


Article

A paleothermometer for the northern Andes based on C₃–C₄ grass phytoliths

Camilla Crifo* , Juan Carlos Berrio, Arnoud Boom, Diego A. Giraldo-Cañas, and Laurent Bremond

Abstract.—Grass-dominated ecosystems cover ~40% of Earth's terrestrial surface, with tropical grasses accounting for ~20% of global net primary productivity. C₃ (cool/temperate) and C₄ (tropical and subtropical) grass distribution is driven primarily by temperature. In this work, we used phytolith assemblages collected from vegetation plots along an elevation and temperature gradient in the northern Andes (Colombia and Ecuador) to develop a paleothermometer for the region. To accomplish this, we created a transfer function based on the inverse relationship between mean annual temperature (MAT) and the phytolith-based climatic index (*Ic*), which is the proportion of C₃ over C₄ grass phytoliths (GSSCP). To evaluate how accurately the index reflects C₄–C₃ grass abundance in vegetation plots, we compared it with semiquantitative floristic estimates of C₄–C₃ grass abundance. To further evaluate the $1 - Ic$ index as a proxy for C₄–C₃ grass abundance, we compared it with corresponding $\delta^{13}C$ values (an independent proxy for C₄–C₃ vegetation). Results indicate that (1) GSSCP assemblages correctly estimate C₄–C₃ grass abundance in vegetation plots; (2) the *Ic* index outperforms the $\delta^{13}C$ record in estimating C₄–C₃ grass abundance, even in open-vegetation types; and (3) our *Ic* index–based model accurately predicts MAT. This new calibrated proxy will help improve paleotemperature reconstructions in the northern Andes since at least the emergence and spread of C₄ grasses in the region during the late Miocene.

Camilla Crifo and Laurent Bremond. Institut des Sciences de l'Évolution de Montpellier (ISEM), EPHE, PSL Research University, Université de Montpellier, CNRS, IRD, Place Eugène Bataillon, CC 065, 34095 Montpellier, France. E-mail: camilla.crifo@gmail.com, laurent.bremond@umontpellier.fr

Juan Carlos Berrio and Arnoud Boom. School of Geography, Geology and Environment, University of Leicester, University Road, Leicester LE1 7RH, U.K. E-mail: jcb34@leicester.ac.uk, ab269@leicester.ac.uk

Diego A. Giraldo-Cañas. Herbario Nacional Colombiano, Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Avenida Ciudad de Quito #55-31, Barrio Nicolás de Federmann, Bogotá D.C., Colombia. E-mail: daginaldoc@unal.edu.co

Accepted: 30 November 2022

*Corresponding author.

Introduction

Grass-dominated ecosystems occupy ~40% of Earth's terrestrial surface (Gibson 2009). Tropical grasses account for ~20% of the global net primary productivity (Lloyd and Farquhar 1994). Most of these grasses employ C₄ photosynthesis (Still et al. 2003; Edwards et al. 2010), and comprise over half of all C₄ plant species (Sage et al. 2011). C₄ photosynthesis operates through several morphological and biochemical modifications of the ancestral C₃ photosynthetic pathway resulting in reduced photorespiration and higher photosynthetic efficiency (Ehleringer and Cerling 2002). C₄ grasses tend to have an advantage and

dominate under high light conditions and in hot, seasonally arid climates such as tropical and subtropical grasslands and savannas—though they are not restricted to these habitats—and under low *p*CO₂. C₃ grasses dominate in temperate and boreal grasslands as well as in high-elevation ecosystems (Ehleringer et al. 1997; Boom et al. 2002; Edwards et al. 2010; Strömberg 2011). Thus, global patterns of C₃ and C₄ grass distribution are well known (Twiss 1992; Edwards and Still 2008; Pau et al. 2013). A number of studies have shown ecological sorting of C₃ and C₄ grasses along temperature and moisture gradients (e.g., Teeri and Stowe 1976; Chazdon 1978; Tieszen et al. 1979; Livingstone and Clayton 1980; Rundel 1980;



Young and Young 1983; Edwards and Still 2008). Large-scale distribution of C_3/C_4 grass seems to be best explained by the temperature crossover model (Ehleringer et al. 1997; Collatz et al. 1998; Still et al. 2003), which predicts that the switch from C_3 to C_4 grass dominance is a function of mean monthly air temperature at a given atmospheric CO_2 . The temperature crossover model, which is widely accepted (but see Winslow et al. 2003), assumes that variation in grass photosynthetic pathways is the main driver of grass distribution. However, several authors have pointed out the important role played by grass evolutionary history on C_3/C_4 grass distribution (e.g., Edwards and Still 2008; Pau et al. 2013). As members of the warm-climate PACMAD clade (i.e., Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, Danthonioideae), C_4 grasses are restricted to warmer areas. Thus, given the shared preference for warm temperatures of C_3 and C_4 PACMAD grasses, as opposed to the temperate-climate Pooideae clade, global patterns of C_3/C_4 grass distribution may be really patterns of Pooideae/PACMAD grass distribution (Edwards and Still 2008).

The spread of grasslands during the Neogene is thought to have had important consequences on faunas (Jacobs et al. 1999) and climate, including the global silicon cycle (Conley and Carey 2015). Geochemical, ecological, paleontological, and systematics studies investigating the drivers of C_4 grass expansion point to two main factors: water availability (aridification and/or changes in growing season precipitations) and pCO_2 levels (e.g., Cotton et al. 2016; Fox et al. 2018; Sage et al. 2018; Zhou et al. 2018). However, there is no consensus on the specific environmental dynamics that have led to the ecological success of C_4 grasses (Cotton et al. 2016). Understanding the timing and drivers of the ecological expansion of grassland ecosystems may help us better understand the factors that control both their present and future distributions. Hence, the evolution and spread of grass-dominated ecosystems during the Cenozoic has been widely investigated by paleoecologists and evolutionary biologists. Yet the grass fossil record is scant and provides limited tools to study the timing and evolution of grasses and their

