supplementary
363 Information SI 9). At PV, plant phytolith yield varied widely among species (see Table S8 in
364 Supplementary Information SI 10), with some quadrats dominated by non- or low-producing woody
365 species (Q17, Q21, Q22,23), some by moderate- to high producers (Q19, Q20, Q24, and Q25), others
366 with an intermediate composition (Q15, Q16); as well as some quadrats dominated by evergreen
367 species (Q17, Q20, Q25), and others by deciduous species (Q15, Q16, Q19, Q21-25), (Table 3).

368 Overall, silica accumulation is significantly higher in woody deciduous species (average $Y_{phyto} =$
369 5.48 ± 5.1) than in woody evergreen species (average $Y_{phyto} = 1.32 \pm 1.57$), (t-test, $p < 0.001$). There is no
370 statistically significant difference in silica accumulation between woody semi-deciduous (average Y_{phyto}
371 $= 2.14 \pm 3.34$) and woody deciduous species (t-test, $p > 0.05$), and between woody semi-deciduous and
372 woody evergreen species (t-test, $p > 0.05$) (Fig. 2).

373 3.2 Differences between vegetation types

374 The two sites differ in average $FI-t$ ratio, D/P_1 , and D/P_2 indices. At LS, $FI-t_{site}$ (average $FI-t$ ratio)
375 is $98.5 \pm 1.4\%$. Confidence intervals are narrow (Fig. 3a), and all $FI-t_{quadrat}$ fall within the CI of $FI-t_{site}$,
376 except for Q3. Average D/P_1 , and D/P_2 are 6.76 ± 6.15 , and 46.63 ± 52 respectively. At PV, $FI-t_{site}$ is
377 $58.8 \pm 21.5\%$; the minimum $FI-t_{quadrat}$ is 19.2% (Q15), and the maximum $FI-t_{quadrat}$ is 84.6% (Q23). $FI-t$
378 ratios are low in quadrats Q16 and Q17 ($37.8-41.5\%$); intermediate in Q20–22 and Q25 ($55.5-62.75\%$);
379 and high in Q19, 23, and Q24 ($81.1-84.6\%$) (Fig. 3b). $FI-t_{quadrat}$ falls within the CI of $FI-t_{site}$ in only four
380 quadrats (Q20–22 and Q25) (Table S5). Average D/P_1 and D/P_2 are 0.76 ± 0.81 and 0.99 ± 0.95 respectively.

381 According to our vegetation census, LS and PV do not share any plant species. NMDS analyses
382 performed on phytolith PFT and morphotype data produce similar results (Figure S4, in Supplementary
383 Information SI 11); here, we present NMDS ordination plots using phytolith PFT only (Fig. 4). For all
384 NMDS ordinations, two dimensions appear adequate to represent the data. Cluster analysis and NMDS
385 (Fig. 4) separate LS and PV phytolith assemblages well, based on the relative abundances of phytolith
386 PFT and morphotypes (not shown). All PFT have significant loadings on NMDS axes 1 and 2 ($p < 0.01$),
387 except for Zingiberales ($p > 0.05$) (Fig. 4a). LS phytolith assemblages cluster together (toward the
388 negative end of axis 1) and are separated from PV assemblages (spread out along axis 2) mainly along
389 axis 1 (Fig. 4a). Along axis 1, the highest loadings correspond to the PFT Palm ($p < 0.01$, negative
390 loading), Chloridoid, and closed habitat (BOE) grasses ($p < 0.01$, positive loadings). On axis 2 the
391 highest loadings correspond to Other FI ($p < 0.01$, negative loading), and various grass PFT (PACMAD,

392 PANI, POOID-ND, and "POOID-D", $p < 0.01$, positive loadings); these latter PFT seem to differentiate
393 phytolith assemblages within PV quadrats (i.e., between the groups highlighted by cluster analysis; Fig.
394 4b) rather than between sites.

395 Our vegetation census indicates that species diversity (of woody angiosperm and palm species)
396 is higher across LS (average $\alpha = 17.57$, $\gamma = 86$) than across PV (average $\alpha = 11.4$, $\gamma = 46$) quadrats.
397 Although the herbaceous vegetation component is not considered in our α and γ calculations, this
398 result agrees with the general trend highlighted by LS and PV global censuses (OTS, 2020). In contrast,
399 non-GSSCP phytolith diversity is lower at LS (average $\alpha_p = 29.8$, $\gamma_p = 62$) than at PV (average $\alpha_p = 37.6$,
400 $\gamma_p = 72$). Within and across the two sites, α and α_p are not related ($p > 0.05$, linear regression).

401 3.3 Variation within vegetation type

402 3.3.1 Vegetation structure

403 At LS, neither LAI nor $FI-t$ ratio, or D/P_1 and D/P_2 follow gradients in elevation or moisture ($p >$
404 0.05 , Spearman rank-order correlation test) (Figs. 5a, b; see also Figure S5; figure not shown for the two
405 D/P indices). At PV, LAI (average = 2.9 ± 0.8) is lower than at LS (average = 5.1 ± 0.4). Along the first
406 transect (T1), LAI is negatively correlated with distance from the water ($p < 0.01$, Spearman rank-order
407 correlation test), but not with elevation ($p > 0.05$, Spearman rank-order correlation test). There are no
408 relationships between LAI and environmental gradients along T2 ($p > 0.05$, Spearman test) (Figs. 5a, b;
409 see also Figure S6; figure not shown for the two D/P indices). The $FI-t$ ratio and the D/P indices show a
410 similar pattern to LAI along T1 although the $FI-t$ ratio shows stronger relationships with environmental
411 variables than the D/P indices. In fact, $FI-t$ decreases with moisture and with elevation ($p < 0.05$,
412 Spearman test), D/P_1 decreases with elevation ($p < 0.05$, Spearman test), and D/P_2 does not follow any
413 moisture or elevation gradient ($p > 0.05$, Spearman test). The $FI-t$ ratio and the D/P indices do not follow
414 any pattern along T2 ($p > 0.05$, Spearman test) (Figs. 5a, b).

415 At LS, the *FI-t* ratio and LAI are negatively correlated ($p < 0.01$, Spearman test) (Fig. 5a, b),
416 while LAI show no significant correlation with either D/P_1 and D/P_2 ($p > 0.05$, Spearman test). At Palo
417 Verde, the *FI-t* ratio, D/P indices, and LAI all decrease along T1 (Fig. 6a; D/P indices not shown), but
418 show no statistically significant correlation along either T1 and T2 ($p > 0.05$, Spearman test). When the
419 data from LS and PV are combined, LAI show a positive, statistically significant relationship with *FI-t* (p
420 < 0.01 , $R^2 = 0.32$, t-test), D/P_1 ($p < 0.01$, $R^2 = 0.29$, t-test), and D/P_2 ($p < 0.001$, $R^2 = 0.55$, t-test) (SI 8, Fig.
421 S1c–e). Finally, *FI-t* show a significant correlation with both D/P_1 ($p < 0.01$, $R^2 = 0.38$, t-test) and D/P_2 (p
422 < 0.001 , $R^2 = 0.55$) (SI 8, Fig. S1a, b).

423 3.3.2 Vegetation composition

424 NMDS ordinations, linear regressions, and Mantel tests ($p > 0.05$) show no evident relationships
425 between environmental gradients, on the one hand, and patterns of either woody plant species
426 compositions or phytolith PFT/morphotype compositions of quadrats along transects on the other.
427 Indeed, in NMDS ordinations of LS quadrats using PFT compositions, quadrats are randomly
428 distributed in ordination space (Fig. 7a). In contrast, at PV, NMDS ordination and cluster analysis (using
429 PFT, morphotypes) resulted in two distinct quadrat groupings on opposite ends of NMDS axis 1 (Fig.
430 7b). The first group includes quadrats along T1 (15-17); the second quadrats from T1 (19-20) and T2 (21-
431 25). Differences along NMDS axis 1 are driven primarily by forest indicator phytoliths (Other FI, positive
432 loading, $p < 0.01$) and GSSCP attributed to the subfamilies Pooideae ("POOID-D"), Panicoideae (PANI),
433 and the PACMAD clade (all negative loadings, $p < 0.01$), and to a lesser extent the subfamily
434 Chloridoideae (CHLOR; negative loading, $p < 0.01$). Linear regression shows a negative relationship
435 between NMDS axis 1 and elevation ($p < 0.05$). NMDS ordination of PV morphotype abundance data
436 gave similar results. In addition, the Mantel test indicates that plant species composition of quadrats is
437 related to their geographic position ($p < 0.01$). At LS, the pattern of spatial change in observed floristic
438 composition and phytolith PFT/raw morphotype composition of quadrats (as estimated by values on

439 NMDS axis 1) seem to partially match, although they show no clear trends along moisture and elevation
440 gradients (Fig. 4c, d, see also Figure S5). Observed and inferred compositional changes at PV are also
441 partially similar, at least along T₁ (Fig. 6c, d; see also Figure S6).

442 4. Discussion

443 4.1 *Preservation biases*

444 Overall, phytolith yields are higher and preservational states are better at La Selva than at Palo
445 Verde. At PV, weakly developed Inceptisols (Powers et al., 2009), and alkaline conditions (which favor
446 silica dissolution) might explain poor phytolith preservation in general, preferential preservation of
447 GSSCP (which are less prone to chemical dissolution; Bartoli and Wilding, 1980), or both. In contrast,
448 the pristine phytolith preservation in the rainforest is consistent with the well-developed Oxisols
449 present at LS (Alvarado Induni, 1990), which typically have a pH at or below neutral (Weil and Brady,
450 2016).