response to climate change (Strömberg 2011). Fossil impressions and compressions of vegetative grass structures are rare and provide little taxonomic information (e.g., Crepet and Feldman 1991). Fossilized grass inflorescences are even rarer. Finally, grass pollen can hardly be identified beyond the family level (Palmer 1976; but see Mander et al. 2013). The $\delta^{13}C$ signature of C_3 and C_4 plants is among the most common indirect proxies used to reconstruct grass and grassland evolution. In fact, differences between C_3 and C_4 photosynthesis result in differences in carbon stable isotope fractionation (Farquhar 1983), with $\delta^{13}C$ values of C_3 plants ranging from ca. -24% to -32% , and $\delta^{13}C$ values of C_4 plants ranging from ca. -10% to -14% (Smith and Epstein 1971; Cerling et al. 1997b). The vegetation $\delta^{13}C$ signature is preserved in pedogenic carbonates (Quade et al. 1994), soil organic matter (Cerling et al. 1989; Boutton 1996), fossil grass pollen (Urban et al. 2010), and herbivore tooth enamel (Cerling et al. 1997a), as well as in lacustrine sediments (Meyers and Ishiwatari 1993) and leaf wax n -alkanes (Rieley et al. 1991), and has been used extensively to reconstruct past vegetation in the absence of direct evidence. However, in lake settings, $\delta^{13}C$ interpretations are complicated by the influence of both terrestrial and aquatic plants, as well as of microorganisms, on the carbon isotope ratio (Meyers and Ishiwatari 1993). The carbon isotope signature of leaf wax n -alkanes, which are generally abundant and well preserved in sediments, especially in lacustrine settings (Castañeda and Schouten 2011), provides finer C_3 – C_4 vegetation reconstructions (e.g., Boom et al. 2002), because n -alkanes from terrestrial plants can be differentiated from those of aquatic plants (but see Diefendorf et al. 2010). Although the $\delta^{13}C$ signature can be used to distinguish between C_3 and C_4 plants, it cannot separate C_3 grasses and trees, limiting the inferences that can be made about vegetation type and habitat openness. In fact, forest–grassland dynamics are not necessarily influenced by the same factors that control C_3 versus C_4 plant species abundance, especially in the tropics. Indeed, a number of studies have shown that open- and closed-canopy systems can persist under the same climatic conditions and that

fire is one of the main controls on the persistence of open-canopy systems (e.g., Bond et al. 2005; Favier et al. 2012; Aleman and Staver 2018).

In the last two decades, phytolith analysis has emerged as one of the most reliable proxies to reconstruct past vegetation, and in particular the evolution and spread of grasses and grass-dominated ecosystems (Thorn 2001; Strömberg 2005; Strömberg et al. 2007a; Cotton et al. 2012; Miller et al. 2012; Chen et al. 2015; Dunn et al. 2015; Strömberg and McInerney 2016; Harris et al. 2017; Loughney et al. 2019). Phytoliths are microscopic bodies of amorphous silica produced by many vascular plants (Trembath-Reichert et al. 2015; Strömberg et al. 2016) at varying abundances by absorption of monosilicic acid (hydrated silica, or opal-A) by plant roots and subsequent polymerization in plant cells and tissues (Piperno 1988). Because of their relatively inert composition, they expand the range of depositional environments that can be studied to include well-oxidized sediments where organic matter such as pollen is typically not preserved. Phytoliths primarily represent local vegetation (Piperno 1988; Crifò and Strömberg 2021) compared with pollen, which is sourced from local and regional vegetation (Prentice 1985), thus allowing finer landscape-level vegetation reconstructions. For instance, the relative proportions of phytoliths produced by grasses versus woody plants is a proxy for canopy structure (Bremond et al. 2005a; Strömberg et al. 2007b; Aleman et al. 2012). Finally, phytoliths produced by grasses in grass silica short cells (GSSCP) are highly diagnostic, allowing for the discrimination of major grass clades (subfamily level) and functional types (open- vs. closed-habitat grasses; C_3 vs. C_4 photosynthesis), thus providing direct insights into grass community composition and ecology (e.g., Alexandre et al. 1997; Barboni et al. 1999; McInerney et al. 2016). Specifically, GSSCP produced by the tropical clade PAC-MAD (i.e., Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, Danthonioideae), which predominantly uses the C_4 photosynthetic pathway, can be distinguished from those produced by C_3 grasses in the BOP clade (Bambusoideae, Oryzoideae,

Pooideae) and by C_3 early-diverging grasses (Anomochlooideae, Pharoideae, Puelioideae). Hence, phytolith analysis uses the proportion of C_4/C_3 GSSCP to estimate C_3 – C_4 grass abundance from fossil phytolith assemblages. Because C_3 – C_4 grass abundance is influenced by climate, phytolith analysis can provide new types of paleoenvironmental reconstruction and improve paleoclimatic inferences. Indeed, the proportion of C_3/C_4 GSSCP has been shown to be strongly related to temperature and consequently represents a potential paleotemperature proxy (Bremond et al. 2008b).

Here we present a grass phytolith-based paleotemperature proxy from across a Neotropical elevation gradient (~300 to 4300 m above sea level) in the northern Andes. The influence of climate and orogeny on the evolution of high-elevation, grass-dominated ecosystems in the Andes is intensely debated (e.g., Simpson and Todzia 1990; Kirschner and Hoorn 2020). The use of a phytolith-based paleotemperature proxy in parallel with traditional phytolith analysis and through comparison with other paleovegetation and paleoclimate proxies in Andean geological records could help shed light on the timing of grasses and grassland expansion into Neotropical high-elevation ecosystems and their relationship with climate variability. Several modern calibration studies have resulted in the development of phytolith-based climatic and vegetation indices in different ecosystems (Parmenter and Folger 1974; Fredlund and Tieszen 1994; Alexandre et al. 1997; Bremond et al. 2005a,b, 2008a; Aleman et al. 2012). However, many of these studies suffer from methodological limitations due to the lack of control on taphonomic processes, spanning from vegetation production to sediment deposition. In addition, only a limited number of studies couple phytolith analysis to other C_3 – C_4 plant archives for interproxy validation (e.g., Lupien et al. 2021). Furthermore, modern reference collections from plants or soil surface samples along C_3 – C_4 gradients are restricted to a few geographic areas and, to date, none have been produced in the Neotropics. Several studies have investigated the relationship between grass phytolith assemblages from soils and

grass communities in different ecosystems but have led to inconsistent results. For instance, Fredlund and Tieszen (1994) found that the GSSCP assemblage composition of soils from 15 modern grasslands in the North American Great Plains broadly reflects the regional east–west moisture gradient (with C_4 GSSCP being associated with drier sites) but also noted the extra-local and regional nature of the phytolith signal. Kerns et al. (2001) found only a weak correspondence between GSSCP assemblages and C_4 – C_3 grass species abundance in three canopy types within a *Pinus ponderosa* forest in North America. In contrast, Bremond et al. (2008b) found a positive relationship between the proportion of C_3/C_4 grasses and the proportion of C_3/C_4 GSSCP along an altitudinal gradient in Kenya. A similar trend was described by An et al. (2015) in the southern Himalayas. More recently, Biswas et al. (2021) developed a temperature calibration based on the phytolith climatic index (*Ic*; Bremond et al. 2008b) in the eastern Himalayan sectors.

The objectives of this study are (1) to assess the accuracy of grass phytolith assemblages in soils as a tool for reconstructing the C_3/C_4 grass ratio in the tropical Andes; (2) to compare this ratio with that of the $\delta^{13}\text{C}$ record; and (3) to develop a transfer function based on modern phytolith assemblages and temperature to estimate paleotemperature from fossil phytolith assemblages. To complete these objectives, floristic data, along with soil surface samples for phytolith and $\delta^{13}\text{C}$ analysis were collected and compared along an altitudinal and aridity gradient in Colombia along the northeastern Andean Cordillera. Based on this work, we propose a new phytolith-based transfer function to reconstruct paleotemperatures, and we discuss the prospects and limitations of this new proxy to improve deep-time paleoclimate inferences.