451 4.2 *Differences between vegetation types*

452 Our results suggest that analysis of multiple soil phytolith samples can adequately reconstruct
453 important functional aspects of vegetation types (structure, composition), which allow distinction
454 between LS rainforest and PV dry forest. Thus, observed structural differences between LS and PV
455 vegetation are broadly reflected by higher and more homogeneous *FI-t* ratios (as well as *D/P* indices;
456 see below for a comparison of the *FI-t* ratio and *D/P* indices) in LS phytolith assemblages. However, the
457 lower and highly variable *FI-t* values across the PV transects exaggerate actual structural heterogeneity,
458 with grasses being overrepresented in some quadrats (e.g., 15, 16, and 17) and woody dicots in others
459 (e.g., 23). These discrepancies likely stem from preservation biases (i.e., soil conditions) or biases in
460 biomass- or phytolith production among taxa, which both vary spatially across PV (see further below).
461 Overall floristic differences between LS and PV are also evident from phytolith assemblage
462 composition, with higher proportions of palm phytoliths and lower proportions of grass phytoliths at

463 LS, and phytolith morphotypes (assigned to PFT Palm/Zingi and Palm/Zingi/Brom) that only occur at
464 PV. This result agrees with previous studies of the composition of surface soil phytolith assemblages
465 showing significant differences between distinct forest types in Amazonia (Dickau et al., 2013; Watling
466 et al., 2016). For instance, although the authors of these studies used different methodological
467 approaches (e.g. surface- instead of subsurface samples and size-fractionation during processing), such
468 that their data are not directly comparable to ours, their phytolith assemblages from dense evergreen
469 humid forest are similar to those from LS, showing overall a dominance of palm and arboreal phytolith
470 types and low frequencies of phytoliths from herbaceous plants. Our study also shows that the more
471 spatially variable floral composition at PV is reflected in compositionally more heterogeneous phytolith
472 assemblages. Similarly, Dickau et al. (2013) found noticeable differences between six surface soil
473 phytolith samples collected in two plots from *terra firme* semi-deciduous forests in Amazonia. Future
474 work to better understand phytolith taxonomic affinity (e.g., for Palm/Zingi, Palm/Zingi/Brom) will
475 further enhance vegetation differentiation using phytoliths.

476 In contrast to vegetation structure and composition, phytolith analysis of non-GSSCP
477 morphotypes cannot capture the substantially higher plant species richness observed herein and
478 previously recorded in LS compared to PV (OTS, 2020). This result is not likely to be a consequence of
479 our methodological approach, which excludes GGSCP morphotypes in order to make phytolith diversity
480 estimates maximally comparable to measurements of floristic diversity (which only included trees). In
481 fact, when GSSCP morphotypes are added to our analysis, it results in even greater phytolith diversity
482 at PV (where grasses are more abundant), thus exacerbating the negative relationship between plant
483 species richness and the number of phytolith morphotypes at the two sites. The lack of correlation
484 between floristic diversity and phytolith richness supports our hypothesis that the two are not
485 meaningfully related. This is not surprising, as multiplicity and redundancy in phytolith production

486 strongly limit phytolith taxonomic resolution. We therefore question inferences of taxonomic diversity
487 based solely on the number of phytolith morphotypes in fossil assemblages (e.g., Hyland et al., 2015).

488 4.3 *Variation within vegetation types*

489 To what extent does spatial heterogeneity in vegetation structure (openness) and floristic
490 composition characterize rainforest vs. dry forest, and how is it manifested in soil phytolith
491 assemblages? The answers to these questions vary between sites, depending both on vegetation type
492 and on differences in environmental factors that influence the representativeness of the soil phytolith
493 record.

494 4.3.1 *Vegetation structure (FI-t ratio/D/P indices)*

495 Along the LS transect, canopy openness (LAI) values vary in a nonlinear way that cannot be
496 directly linked to altitudinal or moisture gradients. Nevertheless, the values fall within the range for late
497 successional stage rainforests (e.g., Murphy and Lugo, 1986), likely reflecting the absence of major
498 structural (and functional) variation within the rainforest. Similarly, estimated canopy openness (*FI-t*)
499 values all indicate a dense, closed-canopy forest and are invariant along environmental gradients. The
500 *D/P* indices follow a substantively similar pattern but with a wider range of variation compared to the *FI-*
501 *t* ratio (see further below for a discussion of the *FI-t* - *D/P* index comparison). However, patterns of LAI
502 and *FI-t* change along the LS transect do not match, indicating that the *FI-t* ratio is insensitive to LAI
503 variation at the scale of this study. An even greater mismatch is observed in patterns of LAI and *D/P*₁
504 and *D/P*₂. The disparities between observed and estimated vegetation structure at LS likely result
505 primarily from temporal averaging of phytolith assemblages by vertical translocation (mixing of older
506 and younger soil, e.g., through bioturbation). Hence, a single *FI-t* (or *D/P*) value represents an average
507 over several tens to hundreds of years, during which canopy structure might have changed slightly
508 compared to today, thus explaining the mismatch between individual *FI-t* (or *D/P*) and LAI values. In

509 contrast, spatial mixing due to lateral transport should matter less in this closed forest (Piperno, 1988;
510 Crifò and Strömberg, 2020).

511 At PV, the highly variable LAI values (range = 1.8–3.9) along the transects are comparable to
512 published values for the site (PAI = 1.7–3.4 in February), (Kalacska et al., 2005), and consistent with
513 relatively heterogeneous canopy cover in the dry forest. The *FI-t* ratio also varies across PV transects,
514 but with a wider range of values ($FI-t = 19.16\text{--}84.45\%$) that encompasses values characteristic of very
515 different vegetation structure: from savannas to closed forests (Strömberg, 2009; Strömberg et al.,
516 2013). *D/P* indices similarly show a wide range of variation ($D/P_1 = 0.06\text{--}1.8$; $D/P_2 = 0.06\text{--}2.2$), but overall
517 markedly underestimate tree cover, producing values that would typically be interpreted as
518 steppe/short grassland to dense savanna (see further below).

519 Despite the fact that *FI-t* and LAI show variation of similar magnitude, thus preserving a general
520 signal of spatial heterogeneity, spatial trends in *FI-t* ratio and LAI only partially match ($p < 0.001$, $R^2 =$
521 0.25 , linear regression analysis). They both decrease along T₁, following a moisture gradient (Fig. 6a),
522 with denser vegetation (with higher LAI, *FI-t*) likely favored in the quadrats closest to the water. Along
523 T₂, LAI and *FI-t* chiefly follow opposite trends. The *D/P* indices follow a substantively similar pattern,
524 also resulting in a mismatch between LAI, and D/P_1 and D/P_2 (see further below for a discussion of the *FI*
525 *-t* vs. *D/P* index comparison).

526 Several factors might explain the mismatches between *FI-t* or *D/P* indices and LAI at PV. First,
527 while *FI-t* and *D/P* values tend to be time-averaged, LAI values are only representative of the season
528 during which they were measured, that is, the 2016 dry season. Second, at PV, lateral transport favored
529 by strong winds might cause spatial mixing of phytolith assemblages (Powers et al., 2009). Third,
530 differential phytolith production or tree leaf litter production (e.g., deciduous versus evergreen tree
531 species) likely contribute. Our study supports previous work in showing that evergreen species deposit
532 less silica in their leaves than deciduous species (Cooke and Leishman, 2011). Therefore, the low *FI-t*

533 ratios (or *D/P* indices) in certain PV quadrats might relate to the fact that their dominant tree species
534 are non-producers or low to moderate producers (e.g., Q15, Q16, Q17, and Q21, as well as Q22 for the
535 *D/P* indices only), or evergreen (e.g., Q16, Q20, and Q25). Fourth, abundant understory grass cover—
536 independent of tree density—might explain low *FI-t* values in some quadrats. Although not observed
537 during field work (conducted in the dry season), the rocky limestone slopes of PV hills reportedly form
538 suitable habitats for grasses (Quirico Jiménez et al., 2016). Quadrats 15, 16, and 17 are located on these
539 slopes, and show low *FI-t* values despite high LAI. In contrast, quadrat 23 has among the highest *FI-t*
540 ratios paired with among the lowest LAI values. We hypothesize that low LAI values in quadrat 23
541 resulted because its predominantly deciduous trees had shed most of their leaves when LAI was
542 measured.

543 In rainforest soils, phytoliths from multiple samples provide a uniform signal in terms of canopy
544 openness, likely mirroring both relatively homogeneous canopy openness (i.e., invariant, high LAI) and
545 temporal averaging of soil phytoliths. At PV, extreme variations in *FI-t* values point to spatial variability
546 within the dry forest that can only be reconstructed through spatial sampling, but they exaggerate
547 actual differences in canopy openness between quadrats (see above). This raises the question whether
548 the lowest *FI-t* values from PV would overlap with those from other vegetation types that produce
549 grass-dominated soil phytolith assemblages (e.g., savannas). Although modern Neotropical savannas
550 have yet to be studied using *FI-t*, surface soil phytolith assemblages from *terra firme* woody savanna
551 (Cerradão) and *terra firme* savanna (*Cerrado*) show GSSCP percentages (mean = 56 and 61%
552 respectively; Dickau et al., 2013) within the range observed at PV (mean = 58%). However, whereas
553 arboreal markers mainly consist of SPHEROID ORNATE and SPHEROID ECHINATE, and of SPHEROID
554 ORNATE morphotypes, respectively, in the *terra firme* wooded savanna (Cerradão) and the *terra firme*
555 savanna (Cerrado), SPHEROID ECHINATE morphotypes are rare or absent in PV phytolith
556 assemblages. In addition, unlike in the Cerradão and Cerrado, phytoliths potentially produced by

557 Zingiberales and Bromeliaceae (Palm/Zingi/Brom) are present at relatively high abundance at
558 PV, and phytoliths produced by the grass subfamily Chloridoideae are extremely abundant
559 ($\leq 50\%$ of the GSSCP assemblage), allowing distinction of this dry forest from both wooded and
560 open savannas. Comparison of our data with the body of work in modern African vegetation types
561 similarly suggest overlap in GSSCP percentages and D/P index values of grass-dominated vegetation
562 types (e.g., savanna) within the range observed at PV. For example, in Congo (Bremond et al., 2005a),
563 the D/P values observed in shrub and tall grass savannas (recalculated from reported values to D/P_1 ,
564 $D/P_2 = 1.1-1.6$), and in savanna-forest transitions (recalculated from reported values to D/P_1 , $D/P_2 = 0.5-$
565 2.2) overlap with those observed in most PV quadrats (i.e., Q19, Q20, Q23, Q24, Q25), whereas values
566 reported from semi-deciduous forests (recalculated from reported values to $D/P_1 = 3-17.4$; $D/P_2 = 3-20.4$)
567 are substantially higher than those observed at Palo Verde. Likewise, in West Africa (Bremond et al.,
568 2005b), D/P values from several savanna biomes (Saharan, Sahelian, and Sudanian zones) (recalculated
569 from reported values to $D/P_1 = 0.01-0.96$; $D/P_2 = 0.01-1$) overlap with those observed at Palo Verde.
570 Finally, in Chad (Novello et al., 2012) D/P indices (recalculated from reported values to D/P_1 , $D/P_2 = 0-$
571 1.3) from grass-dominated vegetation types (i.e., woodlands, edaphic grassland mosaics, wooded
572 grasslands, and deciduous bushlands) overlap with those recorded in many of Palo Verde quadrats
573 (e.g., Q15, Q16, Q17, Q20, Q21, Q22, and Q25).

574 The overlap between D/P values of PV dry forest and these African grass-dominated
575 ecosystems (Barboni et al., 2007; Bremond et al., 2005a, 2005b; Novello et al., 2012) does not allow us
576 to differentiate these vegetation types based on the D/P index. However, floristic differences between
577 these ecosystems should nonetheless be reflected in phytolith assemblages. Indeed, despite the limited
578 taxonomic information from GSSCP assemblages provided in these studies, broad compositional
579 trends indicate major distinctions between these assemblages and Palo Verde samples. Among
580 arboreal markers, SPHEROID ORNATE are more abundant in the African phytolith assemblages

581 (Barboni et al., 2007; Bremond et al., 2005a, 2005b; Novello et al., 2012) than in the PV dry forest (see
582 also further discussion of the comparison between *D/P* index and *FI-t* ratio in section 4.5). Furthermore,
583 Palo Verde and the African GSSCP assemblages with overlapping *D/P* indices differ in term of relative
584 abundance of different GSSCP classes. Thus, at Palo Verde, GSSCP assemblages are dominated by
585 phytoliths diagnostic of the subfamily Chloridoideae (minimum average Chloridoideae = $33.6 \pm 11\%$)
586 whereas the higher proportions of BILOBATE GSSCP in many of the African assemblages (e.g., average
587 Chloridoideae = $20 \pm 14.6\%$ in grasslands-savannas in West Africa, $< 1\%$ in a savanna-forest
588 transition in Congo, and $6.8 \pm 3.3\%$ in Chad woodlands; Bremond et al., 2005a, 2005b; Novello et
589 al., 2012) suggests that grasses of the subfamily Chloridoideae are less abundant in these ecosystems.
590 Note, that direct, quantitative comparison of Chloridoideae phytolith (and other PFT) abundance
591 between PV and African samples is complicated by the use of different methodological
592 approaches among phytolith researchers. Still, the apparent differences in assemblage
593 composition highlight the potential of further phytolith work in modern Neotropical savannas to
594 establish guidelines for distinguishing this vegetation type from dry forest in the fossil record.

595 4.3.2 *Vegetation composition*

596 Observed floristic variation among LS quadrats cannot be linked to any of the measured
597 environmental or geographic gradients. The apparently random distribution of woody species might be
598 due to limitations of our sampling scheme and scale, such as quadrat and transect size relative to scale
599 of environmental gradients. Similarly, redundancy in phytolith production among forest indicator taxa
600 (e.g., Strömberg et al., 2018) and temporal averaging limit the variability in phytolith taxonomic signal,
601 especially when floristic gradients are subtle. Thus, we propose that additional sampling of distinct
602 microhabitats within the LS rainforest, such as riparian zones, forest gaps, and forest edges might yield
603 clearer differentiation in both species make-up and phytolith assemblages.

604 At PV, floral composition (represented by NMDS axes) covaries with both elevation (especially
605 within transects) and geographic position of quadrats, with adjacent quadrats and quadrats of similar
606 elevations (along T₁, T₂, or both) sharing similar plant species, whereas quadrats that are distant or at
607 different elevations show little floristic overlap. The only exceptions are the compositionally similar
608 quadrats Q₁₇ (T₁) and Q₂₄ (T₂). They are similarly distant from the river, but Q₂₄ is established on a
609 regeneration plot and might represent an outlier with its abundant, herbaceous understorey (C. Crifò,
610 2016, personal observation), which was not recorded in our vegetation census. For this reason, floristic
611 resemblance between the two quadrats is surprising but should be interpreted with caution.

612 Phytolith assemblage compositions (NMDS axes) show much less dramatic variation along PV
613 transects than floristic data and do not follow environmental gradients. This discrepancy may result
614 from redundancy in phytolith production and phytolith spatial averaging, both suppressing inter-
615 sample differences. Furthermore, our floral census and phytolith assemblages measure slightly
616 different aspects of vegetation composition. The floral census included only tree and palm species,
617 ignoring potentially important taxonomic differences among herbs, whereas phytolith assemblages
618 should capture some of the variation in herbaceous communities. If trees and herbs respond differently
619 across environmental gradients, the recorded compositional changes may also differ.

620 Overall, trends in floristic and phytolith assemblage composition along transects at LS and PV
621 suggest that the spatial resolution of the phytolith record is tied to vegetation type. Whereas in
622 rainforest habitats—at least at the scale studied here—lateral sampling is redundant because
623 vegetation gradients are not pronounced, in dry forest habitats, spatial sampling is essential for
624 detecting overall heterogeneity.

625 *4.4 Analytical approaches to deep-time vegetation inference*

626 The “general” analytical approach taken in our study includes the use of a set of broad phytolith
627 PFT, and was explicitly designed to be applied to deep-time assemblages without clear modern

628 analogues (Strömberg, 2003, 2004; Strömberg et al., 2018). As stated earlier, this method assigns
629 morphotypes to PFT or taxonomic group based on their observed, overall distribution among plants,
630 regardless of known local flora. In the case of LS and PV, this resulted in the identification of
631 morphotypes linked to PFT/taxonomic groups that are clearly not represented, such as GSSCP
632 diagnostic of Pooideae grasses ("POOID-D"). Pooideae, which today mainly occur in (cool) temperate
633 climates, have only been reported from high elevation ecosystems (i.e., > 3,000 m a.s.l.), such as
634 paramos, in Costa Rica (Bremond et al., 2012; Chazdon, 1978) and it is therefore reasonable to assume
635 that the observed "POOID-D" morphotypes were produced by non-poid grasses (e.g., in the subfamily
636 Panicoideae; OTS, 2020; Quirico Jiménez et al., 2016). Nevertheless, because of the very low relative
637 abundances of these morphotypes in GSSCP assemblages, their "erroneous" PFT assignment does not
638 impact grass community reconstructions at LS and PV. Similarly, we find it unlikely that they would
639 affect inferences in deep time, pointing to the robustness of the "general approach".

640 4.5 Comparison of the *FI-t* ratio with the *D/P* index

641 Our study compared two different, commonly used metrics for tree cover, the *FI-t* ratio (part of
642 the "general approach"; Strömberg et al., 2018) and the *D/P* index (e.g., Alexandre et al., 1997;
643 Bremond et al., 2005a, 2008). We found that both sets of indices (*FI-t*, *D/P*) varied substantially across
644 PV phytolith assemblages, resulting in values that would typically be associated with different
645 vegetation types; however, the *FI-t* ratio overall performs better. Thus, the *D/P* indices (D/P_1 0.06–1.8;
646 D/P_2 0.06–2.2) recorded in PV quadrats would typically be reconstructed as steppe to savanna
647 vegetation (Bremond et al., 2005b), much more open than the dry forest at PV and more open than
648 how the corresponding *FI-t* ratio values would usually be interpreted (savanna to forest; e.g.,
649 Strömberg et al., 2007a, 2013, 2018); although note that *FI-t* was not calibrated on modern floras,
650 hindering direct comparisons with *D/P*. In contrast, phytolith assemblages from a semi-deciduous
651 forest in Congo yielded significantly higher *D/P* values (average $D/P_1 = 2.9 \pm 0.9$ and $D/P_2 = 7.9 \pm 4$

652 (Bremond et al., 2005a). This large discrepancy in *D/P* values between different dry forests is likely due
653 to substantial differences in phytolith assemblage composition between ecosystems in Africa (where
654 the *D/P* index was calibrated) and the Neotropics, reflecting biogeographical floristic differences. For
655 example, palm species and their diagnostic phytolith morphotypes are more abundant in the
656 Neotropics than in African ecosystems (Muscarella et al., 2020). It may also be that the woody taxa that
657 produce the SPHEROID ORNATE morphotype (i.e., the numerator of the *D/P* index), are more abundant in
658 African than in Neotropical soils; this remains to be tested. At our sites, SPHEROID ORNATE morphotypes
659 were generally present at moderate but highly variable abundance, potentially reflecting floristic
660 differences between quadrats rather than differences in canopy openness.