Materials and Methods

Sample and Data Collection.—A total of 35 samples for phytolith and $\delta^{13}\text{C}$ analysis were collected along two elevational transects in the eastern Andean Cordillera of Colombia (Fig. 1). These transects follow a SW-NE orientation and span from 1621 to 3545 and 331 to

2858 m above sea level (m asl), respectively. This set of samples was complemented by five additional samples spanning a higher-elevation range between 3929 and 4303 m asl, and previously collected from a transect oriented SW-SE in the Andean Cordillera of Ecuador (Ledru et al. 2013; Fig. 1). Samples from Colombia encompass several vegetation zones from dry tropical forest to lower and upper montane forest (Hooghiemstra and Van der Hammen 2004). Samples from Ecuador correspond to Paramo vegetation (Fig. 1). All samples were collected from the soil subsurface below the litter layer. Each homogenized soil sample consists of about a dozen subsamples (measuring $\sim 1\text{ cm}^3$ each) collected randomly within a plot measuring about 100 m^2 and then mixed together (Bremond et al. 2005b). Plots were installed within open-vegetation areas. Forested areas in closed- or mosaic-vegetation types were avoided, in order to minimize the influence of woody vegetation (mainly C_3) on the soil $\delta^{13}\text{C}$ signal, as well as to maximize the amount of GSSCP present in the samples. Furthermore, to limit potential biases in the composition of the GSSCP assemblages linked to the presence of exotic grass species, disturbed areas were also avoided. These mainly include pastures, where non-native grass species might have been introduced. A semiquantitative floristic survey was conducted at each sample site to estimate the relative abundance of C_3 versus C_4 grasses (see Supplementary Tables 1, 2). In each plot, two $5 \times 5\text{ m}$ subplots were installed; all grass species present in each subplot were identified and their percent cover was estimated visually. Percent cover values for each species at each site were calculated by averaging the two subplot estimations. The relative abundance of C_3 and C_4 grasses at each site corresponds to the sum of the percent cover of all C_3 and C_4 grasses, respectively (and does not take into account non-grass vegetation cover). The dominant vegetation stratum was also noted for each site. Based on this information, sampling sites were classified into six vegetation types (C_3 grassland, C_4 grassland, C_3 shrubland, C_4 shrubland, C_3 grass forest, and C_4 grass forest). The “grass forest” vegetation type refers to any closed-forest vegetation type in which grasses

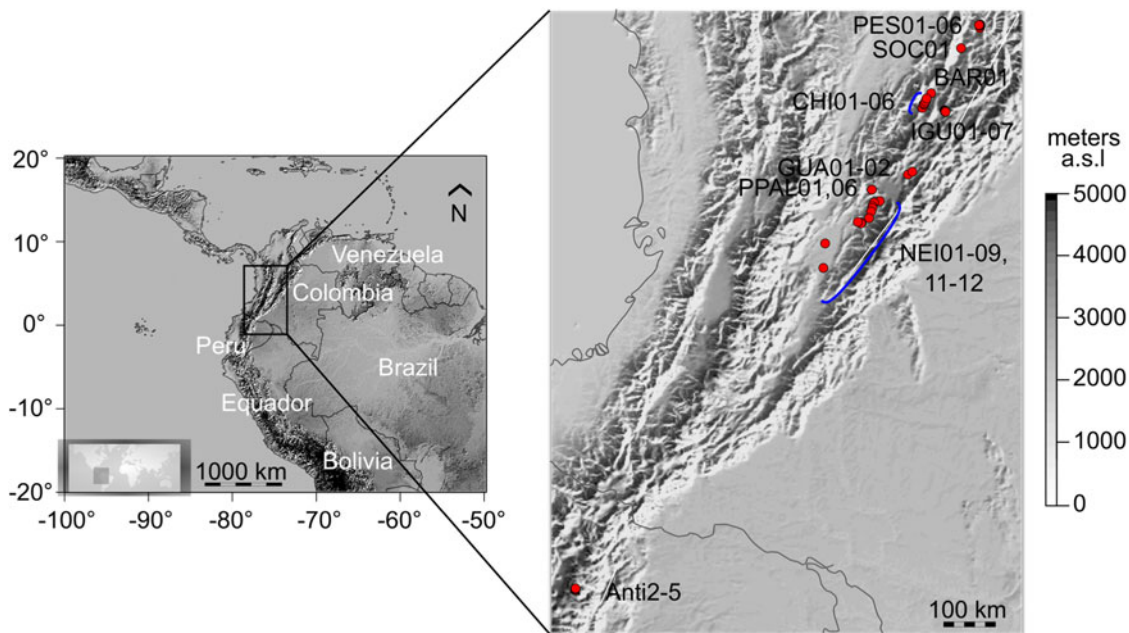


FIGURE 1. Map of the study site. Red dots represent sample sites. Site names are given in black and correspond to: PES01-06; SOC01; BAR01; CHI01-06; IGU01-07; GUA01-02; PPAL01, 06; NEI01-9, 11-2; Anti2-5. Map was created using the ETOPO2 Global Relief Model map (NOAA) in R v. 4.0.3.

are present in the understory. Geographic coordinates and elevation of each sample site were recorded in the field; mean annual precipitation (MAP) and mean annual temperature (MAT) data were gathered from the WorldClim data website (Supplementary Table 3). We extracted regional C_3 and C_4 grass distribution and MAT data from a compilation of grass species richness data from four intertropical countries in South America (Bolivia, Colombia, Ecuador, and Peru) and a compilation of climatic variables obtained from the Worldclim 1.4 1-km database by Bremond et al. (2012) (in Supplementary Table 4).

Stable Isotope Analysis ($\delta^{13}C$).—Stable carbon isotope analyses were performed on bulk soil from the 35 soil samples collected in Colombia. These analyses were not performed on the six samples collected in Ecuador due to the use of a different sampling protocol for their collection (Ledru et al. 2013). Approximately 5 g samples were pretreated with 10% hydrochloric acid (HCl) to remove carbonates and then rinsed in deionized water until neutral, freeze-dried, and ball milled. The stable isotope ratio of total organic carbon ($\delta^{13}C_{TOC}$) was

measured using a SerCon ANCA GSL elemental analyzer interfaced to a Hydra 20-20 continuous flow isotope ratio mass spectrometer. Carbon isotopic values ($\delta^{13}C$) are expressed in per mil with respect to the international VPDB (Vienna Pee Dee Belemnite) standard (see Supplementary Table 2). The precision/standard deviation is $\pm 0.16\%$.

Phytolith Extraction.—For each soil sample we used 3–5 g for phytolith extraction. We used the protocol described by Aleman et al. (2013), which includes treatment with HCl for carbonate removal, potassium hydroxide for humic acid removal, hydrogen peroxide for removal of organic material, and sodium citrate and sodium dithionite for removal of iron oxides.