661 As a result of these difference between African and Neotropical vegetation and soil phytolith
662 assemblages, we suggest that in Neotropical ecosystems, palm morphotypes should be used not only
663 as floristic indicators but also as an important variable within indices of vegetation structure, such as
664 the *D/P* index (as previously suggested by Dickau et al., 2013). Our study, which constitutes the first test
665 of the “general approach” in modern ecosystems support its utility in the Neotropics. However, we note
666 that more modern analog work using this method (and the *Fl-t* ratio) is needed to better understand its
667 ability to reconstruct vegetation types in absolute terms (as opposed to relative changes; see
668 Strömberg et al., 2018) and its performance across ecosystems and geographic areas.

669

670 5. Conclusions

671 Our study is the first to characterize and compare phytolith assemblages across vegetation
672 transects in different Neotropical vegetation types using phytoliths exclusively from subsurface soil
673 samples, which tend to be temporally averaged, but represent a better analogue to pre-Quaternary
674 phytolith assemblages. It also differs from previous work in its use of the *Fl-t* ratio and phytolith-based
675 plant functional types (PFT) coupled with vegetation surveys, and a focus on multiple aspects of

676 vegetation communities (i.e., structure, composition, and diversity). We show that, in spite of likely
677 time- and/or spatial averaging, soil subsurface assemblages preserve information about overall
678 vegetation structure and composition, as well as the degree to which they vary spatially. These results
679 highlight the clear utility of phytolith analysis to differentiate past vegetation formations based on
680 major structural and compositional differences. Accordingly, the homogeneous rainforest vegetation of
681 La Selva produces a very homogeneous soil phytolith signal that also reflects the extremely high palm
682 abundance in this rainforest, whereas the more heterogeneous dry forest vegetation of Palo Verde
683 tends to produce phytolith assemblages characterized by variable proportion of grass- and forest
684 indicator phytolith. A spatially variable phytolith signal in the fossil record can therefore be reasonably
685 interpreted as representing a heterogeneous habitat, confirming that analysis of multiple phytolith
686 samples collected along chronologically constrained stratigraphic levels form a vital tool for
687 reconstructing local vegetation heterogeneity and for refining comparisons of overall vegetation types
688 through time and space. Our work also suggests that other factors, such as phytolith production and
689 taphonomy contribute to shaping the phytolith assemblage signal in ways that vary according to
690 vegetation type. For example, the presence of taxa that produce fewer phytoliths (e.g., evergreen
691 species), potential phytolith transport, or particular soil conditions (i.e., alkaline pH) in some quadrats
692 might result in underrepresentation of forest indicator morphotypes. In future studies, even more
693 detailed spatial gradients in vegetation type might be tractable by coupling phytolith analysis with
694 other proxies for canopy openness, (e.g., stable carbon isotopes, phytolith undulation index) and/or
695 vegetation (e.g., pollen), and by including herbaceous species in vegetation surveys in order to estimate
696 the impact of herbaceous communities on phytolith assemblages.

697

698 **Acknowledgments**

699 Samples were collected, exported, and imported under research permit #006-2016-INV-
700 ACAT, export permits #DGVS-076- 2016, #DGVS-082-2016, and import permit USDA#P330–
701 14-00285. Assistance with research and export permits was provided by MINAE (Ministry of the
702 Environment and Energy) Francisco Campos, and Davinia Beneyto. We are grateful to the OTS
703 staff of La Selva and Palo Verde Biological stations for helping with field work logistics. We
704 thank the Burke Museum of Natural History and Culture (UWBM), Ron Eng, Katie Anderson,
705 UWBM Botany Collections (WTU herbarium), David Giblin and Soojeong Shin, at the University
706 of Washington (Seattle) for helping with sample cataloguing. We also thank undergraduate
707 research assistants Claire Grant and Kailyn Zard for their help processing modern plant
708 samples; and Janneke Hille Ris Lambers, and Richard Olmstead for comments on an earlier
709 version of this manuscript. This research was supported by American Philosophical Society
710 [C.C.], Geological Society of America [C.C.], Organization for Tropical Studies [C.C.],
711 Paleontological Society [C.C.], Quaternary Research Center [C.C.], University of Washington
712 [C.C.], Washington Research Foundation [C.C.] and National Science Foundation [EAR-
713 1253713 and EAR-1349530; C.A.E.S]. This work partially fulfilled the Ph.D. in Biology degree
714 requirements of C.C. at University of Washington.

715

References

- 716 Aleman, J.C., Leys, B., Apema, R., Bentaleb, I., Dubois, M.A., Lamba, B., Lebamba, J., Martin,
717 C., Ngomanda, A., Truc, L., Yangakola, J.-M., Favier, C., Bremond, L., Woods, K., 2012.
718 Reconstructing savanna tree cover from pollen, phytoliths and stable carbon isotopes.
719 *Journal of Vegetation Science* 23, 187-197.
- 720 Alexandre, A., Meunier, J.-D., Lezine, A.-M., Vincens, A., Schwartz, D., 1997. Phytoliths:
721 indicators of grassland dynamics during the late Holocene in intertropical Africa.
722 *Palaeogeography, Palaeoclimatology, Palaeoecology* 136, 213-229.
- 723 Alvarado Induni, G.E., 1990. Características geológicas de la Estación Biológica La Selva,
724 Costa Rica. *Tecnología en marcha* 10, 11-22.
- 725 Barboni, D., Bremond, L., Bonnefille, R., 2007. Comparative study of modern phytolith
726 assemblages from inter-tropical Africa. *Palaeogeography, Palaeoclimatology,*
727 *Palaeoecology* 246, 454-470.
- 728 Bartoli, F., Wilding, L.P., 1980. Dissolution of biogenic opal as a function of its physical and
729 chemical properties. *Soil Science Society of America Journal* 44, 873-878.
- 730 Bremond, L., Alexandre, A., Hély, C., Guiot, J., 2005a. A phytolith index as a proxy of tree cover
731 density in tropical areas: calibration with Leaf Area Index along a forest–savanna
732 transect in southeastern Cameroon. *Global and Planetary Change* 45, 277-293.
- 733 Bremond, L., Alexandre, A., Peyron, O., Guiot, J., 2005b. Grass water stress estimated from
734 phytoliths in West Africa. *Journal of Biogeography* 32, 311-327.
- 735 Bremond, L., Alexandre, A., Peyron, O., Guiot, J., 2008. Definition of grassland biomes from
736 phytoliths in West Africa. *Journal of Biogeography* 35, 2039-2048.
- 737 Bremond, L., Boom, A., Favier, C., 2012. Neotropical C3/C4 grass distributions – present, past
738 and future. *Global Change Biology* 18, 2324-2334.
- 739 Chazdon, R.L., 1978. Ecological aspects of the distribution of C4 grasses in selected habitats of
740 Costa Rica. *Biotropica*, 265-269.

741 Clark, D.B., Olivas, P.C., Oberbauer, S.F., Clark, D.A., Ryan, M.G., 2008. First direct
742 landscape-scale measurement of tropical rain forest Leaf Area Index, a key driver of
743 global primary productivity. *Ecol Lett* 11, 163-172.

744 Cooke, J., Leishman, M.R., 2011. Silicon concentration and leaf longevity: is silicon a player in
745 the leaf dry mass spectrum? *Functional Ecology* 25, 1181-1188.

746 Crifò, C., Strömberg, C.A.E., 2020. Small-scale spatial resolution of the soil phytolith record in a
747 rainforest and a dry forest in Costa Rica: applications to the deep-time fossil phytolith
748 record. *Palaeogeography, Palaeoclimatology, Palaeoecology* 537, 109107.

749 Dickau, R., Whitney, B.S., Iriarte, J., Mayle, F.E., Soto, J.D., Metcalfe, P., Street-Perrott, F.A.,
750 Loader, N.J., Ficken, K.J., Killeen, T.J., 2013. Differentiation of neotropical ecosystems
751 by modern soil phytolith assemblages and its implications for palaeoenvironmental and
752 archaeological reconstructions. *Review of Palaeobotany and Palynology* 193, 15-37.

753 Dunn, R.E., Strömberg, C.A.E., Madden, R.H., Kohn, M.J., Carlini, A.A., 2015. Linked canopy,
754 climate, and faunal change in the Cenozoic of Patagonia. *Science* 347, 258-261.

755 Fredlund, G.G., Tieszen, L.T., 1994. Modern Phytolith Assemblages from the North American
756 Great Plains. *Journal of Biogeography* 21, 321-335.

757 Gentry, A.H., 1993. *Four neotropical rainforests*. Yale University Press.

758 Gillespie, T.W., Grijalva, A., Farris, C.N., 2000. Diversity, composition, and structure of tropical
759 dry forests in Central America. *Plant Ecology* 147, 37-47.

760 Harris, E.B., Strömberg, C.A.E.S., N.D., Smith, S.Y., Vilhena, D.A., 2017. Vegetation response
761 during the lead-up to the middle Miocene warming event in the Northern Rocky
762 Mountains, USA. *Palaeogeography, Palaeoclimatology, Palaeoecology* 485, 401-415.

763 Hodson, M.J., White, P.J., Mead, A., Broadley, M.R., 2005. Phylogenetic variation in the silicon
764 composition of plants. *Ann Bot* 96, 1027-1046.