We removed soil particles $>200 \mu m$ by sieving, and deflocculated the remaining soil particles using sodium hexametaphosphate. We used heavy-liquid floatation to isolate the biogenic silica (sponge spicules, diatoms, chrysophyte cysts, and phytoliths) using a heavy-liquid solution of zinc bromide, HCl, and water prepared at a specific gravity of 2.38 g/ml.

TABLE 1. List of phytolith types included in this study. Names used in column 1 come from Bremond et al. (2005a,b, 2008a, b, 2012). Full names and acronyms according to the International Code for Phytolith Nomenclature (ICPN 2.0) are given in columns 2 and 3 (numbers in parentheses in column 3 refer to morphotype subtypes). The dominant photosynthetic pathway of producing taxa (grass silica short cell phytoliths [GSSCP] only) is given in column 4. Phytolith classes (1–5) are given in column 5.

| Morphotype name (this study) | ICPN 2.0 name | ICPN 2.0 code | Dominant metabolic pathway (C ₃ /C ₄) | Class | |
|--|-------------------|---------------|--|-----------------------|------------------|
| <i>Stypha</i> -type bilobate | BILOBATE | BIL (1) | C ₃ | 1. GSSCP | |
| Short bilobate | | BIL (2) | C ₄ | | |
| Long bilobate | | BIL (3) | C ₄ | | |
| Rondel | RONDEL | RON (1) | C ₃ | | |
| Horned rondel | | RON (2) | C ₃ | | |
| Truncated rondel | | RON (3) | C ₃ | | |
| Keeled rondel | | RON (4) | C ₃ | | |
| Trapeziform trilobate | CRENATE | CRE (1) | C ₃ | | |
| Trapeziform polylobate | | CRE (2) | C ₃ | | |
| Collapsed saddle (<i>Chusquea</i> type) | SADDLE | SAD (1) | C ₃ | | |
| Short saddle | | SAD (2) | C ₄ | 2. Woody dicotyledons | |
| Long saddle | | SAD | C ₄ | | |
| Cross | CROSS | CRO | C ₄ | | |
| Globular smooth | SPHEROID PSILATE | SPH_PSI | — | | |
| Globular granulate | SPHEROID ORNATE | SPH_ORN | — | | |
| Globular echinate | SPHEROID ECHINATE | SPH_ECH | — | | |
| Cuneiform bulliform cell | BULLIFORM | BUL_FLA | — | | 4. Other Poaceae |
| | FLABELLATE | | | | |
| Parallelepipedal bulliform cell | BLOCKY | BLO | — | | 5. Nondiagnostic |
| Acicular air cell | ACUTE BULBOSUS | ACU_BUL | — | | |
| Papillae | PAPILLATE | PAP | — | | |
| Parallelepipedal long cells | ELONGATED ENTIRE | ELO_ENT | — | | |
| Thick trapezoidal rectangles | | | — | | |
| Elongates | | | — | | |
| Unclassified | Undetermined | n.a. | — | | |

Phytolith Classification and Counting.—Each dry bio-silica sample was mounted for microscopy using immersion oil to allow for rotation during identification and counting. Phytolith morphotypes were classified according to the International Code for Phytolith Nomenclature (ICPN 2.0; Neumann et al. 2019) into

23 morphotypes divided into the following classes: (1) GSSCP, exclusively produced by and diagnostic of several Poaceae (grass) sub-families; (2) phytoliths diagnostic of woody dicotyledons; (3) phytoliths distinctive of the monocot family Arecaceae (palms) (SPHEROID ECHINATE), but also possibly diagnostic of the

TABLE 2. Statistics of the linear regression between C₄/C₃ grass proportion of the samples, and MAT, elevation of the sites, C₄/C₃ GSSCP (grass silica short cell phytoliths) proportion of the samples, and the δ¹³C of the samples. **p* < 0.05; ***p* < 0.001; ****p* ≈ 0.

| Variable | C ₄ /C ₃ grasses | | | | | | <i>p</i> -value |
|--------------------------------------|--|----------|-----------------|--------------------|----------------|-------------------------|-----------------|
| | Coefficient | SE | <i>t</i> -value | Pr(> <i>t</i>) | R ² | Adjusted R ² | |
| (Intercept) | −29.0226 | 10.2961 | −2.819 | 0.00753** | | | |
| MAT | 5.2404 | 0.5773 | 9.078 | 3.68e−11*** | 0.6788 | 0.6705 | 3.676e−11 |
| (Intercept) | 131.033145 | 8.672821 | 15.108 | <2e−16*** | | | |
| Elevation | −0.033227 | 0.003496 | −9.503 | 1.06e−11*** | 0.6984 | 0.6907 | 1.062e−11 |
| (Intercept) | 5.663 | 6.458 | 0.877 | 0.386 | | | |
| C ₄ /C ₃ GSSCP | 118.327 | 12.049 | 9.820 | 4.27e−12*** | 0.712 | 0.7047 | 4.266e−12 |
| (Intercept) | 213.539 | 48.962 | 4.365 | 0.000113*** | | | |
| δ ¹³ C | 6.684 | 2.186 | 3.058 | 0.004324** | 0.2157 | 0.1926 | 0.004324 |

family Bromeliaceae (bromeliads) and of the order Zingiberales (e.g., spiral gingers) (i.e., when the phytolith diameter is $\leq 5\mu\text{m}$; see Benvenuto et al. 2015; Crifò and Strömberg 2021); (4) other phytoliths diagnostic of the Poaceae family (including forms produced mainly but not exclusively by the Poaceae family), as well as other morphotypes belonging to the family Poaceae but whose production is influenced by environmental factors such as water availability (e.g., Bremond et al. 2005b; Madella et al. 2009); and (5) other nondiagnostic forms. Morphotype names (according to Bremond et al. 2005a), corresponding ICPN names and codes, and other information are provided in Table 1. Details regarding GSSCP morphotype assignment to C_3 or C_4 grasses are given in the Supplementary Appendix. Morphotypes used in our classification are illustrated in Figure 2. Each phytolith count includes all morphotypes present in the assemblage from the five classes mentioned above, and a minimum number (n) of 200 GSSCP diagnostic morphotypes (except for one sample with $n = 139$) for statistical significance (Pearson 2000; Strömberg 2009). Phytolith counts are provided in Supplementary Table 2.

Statistical Analysis.—All statistical analyses were performed in R v. 4.0.3 (R Development Core Team 2019). All R codes are provided in the Supplementary Appendix.