765 Hyland, E., Smith, S.Y., Sheldon, N.D., 2013. Representational bias in phytoliths from modern
766 soils of central North America: Implications for paleovegetation reconstructions.
767 *Palaeogeography, Palaeoclimatology, Palaeoecology* 374, 338-348.

768 Hyland, E.G., Sheldon, N.D., Cotton, J.M., 2015. Terrestrial evidence for a two-stage mid-
769 Paleocene biotic event. *Palaeogeography, palaeoclimatology, palaeoecology* 417, 371-
770 378.

771 Hyland, E.G., Sheldon, N.D., Smith, S.Y., Strömberg, C.A.E., 2018. Late Miocene rise and fall of
772 C4 grasses in the western United States linked to aridification and uplift. *GSA Bulletin*
773 131, 224-234.

774 Kalacska, M., Calvo-Alvarado, J.C., Sanchez-Azofeifa, G.A., 2005. Calibration and assessment
775 of seasonal changes in leaf area index of a tropical dry forest in different stages of
776 succession. *Tree Physiol* 25, 733-744.

777 Kerns, B.K., Moore, M.M., Hart, S.C., 2001. Estimating forest-grassland dynamics using soil
778 phytolith assemblages and $\delta^{13}C$ of soil organic matter. *Écoscience* 8, 478-488.

779 Kruskal, J.B., 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric
780 hypothesis. *Psychometrika* 29, 1-27.

781 Loughney, K.M., Hren, M.T., Smith, S.Y., Pappas, J.L., 2020. Vegetation and habitat change in
782 southern California through the Middle Miocene Climatic Optimum: Paleoenvironmental
783 records from the Barstow Formation, Mojave Desert, USA. *GSA Bulletin* 132, 113-129.

784 Mercader, J., Bennett, T., Esselmont, C., Simpson, S., Walde, D., 2011. Soil phytoliths from
785 miombo woodlands in Mozambique. *Quaternary Research* 75, 138-150.

786 Mulholland, S.C., 1989. Phytolith shape frequencies in North Dakota grasses: a comparison to
787 general patterns. *Journal of Archaeological Science* 16, 489-511.

788 Murphy, P.G., Lugo, A.E., 1986. Ecology of tropical dry forest. *Annual review of ecology and*
789 *systematics* 17, 67-88.

790 Muscarella, R., Emilio, T., Phillips, O.L., Lewis, S.L., Slik, F., Baker, W.J., Couvreur, T.L.,
791 Eiserhardt, W.L., Svenning, J.C., Affum-Baffoe, K., 2020. The global abundance of tree
792 palms. *Global Ecology and Biogeography* 29, 1495-1514.

793 Neumann, K., Strömberg, C.A.E., Ball, T., Albert, R.M., Vrydaghs, L., Cummings, L.S., 2019.
794 International code for phytolith nomenclature (ICPN) 2.0. *Annals of Botany* 124, 189-
795 199.

796 Novello, A., Barboni, D., Berti-Equille, L., Mazur, J.-C., Poilecot, P., Vignaud, P., 2012. Phytolith
797 signal of aquatic plants and soils in Chad, Central Africa. *Review of Palaeobotany and*
798 *Palynology* 178, 43-58.

799 Pearsall, D.M., 2000. *Paleoethnobotany: a handbook of procedures* Academic Press. Google
800 Scholar.

801 Piperno, D.R., 1985. Phytolith taphonomy and distributions in archaeological sediments from
802 Panama. *Journal of Archaeological Science* 12, 247-267.

803 Piperno, D.R., 1988. *Phytolith analysis: an archaeological and geological perspective*.
804 Academic Press, San Diego.

805 Piperno, D.R., 2006. *Phytoliths: a comprehensive guide for archaeologists and paleoecologists*.
806 Rowman Altamira.

807 Powers, J.S., Becknell, J.M., Irving, J., Pèrez-Aviles, D., 2009. Diversity and structure of
808 regenerating tropical dry forests in Costa Rica: Geographic patterns and environmental
809 drivers. *Forest Ecology and Management* 258, 959-970.

810 Quírico Jiménez, M., Carrillo, E., Kappelle, M., 2016. The Northern Pacific Lowland Seasonal
811 Dry Forests of Guanacaste and the Nicoya Peninsula, in: Kappelle, M. (Ed.), *Costa*
812 *Rican Ecosystems*. The University of Chicago Press, Chicago and London, pp. 247-289.

813 R Development Core Team, 2019. *R: A language and environment for statistical computing*. R
814 Foundation for Statistical Computing, Vienna, Austria.

815 Rovner, I., 1971. Potential of Opal Phytoliths for use in Paleoeological Reconstruction.
816 Quaternary Research 1, 343-359.

817 Strömberg, C.A.E., 2003. The origin and spread of grass-dominated ecosystems during the
818 Tertiary of North America and how it relates to the evolution of hypsodonty in equids.
819 University of California, Berkeley.

820 Strömberg, C.A.E., 2004. Using phytolith assemblages to reconstruct the origin and spread of
821 grass-dominated habitats in the great plains of North America during the late Eocene to
822 early Miocene. *Palaeogeography, Palaeoclimatology, Palaeoecology* 207, 239-275.

823 Strömberg, C.A.E., 2005. Decoupled taxonomic radiation and ecological expansion of open-
824 habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy*
825 *of Sciences of the United States of America* 102, 11980-11984.

826 Strömberg, C.A.E., 2009. Methodological concerns for analysis of phytolith assemblages: Does
827 count size matter? *Quaternary International* 193, 124-140.

828 Strömberg, C.A.E., Dunn, R.E., Crifò, C., Harris, E.B., 2018. Phytoliths in paleoecology:
829 analytical considerations, current use, and future directions, in: Croft, D.A., Simpson,
830 S.W., Su, D.F. (Eds.), *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial*
831 *Environments and Ecological Communities*. Springer Cham, Switzerland, pp. 233-285.

832 Strömberg, C.A.E., Dunn, R.E., Madden, R.H., Kohn, M.J., Carlini, A.A., 2013. Decoupling the
833 spread of grasslands from the evolution of grazer-type herbivores in South America. *Nat*
834 *Commun* 4, 1478.

835 Strömberg, C.A.E., Werdelin, L., Friis, E.M., Sarac, G., 2007a. The spread of grass-dominated
836 habitats in Turkey and surrounding areas during the Cenozoic: Phytolith evidence.
837 *Palaeogeography Palaeoclimatology Palaeoecology* 250, 18-49.

838 Strömberg, C.A.E., Werdelin, L., Friis, E.M., Saraç, G., 2007b. The spread of grass-dominated
839 habitats in Turkey and surrounding areas during the Cenozoic: Phytolith evidence.
840 *Palaeogeography, Palaeoclimatology, Palaeoecology* 250, 18-49.

- 841 Watling, J., Iriarte, J., Whitney, B.S., Consuelo, E., Mayle, F., Castro, W., Schaan, D.,
842 Feldpausch, T.R., 2016. Differentiation of neotropical ecosystems by modern soil
843 phytolith assemblages and its implications for palaeoenvironmental and archaeological
844 reconstructions II: Southwestern Amazonian forests. *Review of Palaeobotany and*
845 *Palynology* 226, 30-43.
- 846 Weil, R., Brady, N.C., 2016. *The nature and properties of soils: Pearson new international*
847 *edition*. Pearson Higher Ed.
- 848 Zhao, Z., Pearsall, D.M., 1998. Experiments for Improving Phytolith Extraction from Soils.
849 *Journal of Archaeological Science* 25, 587-598.

850

Table captions

851 **Table 1.** Differences between palynological, macrofloral and phytolith deep-time fossil assemblages,
852 with focus on Neotropical environments.

853

854 **Table 2.** List of phytolith types included in this study, phytolith-based plant functional types (PFT),
855 producing taxa, vegetation type for which they are diagnostic, and ecology. Full names for the
856 acronyms in column 2 are given in column 3.

857

858 **Table 3.** Phytolith yield rank of species present in Palo Verde quadrats, their leaf phenology and relative
859 contribution to the total tree biomass (BA_{sp}), total deciduous biomass in each quadrat (%), and
860 estimated $Fl-t$ ratio, and LAI values for each quadrat.

861

Figure captions

862 **Figure 1.** Map of the study sites. Map was created using the biome layer in ArcGIS® software by ©Esri.

863

864 **Figure 2.** Violin plot of phytolith yield (Y_{phyto}) of Palo Verde woody deciduous (yellow), semi-deciduous

865 (blue), and evergreen (green) species. Numbers in parenthesis indicate sample size. Statistically

866 significant differences between groups (deciduous and evergreen species) are indicated by the symbol

867 *; n.s. indicates non-significant difference between groups (deciduous and semi-deciduous, and semi-

868 deciduous and evergreen species). (For interpretation of the references to color in this figure

869 caption, the reader is referred to the web version of this article.)

870

871 **Figure 3.** Phytolith assemblages from **A)** La Selva and **B)** Palo Verde. The upper graphs show the $FI-t$

872 ratio ($FI-t_{quadrat}$) expressed as the percent forest indicator (FI) phytoliths in a sum of FI and grass

873 phytolith morphotypes ($\%FI/(FI+GSSCP)$). Diamonds represents estimates of $FI-t$ ratios from phytolith

874 counts; vertical bars represent 95% confidence intervals based on bootstrap analysis of each sample.

875 Grey dotted lines represent average $FI-t$ ratios ($FI-t_{site}$). The bar graphs (lower part of the figure)

876 represent phytolith assemblage composition of the samples indicating percent phytoliths in each of the

877 seven defined functional groups (in legend). Percent forest indicators in the upper graphs refer to the

878 sum of all forest indicators (green and blue gradient) in the legend. Pie charts (middle part of the figure)

879 represent relative abundance of grass PFT within GSSCP (in yellow in the bar graphs) in the

880 assemblages. Number in parenthesis above pie charts represent the number of GSSCP counted, if less

881 than 200. "POOID" = "POOID-D" + POOID-ND. BOE = Bambusoideae, Oryzoideae and members of the

882 early-diverging grass clades Anomochlooideae, Pharoideae, and Puelioideae. (For interpretation of

883 the references to color in this figure caption, the reader is referred to the web version of this

884 article.)

885

886 **Figure 4.** Compositional differences of soil phytolith assemblages between the rainforest and dry
887 forest. **A)** NMDS ordination (stress value = 0.047) of the soil samples across La Selva (Q1-Q14) and Palo
888 Verde (Q15-Q25) quadrats, and phytolith classes (black vectors). Only phytolith classes (i.e., compound
889 variables) with significant ($p < 0.01$) loadings on the ordination axes are shown here. Vector direction
890 indicates maximum correlation between sample scores in the quadrats (soil phytolith assemblages) and
891 the phytolith classes (PFT). **B)** NMDS ordination [as in A)], overlaid by the dendrogram of the quadrats
892 at the two sites obtained from cluster analysis based on phytolith PFT. Blue squares represent La Selva
893 quadrats; dark green and light green circles represent Palo Verde quadrats belonging to, respectively,
894 transect 1 (T1) and 2 (T2). (For interpretation of the references to color in this figure caption, the
895 reader is referred to the web version of this article.

896

897 **Figure 5.** Vegetation and phytolith gradients at La Selva, with quadrats ordered according to **A), C)**
898 increasing distance from the water, and **B), D)** increasing elevation. **A), B)** Observed (Leaf Area Index;
899 LAI; solid line) and estimated (FI-t ratio; dotted line) vegetation structure. **C), D)** Observed (solid line),
900 and estimated (dotted line) vegetation composition (represented by loadings on axis 1 of the NMDS
901 ordination according to, respectively, plant species composition and phytolith morphotype
902 composition). Numbers indicate quadrat numbers.

903

904 **Figure 6.** Vegetation and phytolith gradients at Palo Verde transect 1 (T1) and 2 (T2), with quadrats
905 ordered according to **A), C)** increasing distance from the water, and **B), D)** increasing elevation. **A), B)**
906 Observed (Leaf Area Index; LAI, solid line) and estimated (FI-t ratio; dotted line) vegetation structure.
907 **C), D)** Observed (solid line), and estimated (dotted line) vegetation composition (represented by

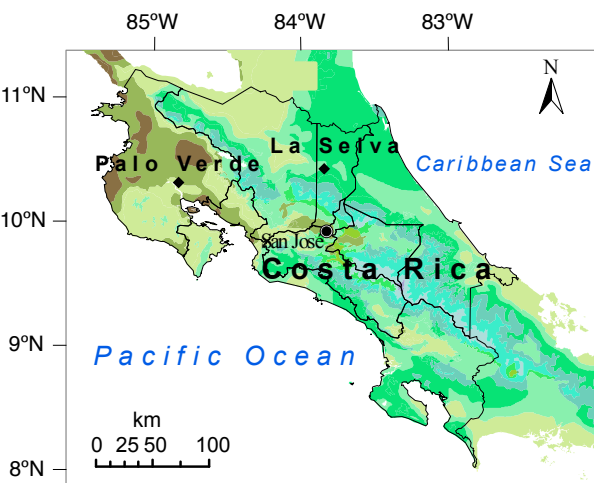
908 loadings on axis 1 of the NMDS ordination according to plant species composition, and phytolith
909 morphotype composition respectively). Numbers indicate quadrat numbers.













910

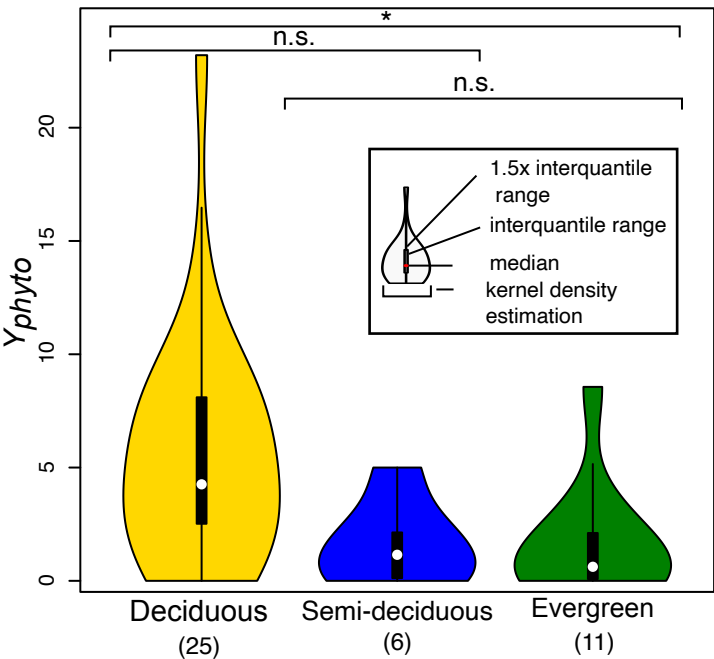
911 **Figure 7.** Compositional variation of soil phytolith assemblages within forest types. **A)** NMDS
912 ordinations of the soil samples across La Selva quadrats, and phytolith PFT (black vectors). **B)** NMDS
913 ordinations of the soil samples across Palo Verde quadrats, and phytolith PFT (black vectors), with
914 superimposed dendrogram (grey lines) obtained by cluster analysis. Only phytolith PFT with significant
915 ($p < 0.01$) loadings on the ordination axes are shown here. The direction of the vectors indicates
916 maximum correlation between the score of the samples in the quadrats (soil phytolith assemblages)
917 and the phytolith PFT.

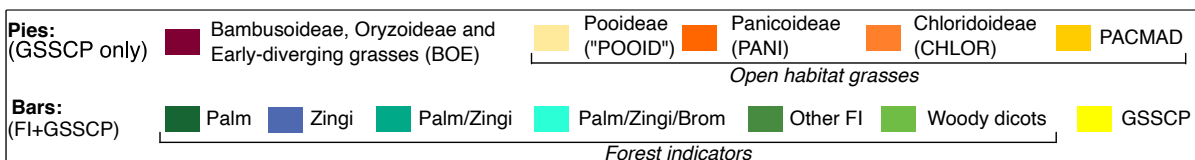
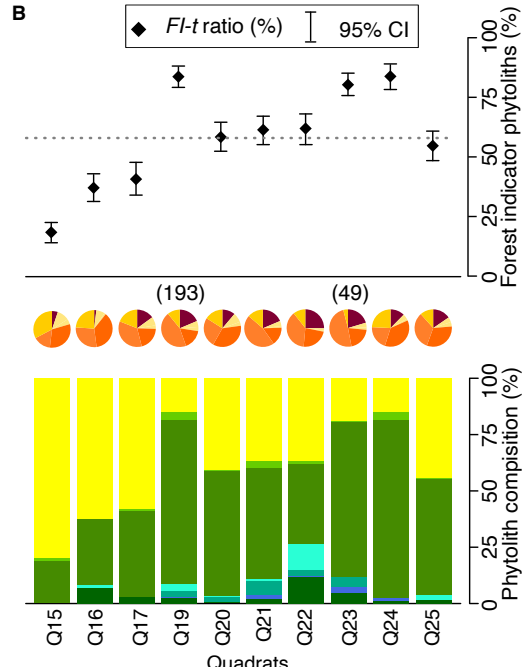
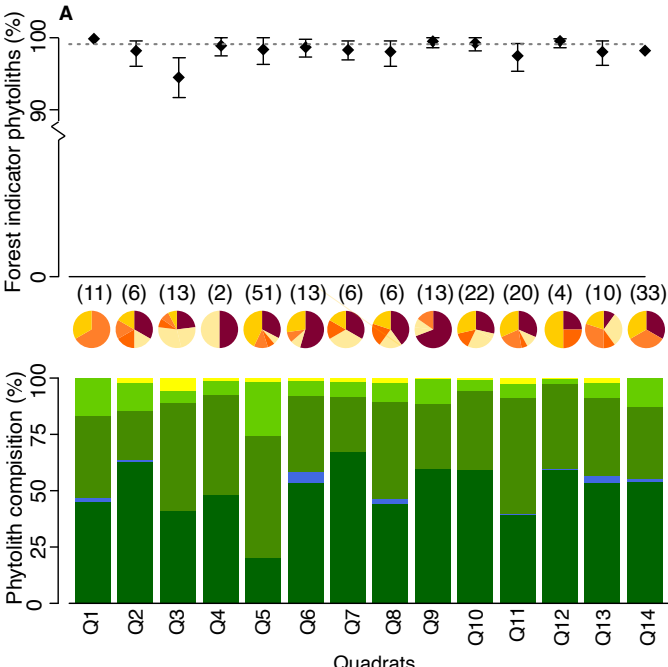
918 **Data Availability Statement**

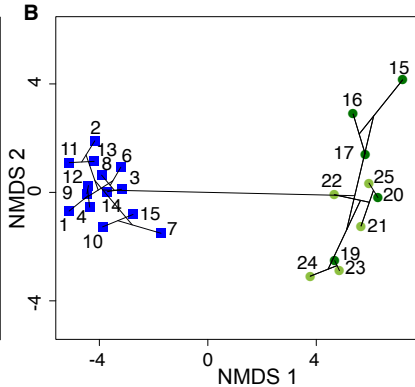
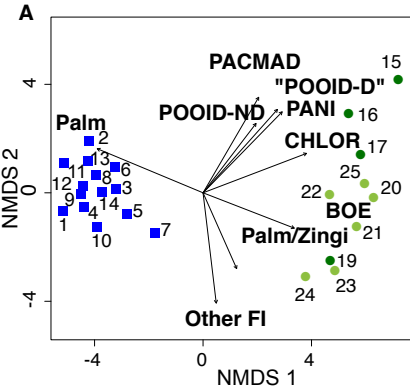
919 All data are available in the Supporting Information (SI 1-11).