The C_3/C_4 GSSCP ratio, corresponding to the climate index Ic (Bremond et al. 2008b), was calculated as follows:

$$Ic = \frac{\text{Stipa type bilobate} + \text{trapeziform trilobate} + \text{trapeziform polylobate} + \text{collapsed saddle} + \text{rondel}}{\text{short bilobate} + \text{long bilobate} + \text{short saddle} + \text{long saddle} + \text{cross}} \quad (1)$$

To assess robustness of the calculated Ic values, 95% confidence intervals (CIs) were constructed using bootstrap analysis of 10,000 replicates (Strömberg 2009). Linear regression was used to test the significance of the relationship between the proportion of C_4 grass phytoliths over the total GSSCP count (i.e., $1 - Ic$) and the proportion of C_4 grasses at our sites. We used a generalized linear model (R function *glm*) with a logit-link function to build a transfer function for MAT. The logit-link function converts the phytolith C_4 proportion Y to the

logarithm of the ratio $\frac{Y}{(1-Y)}$ in order to make the data suitable for regression. Thus, we fit the following equation to our data:

$$\log\left(\frac{Y_i}{1 - Y_i}\right) = \beta_0 + \beta_1 T + w \quad (2)$$

where β values correspond to the intercepts, and T corresponds to MAT. To account for potential error due to differences in sample size, we assigned a weight (w) to samples according to their CIs so that phytolith assemblages with wider CIs have smaller weight in the regression fit than assemblages with narrower CIs (see Harris et al. 2017). We calculated w as follows:

$$w = \frac{1}{((Y_i^{\text{upper}} - Y_i) + 0.1)} \quad (3)$$

where Y_i corresponds to $1 - Ic$ and Y_i^{upper} is the upper bound (upper CI) of Y_i . To deal with cases where $Y_i = 0$, which lead to $Y_i^{\text{upper}} = Y_i$, and $w = 0$, we added an additional small weight (0.1) into the denominator of the equation.

Differences in GSSCP compositions between sites were evaluated using nonmetric multidimensional scaling (NMDS; R function *metaMDS*) on the GSSCP morphotype relative abundance matrix (see Crifò and Strömberg 2021). We used tests of stress value randomization (R function *nmds.monte*) to determine the

minimum number of dimensions required to capture the variation in the data. Further, we used correspondence analysis (CA; R function *cca*) to evaluate the influence of environmental gradients (i.e., MAT, MAP, and elevation) on GSSCP compositional differences between sites. CA was carried out using the GSSCP morphotype relative abundance data and the log-transformed environmental variables (elevation, MAT, and MAP). For visual comparison of the stable carbon isotope and phytolith ($1 - Ic$) proxies, $\delta^{13}\text{C}$ values obtained

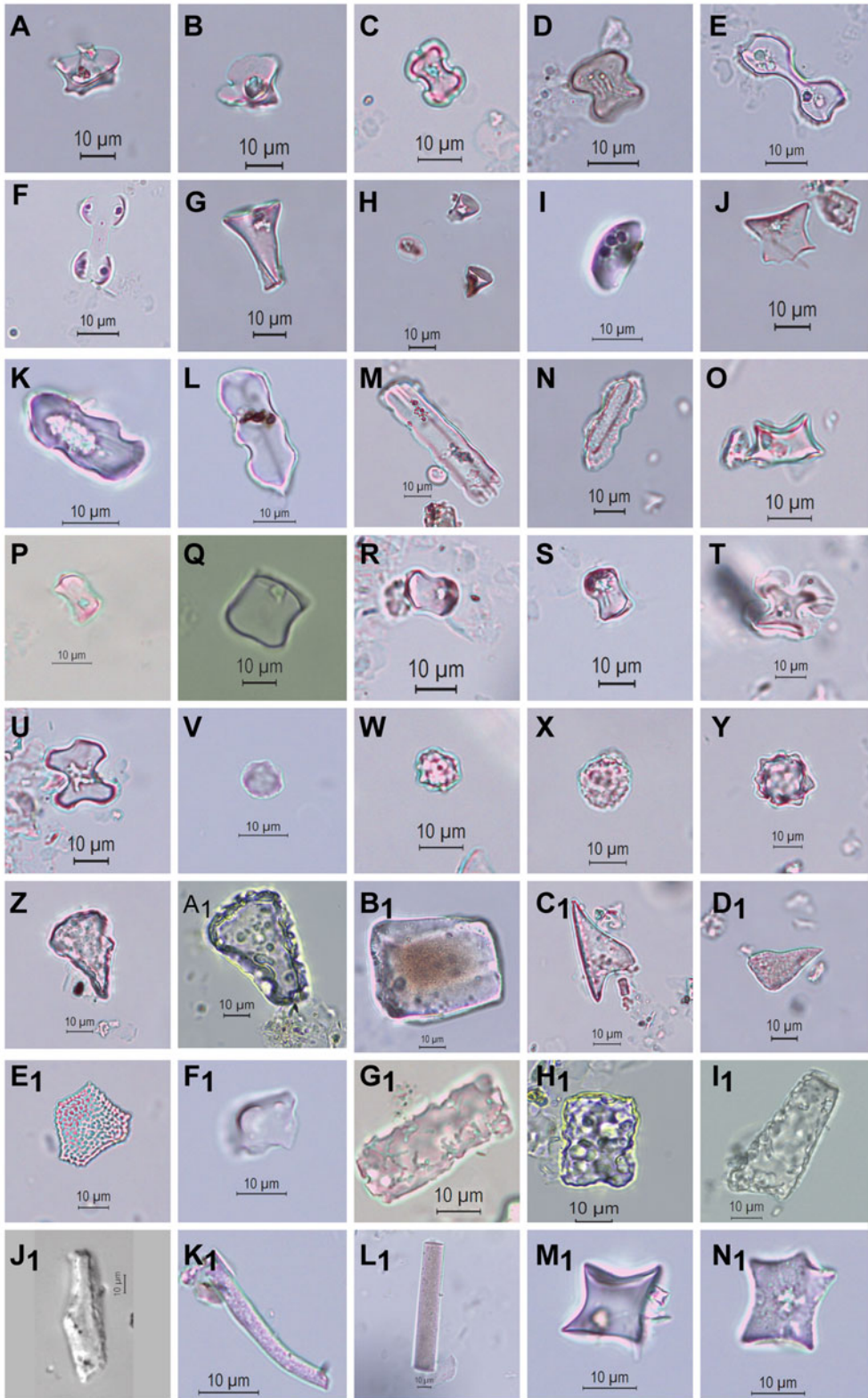


FIGURE 2. Examples of phytolith morphotypes from soil assemblages diagnostic of: Poaceae (A–U), woody dicotyledons (V–X), palms/bromeliads/Zingiberales (Y), other Poaceae (Z–F₁), nondiagnostic morphotypes (G₁–L₁), and morphotypes of undetermined affinity (M₁–N₁). Morphotype names in small capitals (e.g., BILOBATE), and acronyms (e.g., BIL) signify morphotypes as formally defined in the International Code for Phytolith Nomenclature 2.0 (Neumann et al. 2019); morphotype names in square brackets correspond to the classification scheme by Bremond et al. (2005a,b, 2008a,b, 2012). A–F, BILOBATE (BIL) morphotypes, including: A, B, BIL (1) [*Stipa* type bilobate]; C,D, BIL (2) [short bilobate]; and E,F, BIL (3) [long bilobate]. G–J, RONDEL (RON) morphotypes. K–N, CRENATE (CRE) morphotypes, including K, L, CRE (1) [trapeziform trilobate]; and M, N, CRE (2) [trapeziform polylobate]. O–S, SADDLE (SAD) morphotypes, including: O, P, SAD (1) [collapsed saddles]; Q, R, SAD (2) [short saddle]; and S, SAD (3) [long saddle]. T, U, CROSS morphotypes. V, SPHEROID PSILATE (SPH_PSI) morphotypes [globular smooth]. W, X, SPHEROID ORNATE (SPH_ORN) morphotypes [globular granulate]. Y, SPHEROID ECHINATE (SPH_ECH) morphotypes [globular echinate]. Z, A₁, BULLIFORM FLABELLATE (BUL_FLA) morphotypes [cuneiform bulliform cell]. B₁, BLOCKY (BLO) morphotypes [parallelepipedal bulliform cell]. C₁, D₁, ACUTE BULBOSUS (ACU_BUL) morphotypes [acicular air cell]. E₁, F₁, PAPILLATE (PAP) morphotypes [papillae]. G₁–L₁, ELONGATE ENTIRE (ELO_ENT) morphotypes, including: G₁, H₁, ELO_ENT (1) [parallelepipedal long cell]; I₁, J₁, ELO_ENT (2) [thick trapezoidal rectangle]; and K₁, L₁, ELO_ENT (3) [elongates]. M₁, N₁, Undetermined grass silica short cell phytoliths (GSSCP) morphotypes (n.a.) [unclassified].

from the samples were plotted together with C₄ proportions against MAT values. We used linear regression to test for a relationship between the proportion of C₄ GSSCP and δ¹³C values of the samples and between the proportion of C₄ GSSCP in the samples and the proportion of C₄ grasses recorded by the floristic survey.

Results

C₄ Grass Proportion from Floristic Survey and Phytolith Counts along Elevation and MAT Gradients.—The proportion of C₄ grasses shows a positive relationship with MAT and a negative relationship with elevation ($p < 0.001$; $R^2 = 0.68$ and 0.7 , respectively; Table 2). Linear regression analysis also indicates a positive relationship between the C₄ grass proportion from the floristic survey and the C₄ grass proportion estimated from phytolith assemblages (i.e., $1 - I_c$) ($p < 0.001$; $R^2 = 0.71$; Table 2). In addition, the C₄/C₃ GSSCP proportion is positively related to MAT and negatively related to elevation ($p < 0.001$; $R^2 = 0.88$ for both regressions; Table 3).

Influence of Environmental Gradients on the Distribution of GSSCP Morphotypes across Sites.—NMDS ordination of the sites performed on GSSCP assemblage composition adequately represents the data using two dimensions (NMDS, stress value = 0.14; test of stress value randomization, $p < 0.01$). Phytolith morphotype counts are provided in Supplementary Table 2. On NMDS axis 1 (Fig. 3), positive loadings are represented by “regular” (RON_1) “truncated” (RON_3), and “keeled” (RON_4) RONDEL morphotypes ($p < 0.001$), “trapeziform polylobate” (CRE_2) CRENATE morphotypes ($p < 0.001$), and by “*Stipa*-type” (BIL_1) BILOBATE morphotypes ($p < 0.05$); while negative loadings are represented by “short” (BIL_2) and “long” (BIL_3) BILOBATE and CROSS (CRO) morphotypes ($p < 0.001$). On NMDS axis 2, positive loadings are represented by the “short” (SAD_2) and “long” (SAD_3) SADDLE morphotypes ($p < 0.001$); while negative loadings are represented by “trapeziform trilobate” (CRE_1) CRENATE morphotypes ($p < 0.001$), “short” (BIL_2) and “long” (BIL_3) BILOBATE morphotypes ($p < 0.001$), and CROSS (CRO)

TABLE 3. Statistics of the linear regression between C₄/C₃ grass silica short cell phytoliths (GSSCP) proportion of the samples, and MAT, elevation of the sites, and the δ¹³C of the samples. * $p < 0.05$; ** $p < 0.001$; *** $p \approx 0$.

| C ₄ /C ₃ GSSCP | | | | | | | |
|--------------------------------------|-------------|-----------|-----------------|-------------------|----------------|-------------------------|-----------------|
| Variable | Coefficient | SE | <i>t</i> -value | Pr(> <i>t</i>) | R ² | Adjusted R ² | <i>p</i> -value |
| (Intercept) | −0.263601 | 0.045328 | −5.815 | 9.35e−07*** | | | |
| MAT | 0.042492 | 0.002541 | 16.720 | 2e−16*** | 0.8776 | 0.8744 | 2.2e−16 |
| (Intercept) | 1.025e+00 | 3.973e−02 | 25.80 | 2e−16*** | | | |
| Elevation | −2.653e−04 | 1.602e−05 | −16.56 | 2e−16*** | 0.8755 | 0.8724 | 2.2e−16 |
| (Intercept) | 0.83763 | 0.38503 | 2.175 | 0.0366* | | | |
| δ ¹³ C | 0.01547 | 0.01720 | 0.899 | 0.3747 | 0.2324 | −0.005 | 0.3747 |

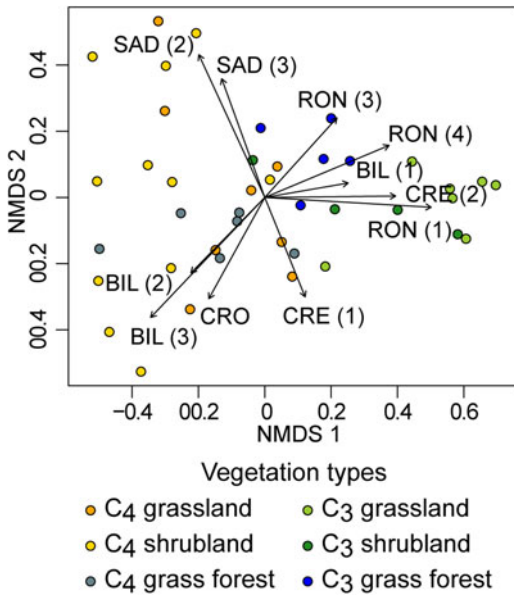


FIGURE 3. Nonmetric multidimensional scaling (NMDS) ordination of the phytolith assemblages across sites (colored dots), and phytolith classes (black vectors). Only phytolith classes with significant ($p < 0.05$) loadings on the ordination axes are shown here. Vector direction indicates maximum correlation of the phytolith classes with between-sample scores in the ordination space. Sample sites are colored according to vegetation type (see figure key).

morphotypes ($p < 0.001$). Sites dominated by C_3 grasses tend to group toward the positive end of NMDS axis 1 as opposed to C_4 grass-dominated sites, which are grouped toward the negative end. Among C_4 grass-dominated sites, C_4 grasslands and C_4 shrublands are spread along NMDS axis 2, while C_4 grass forests tend to be grouped toward the negative side of the axis.