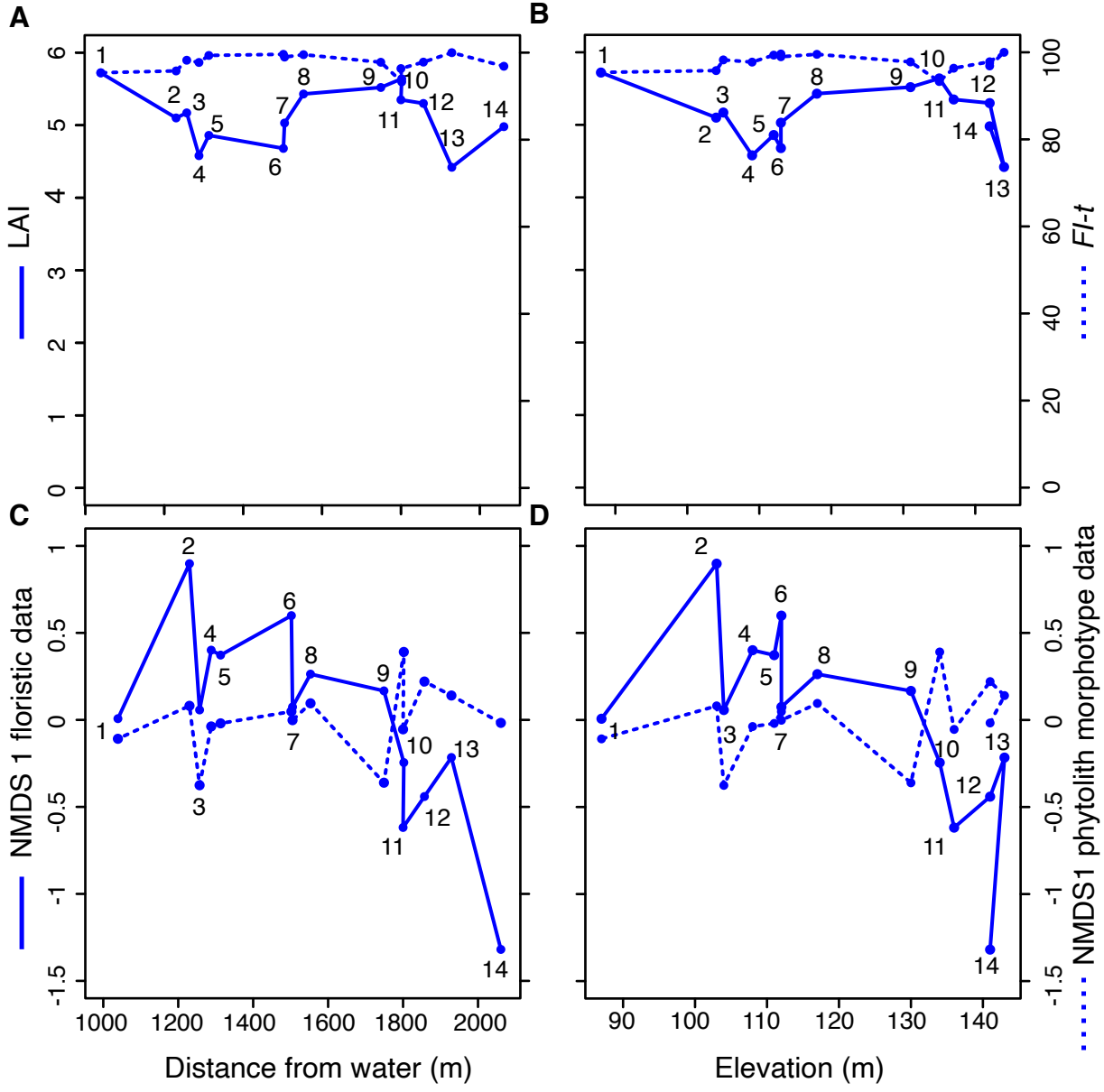


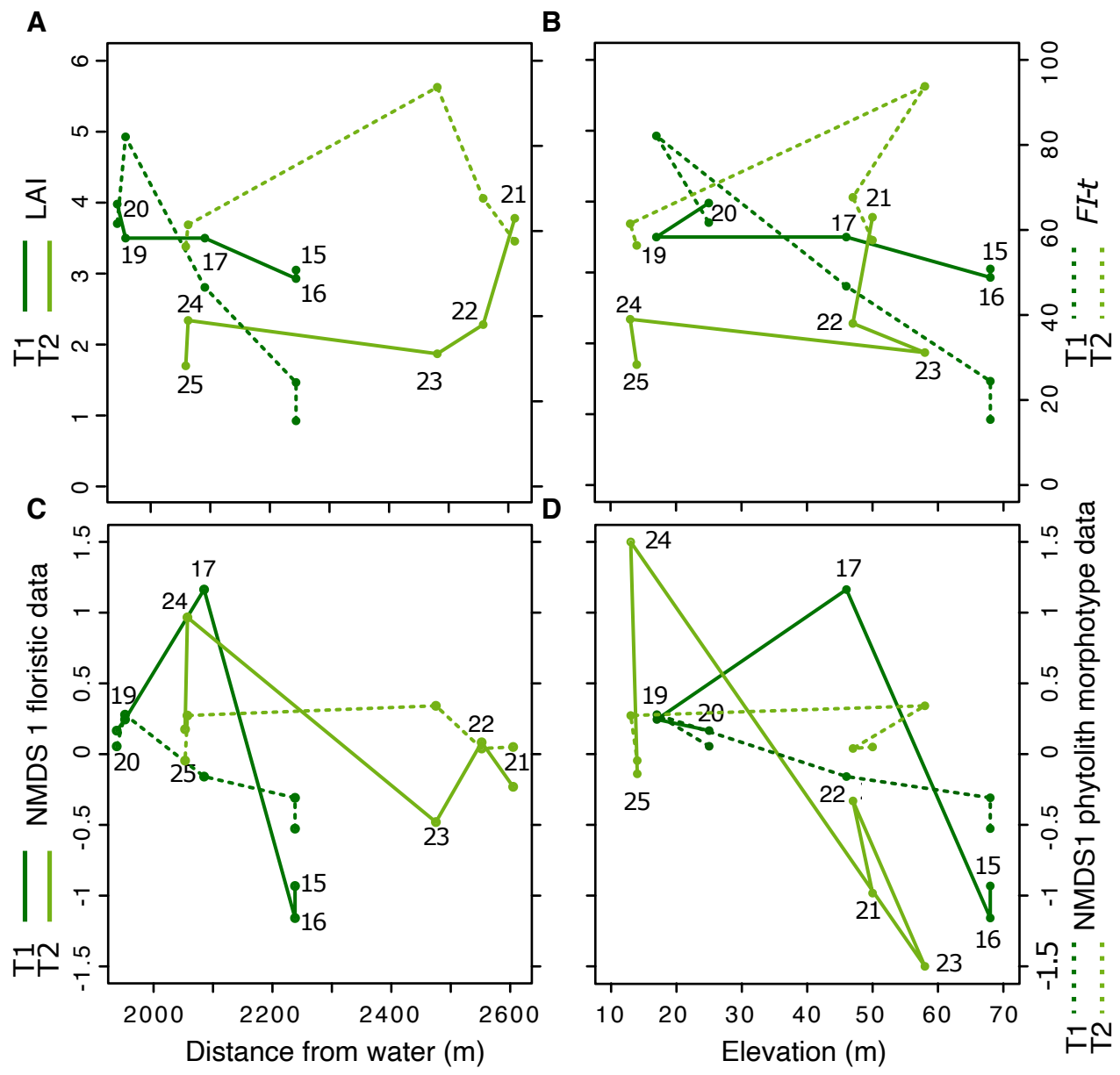
- | | |
|---|--|
|  Tropical dry forest |  Lower montane moist forest |
|  Tropical moist forest |  Lower montane wet forest |
|  Tropical moist-dry forest transition |  Lower montane rainforest |
|  Tropical wet forest |  Montane wet forest |
|  Premontane wet forest |  Montane rainforest |
|  Premontane rainforest |  Subalpine rain paramo |











	Pollen and spores	Leaf macrofossils	Phytoliths
Spatial resolution	Regional	Local	Local-regional*
Temporal resolution	1–many years	Days-years	1-10 ³ years**
Type of inferences	Arboreal cover, species richness, qualitative climate reconstruction	Species richness, ecophysiology, quantitative climate reconstruction	Arboreal cover, grass-woody plant dynamics, plant functional groups, grass community composition

*likely dependent on habitat type, **dependent on depositional environment: in paleosols usually 10²-10³ years (see Strömberg et al., 2018), in lacustrine environments 1-many years (same as for pollen)