CA ordination of the sites performed on the site GSSCP assemblage composition (counts) by environmental variables (MAP, MAT, and elevation) shows a similar pattern of site distribution (Fig. 4). The first two axes of the CA ordination account for respectively 43% and 15% of the total variation. CA axis 1 describes a negative MAT gradient ($p < 0.001$) opposite to a positive elevation gradient ($p < 0.01$). C_4 -dominated sites are distributed toward the negative end of this axis, as opposed to C_3 -dominated sites, which are distributed toward the positive end. Most C_4 -dominated sites tend to group toward the positive end of

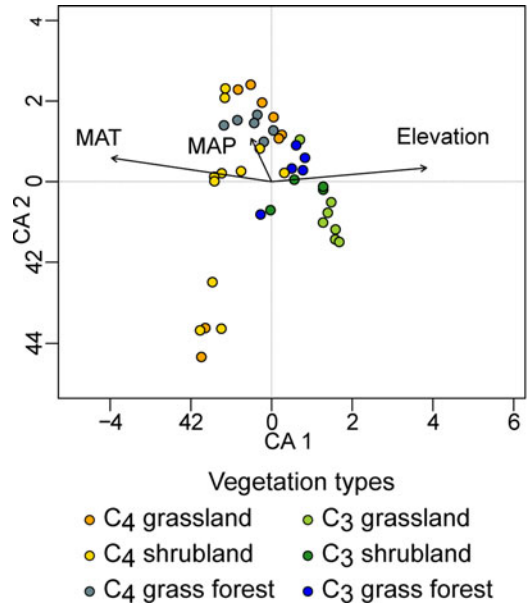


FIGURE 4. Correspondence analysis (CA) ordination of the phytolith assemblages across sites (colored dots), and environmental variable (black vectors). Sample sites are colored according to vegetation type (see figure key).

CA axis 2, with the exception of five sites that appear to form an isolated cluster toward the negative end of CA axis 2. The positive loading of MAP on this axis is not significant ($p > 0.05$).

Predicted Temperature as a Function of C_4 Phytolith and C_4 Grass Proportion.—Logistic regression shows that the C_4 GSSCP proportion is a good predictor of MAT. C_4 phytolith and grass proportions increase with MAT following a sigmoidal growth pattern ($p < 0.001$; Fig. 5). Accordingly, at lower temperatures (i.e., MAT $< 12^\circ\text{C}$) and high elevation, the C_4 GSSCP proportion is low, and close to zero in several assemblages (Fig. 6). These assemblages are dominated by RONDEL, “*Stipa*-type” BILOBATE, and CRENATE phytolith morphotypes (Fig. 3), which are assigned to grass subfamilies in the C_3 BEP clade (Bambusoideae, Ehrhartoideae, and Pooideae). The proportion of C_4 phytoliths sharply increases at higher temperatures. SADDLE, CROSS, and “short” and “long” BILOBATE phytolith morphotypes, assigned to the C_4 clade PACMAD, dominate these assemblages (Figs. 3, 6). However, unlike C_4 GSSCP, C_3 GSSCP are present across the entire temperature (and elevation) range.

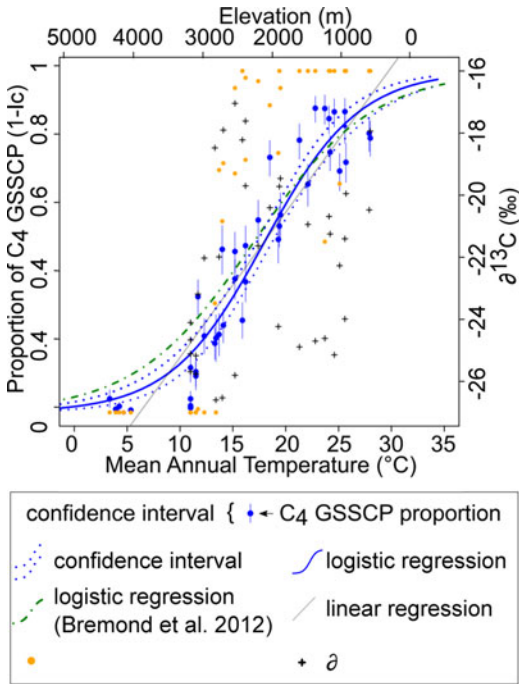


FIGURE 5. Proportion of C_4 grass phytoliths (GSSCP) and carbon isotope ratios ($\delta^{13}C$) of samples against mean annual temperature (MAT; lower x-axis) and mean annual precipitation (MAP; upper x-axis). Blue dots represent estimates of C_4 GSSCP proportion from phytolith counts; vertical blue bars correspond to 95% confidence intervals (CIs) calculated by bootstrap analysis of each sample. Black crosses represent carbon isotope ratios ($\delta^{13}C$) of the samples. Solid and dotted blue lines represent, respectively, the logistic regression curve performed between the proportion of C_4 GSSCP in phytolith assemblages and MAT ($p \ll 0.01$), and the upper and lower CI bounds. The solid grey line represents the linear regression ($p \ll 0.001$, $R^2 = 0.87$) performed between the proportion of C_4 GSSCP in phytolith assemblages and MAT. Orange dots represent the C_4 proportion in the study sites (calculated from the floristic survey). The dot-dashed green line represents the logistic regression based on a compilation of floristic data for Colombia, Ecuador, Peru, and Bolivia in Bremond et al. (2012) ($p \ll 0.01$).

Relationship between C_4 GSSCP Proportion and Carbon Isotope Ratio.—The $\delta^{13}C$ values of the samples span -17.04‰ to -26.59‰ (Fig. 5B), corresponding mainly to intermediate C_3 – C_4 , and C_3 vegetation signatures (Cerling et al. 1997b). Linear regression indicates no significant relationship between the carbon $\delta^{13}C$ values and the proportion of C_4 GSSCP in the samples ($p \gg 0.05$; Table 3). Likewise, sample $\delta^{13}C$ values have no significant relationship with environmental factors (elevation, MAT, and MAP) ($p \gg 0.05$; Supplementary Table 5)

but are weakly related to the C_4 grass proportion of the sites ($p < 0.01$; $R^2 = 0.21$; Table 2).

Discussion

Relationship between the Ic Index and C_3 – C_4 Grass Distribution.—Overall, the grass floristic survey and phytolith counts reflect C_3 – C_4 grass distribution along elevation and temperature gradients, as previously documented in the region (Bremond et al. 2012). However, MAT and elevation explain most of the variation in 1-*Ic* phytolith index but explain variations in C_4/C_3 grass proportions from the floristic surveys only to a lesser extent. Further, at least some of the variation in 1-*Ic* seems to not be explained by variation in C_4/C_3 grass proportions from the floristic surveys. Floristic data indicate higher C_4/C_3 grass proportions at high temperatures and lower C_4/C_3 grass proportions at lower temperatures compared with the phytolith data. These inconsistencies might have several explanations. First, the floristic survey provides instantaneous estimates of C_4/C_3 grass proportion on a small spatial scale. The C_4/C_3 grass proportion can vary in relation to the botanical survey season and to local conditions, such as microclimate and topography, and in response to short-term climate fluctuations. Thus, it might not be representative of the vegetation on a wide spatial and temporal scale; instead, it might reflect super-local vegetation and environmental conditions. In contrast, C_4/C_3 phytolith abundance from soil surface samples does not represent a real-time snapshot of the vegetation, because a phytolith assemblage typically reflect tens to hundreds of years of vegetation occurrence (Song et al. 2016). Further, our floristic survey was conducted in a semiquantitative way and only provides estimates of relative C_4/C_3 grass cover. Error associated with these estimates could explain at least in part the mismatch between the floristic survey data and the phytolith signal. Furthermore, although phytolith production should be similar among C_3 and C_4 grasses growing in the same habitat and under identical environmental conditions (Brightly et al. 2020), the relative proportion of different diagnostic phytolith morphotypes produced for a given plant biomass varies