Phytolith types	#	PFT acronym	PFT full name	Producing taxa	Vegetation type	Climate preferences
Non-GSSCP	1	Palm	Palms	Arecaceae	FI	Mostly wet, tropical
	2	Woody dicots	Woody dicotyledons	Woody dicotyledonous angiosperms	FI	Various
	3	Other FI	Other forest indicators	Woody and herbaceous dicotyledonous angiosperms, conifers, ferns, cycads, ginkgo	FI	Various
	4	Zingi	Zingiberales	Zingiberales	FI	Mostly wet, tropical
	5	Palm/Zingi	Palms or Zingiberales	Arecaceae and Zingiberales	FI	Mostly wet, tropical
	6	Palm/Zingi/Brom	Palms, Zingiberales, Bromeliads	Arecaceae, Zingiberales, Bromeliaceae	FI	Mostly wet, tropical
	7	OTH	Other groups	Unknown or multiples taxa	N/A	N/A
GSSCP	8	BOE	Bambusoideae, Oryzoideae, and early diverging grasses (C ₃)	Bambusoideae, Oryzoideae, and early-diverging grasses in subfamilies Anomochlooideae, Puelioideae, and Pharioideae	FI (wetlands)	Mostly wet, tropical–temperate
	9	“POOID-D”	Diagnostic Pooideae (C ₃)	Pooideae (but possible redundancy with, e.g., PACMAD morphotypes)	Open habitat	Arid, temperate
	10	POOID-ND	Non-diagnostic Pooideae (C ₃)	Mainly Pooideae but other grass subfamilies too	Open habitat	Arid, temperate
	11	CHLOR	Chloridoideae (C ₄)	Chloridoideae	Open habitat	Arid, warm–tropical
	12	PANI	Panicoideae (mostly C ₄)	Panicoideae	Open habitat	Rel. wet, tropical
	13	PACMAD	Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, and Danthonioideae (C ₃ and C ₄)	Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, and Danthonioideae	Open habitat	Wet–arid, warm–tropical
	14	OTHG	Other grasses	Unknown or multiple grass taxa	N/A	N/A

Quadrat	Yield Rank ^a	Species ^b	BA _{sp} (%) ^c	% deciduous biomass ^d	Estimated FI-t ratio ^e	LAI ^f
	0	<i>Machaerium biovolatum</i> ¹ (24.33)	24.33			
	1	<i>Aphelandra scabra</i> ³ (1.87), <i>Astronium graveolens</i> ² (11.8), <i>Erythroxylum havanense</i> ¹ (4.38)	18.05			
15	2	<i>Casearia corymbosa</i> ¹ (0.12), <i>Tabebuia ochracea</i> ¹ (5.54)	5.66	~85%	19.16	3.1
	3	<i>Acacia centralis</i> ¹ (50.64), <i>Lonchocarpus parviflorus</i> ¹ (0.32)	50.96			
	4	<i>Alophylus racemosus</i> ¹ (1)	1			
	0	n.a.	0			
	1	<i>Erythroxylum havanense</i> ¹ (4.15), <i>Manilkara chicle</i> ³ (2.66)	6.81			
16	2	<i>Casearia corymbosa</i> ¹ (0.46), <i>Luehea candida</i> ² (2.17), <i>Tabebuia ochracea</i> ¹ (10.3)	12.93	~77%	37.83	2.9
	3	<i>Cordia allococca</i> ¹ (0.84), <i>Guazuma ulmifolia</i> ¹ (16.15), <i>Lonchocarpus fasciculatus</i> ¹ (16.43), <i>Lonchocarpus parviflorus</i> ¹ (2.31)	35.73			
	4	<i>Alophylus racemosus</i> ¹ (24.24), <i>Cordia dentata</i> ² (0.16), Indetermined Malvaceae (2.53), <i>Simarouba glauca</i> ³ (17.6)	44.53			
	0	<i>Coccoloba caracasana</i> ² (5.86), <i>Thouinidium decandendrum</i> ² (0.84)	4.32			
	1	<i>Acacia collinsii</i> ¹ (0.44), <i>Astronium graveolens</i> ² (7.28), <i>Manilkara chicle</i> ³ (0.11), <i>Triplaris melanodendron</i> ³ (76.87)	86.85			
17	2	<i>Psychotria carthagenensis</i> ³ (0.48)	0.49	~9%	41.46	3.5
	3	n.a.	0			
	4	<i>Alophylus racemosus</i> ¹ (0.5), <i>Spondias mombin</i> ¹ (7.62)	8.34			
	0	n.a.	0			
	1	<i>Astronium graveolens</i> ² (5.01)	5.01			
19	2	<i>Calycophyllum candidissimum</i> ¹ (6.33), <i>Capparidastrum frondosum</i> ³ (0.03)	6.36	~91%	84.45	3.5
	3	<i>Acacia centralis</i> ¹ (64.3), <i>Guazuma ulmifolia</i> ¹ (19.96), <i>Lonchocarpus parviflorus</i> ¹ (0.3), <i>Rourea glabra</i> ³ (0.65), <i>Tabebuia rosea</i> ¹ (0.21)	85.42			
	4	<i>Garcia nutans</i> ³ (3.21)	3.21			
	0	n.a.	0			
	1	<i>Manilkara chicle</i> ³ (1.77)	1.77			
20	2	<i>Tabebuia ochracea</i> ¹ (0.89)	0.89	~1%	59.13	4
	3	<i>Rourea glabra</i> ³ (39.21)	39.21			
	4	<i>Garcia nutans</i> ³ (58.13)	58.13			
	0	n.a.	0			
	1	n.a.	0			
21	2	<i>Cochlospermum vitifolium</i> ¹ (78.83), <i>Tabebuia ochracea</i> ¹ (0.35)	79.18	~97%	62.15	3.8
	3	<i>Acacia centralis</i> ¹ (3.16), <i>Guazuma ulmifolia</i> ¹ (12.72)	15.88			
	4	<i>Alophylus racemosus</i> ¹ (0.18), <i>Bonellia nervosa</i> ¹ (0.2), <i>Croton niveus</i> ¹ (0.66), <i>Lonchocarpus fasciculatus</i> ¹ (0.95), <i>Pithecellobium unguis-catii</i> ¹ (0.14), <i>Simarouba glauca</i> ³ (2.81)	4.94			
	0	n.a.	0			
	1	n.a.	0			
22	2	<i>Calycophyllum candidissimum</i> ¹ (28.17), <i>Casearia corymbosa</i> ¹ (1.17), <i>Cochlospermum vitifolium</i> ¹ (30.55),	59.89	~80%	62.71	2.3
	3	<i>Acacia centralis</i> ¹ (1.07), <i>Guazuma ulmifolia</i> ¹ (10.78)	11.85			
	4	<i>Bonellia nervosa</i> ¹ (2.65), <i>Croton niveus</i> ¹ (2.01), <i>Pithecellobium unguis-catii</i> ¹ (2.85), <i>Simarouba glauca</i> ³ (20.75)	28.26			
	0	<i>Randia aculeata</i> ³ (0.04), <i>Randia armata</i> ² (0.78)	0.82			
	1	<i>Acacia collinsii</i> ¹ (0.78), <i>Erythroxylum havanense</i> ¹ (3.58), <i>Guettarda macrocarpa</i> ¹ (18.48)	22.84			
23	2	<i>Casearia corymbosa</i> ¹ (14.67), <i>Cochlospermum vitifolium</i> ¹ (49.56), <i>Guaiacum sanctum</i> ³ (0.7)	64.93	~99%	81.07	1.9
	3	<i>Acacia centralis</i> ¹ (0.43), <i>Lonchocarpus parviflorus</i> ¹ (1.52), <i>Trichilia hirta</i> ¹ (2.79)	4.74			
	4	<i>Bonellia nervosa</i> ¹ (0.78), <i>Croton niveus</i> ¹ (2.27), <i>Lonchocarpus minimiflorus</i> ¹ (3.57), <i>Pithecellobium unguis-catii</i> ¹ (0.05)	6.67			
	0	<i>Quadrella indica</i> ³ (0.55)	0.55			
	1	<i>Acacia collinsii</i> ¹ (5.12)	5.12			
24	2	<i>Capparidastrum frondosum</i> ³ (0.06),	0.06	~99.5%	84.58	2.3
	3	<i>Guazuma ulmifolia</i> ¹ (69.76)	69.76			
	4	<i>Bonellia nervosa</i> ¹ (0.69), <i>Pithecellobium unguis-catii</i> ¹ (23.82)	24.51			
	0	<i>Quadrella indica</i> ³ (2.35), <i>Randia armata</i> ² (3.01)	5.36			
	1	n.a.	0			
25	2	<i>Calycophyllum candidissimum</i> ¹ (9.8), <i>Casearia corymbosa</i> ¹ (8.9), <i>Guaiacum sactum</i> ³ (1.46)	20.16	~66%	55.47	1.7
	3	n.a.	0			
	4	<i>Bonellia nervosa</i> ² (58.89), <i>Croton niveus</i> ¹ (13.04), <i>Lonchocarpus costaricensis</i> ¹ (2.55),	74.48			

^aPhytolith yield, ranked according to four categories (0-4) where 0 = non-productive, 1 = very low producer, 2 = low producer, 3 = moderate producer, 4 = high producer.

^bHabit, where superscripts next to species names indicate ¹ deciduous species, ² semi-deciduous species, and ³ evergreen species. Numbers in parenthesis next to species names represent the species relative biomass (%) within the quadrat.

^cBA_{sp} (%) = percent biomass represented in each quadrat by all the species in each of the yield ranking categories (0-4).

^d% non-deciduous biomass = the percentage of the total tree biomass in each quadrat that corresponds to evergreen and semi-deciduous trees.

^eFI-t ratio estimated by bootstrap analysis for each quadrat.

^fLAI values calculated from hemispheric photographs for each quadrat.