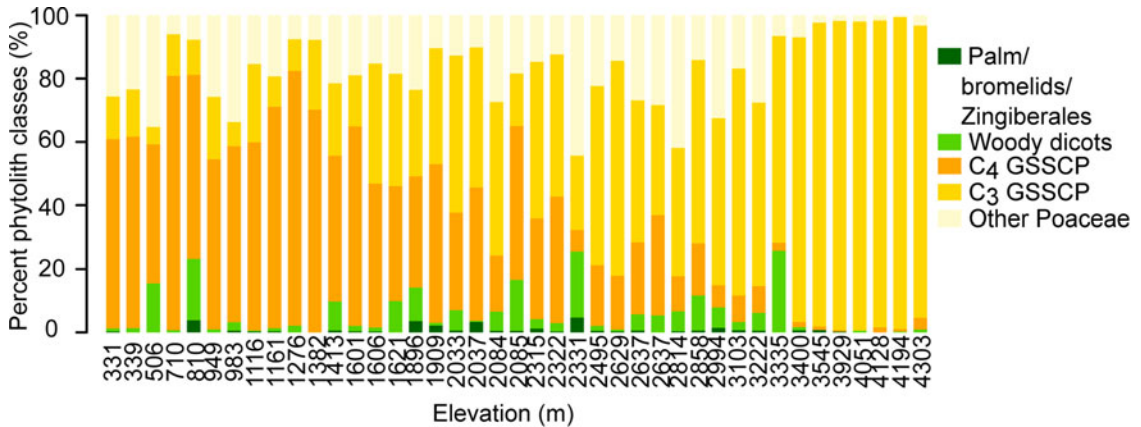


FIGURE 6. Phytolith assemblages organized according to increasing elevation. Bars represent percentage of phytolith classes. Palm phytoliths include SPHEROID ECHINATE morphotypes; woody dicot phytoliths include SPHEROID PSILATE and SPHEROID ORNATE morphotypes; C_3 grass silica short cell phytoliths (GSSCP) include “*Stipa*-type” BILOBATE [BIL (1)], RONDEL (RON), CRENATE (CRE), and “collapsed” SADDLE [SAD (1)] morphotypes; C_4 GSSCP include “short” and “long” BILOBATE [BIL (2) and BIL (3)], short” and “long” SADDLE [SAD (2), and SAD (3)], and CROSS (CRO) morphotypes; other Poaceae morphotypes include BULLIFORM FLABELLATE (BUL_FL_A), BLOCKY (BLO), ACUTE BULBOSUS (ACU_BUL), and PAPILLATE (PAP) morphotypes.

among species and might result in C_3/C_4 proportion under- or overestimation. In addition, coarse classification of RONDEL morphotypes (which are all grouped at the numerator of the I_c equation, as diagnostic of C_3 grasses) might result in potential underestimation of C_4 grass abundance (which also produces these morphotypes). However, the error associated to potentially erroneous RONDEL classification is unknown, because data on the production of different morphotypes by different species are not available (see also Supplementary Appendix). Finally, the influence of taphonomic biases on C_3/C_4 GSSCP proportion can be reasonably excluded, because the phytolith assemblages extracted from the samples are all finely preserved and do not present evidence of differential preservation or dissolution.

The C_4/C_3 GSSCP Proportion as a Paleotemperature Proxy.—Our model, based on logistic regression, performs better than models based on linear regression (i.e., Biswas et al. 2021) and is in line with the temperature crossover model (Ehleringer et al. 1997; Collatz et al. 1998; Still et al. 2003) and with previous work (Bremond et al. 2012). Indeed, the relationship between C_4/C_3 GSSCP proportion and temperature is not linear, due to the combined effect of the presence of C_3 grasses across the entire elevation and temperature range and of

a crossover temperature (at which the proportion of C_4 grasses, and thus of C_4 GSSCP, equals that of C_3 grasses, and thus of C_3 GSSCP). This results in a sigmoidal “growth” pattern of the C_4/C_3 GSSCP proportion that is best described by logistic regression. Thus, the transfer function developed herein can be used to obtain paleotemperature estimates in the fossil record of the Andean region. Nevertheless, the accuracy of our model also relies on the fact that phytolith assemblages constitute a local vegetation signal and on our choice of restricting soil sampling to open areas. In fact, in mountain forests, grass species belonging to the subfamily Panicoideae (which has both C_3 and C_4 representatives) and inhabiting forest understory often use C_3 photosynthesis. At intermediate elevation (where mountain forest occurs), their presence can therefore bias the C_4/C_3 proportion toward higher values. While our sampling strategy enabled us to limit such bias associated to the presence of C_3 grasses of the subfamily Panicoideae (by sampling preferentially in open-vegetation areas), it might complicate the interpretation and subsequent use of the transfer function to estimate MAT (which might be potentially overestimated) from fossil phytolith assemblages, where the vegetation type is unknown. As a consequence, for correct application of this

method for paleotemperature reconstructions, fossil phytolith assemblages should first be studied in detail (including forest indicator phytoliths) in order to preferentially select those representing open-canopy vegetation. Phytolith-based vegetation indices commonly used to assess canopy cover, such as the *D/P* index (the sum of SPHEROID ORNATE phytoliths, produced by many tropical ligneous dicotyledonous, over the sum of Poaceae phytoliths; Alexandre et al. 1997) and the *FI-t* ratio (the sum of forest indicator phytoliths morphotypes over the sum of all morphotypes; Strömberg et al. 2007b), should be used.

Relationship between the C_4/C_3 GSSCP Proportion and the Carbon Isotope Ratio ($\delta^{13}C$).—Soil carbon isotopic ratios ($\delta^{13}C$), are not consistent with the *Ic* index and the C_3 – C_4 grass composition from floristic surveys along elevation and temperature gradients. In fact, most $\delta^{13}C$ values indicate C_3 or mixed C_3 – C_4 vegetation, while none of the samples is characterized by $\delta^{13}C$ values indicative of C_4 -dominated vegetation. This discrepancy can be explained by several factors. First, the soil $\delta^{13}C$ signature represents a mixed vegetation signal (which reflects the presence of grasses as well as forbs, shrubs, and trees) rather than a pure grass signal. Because the relationship between the prevalence of C_3 versus C_4 photosynthesis and temperature differs in grasses and trees, a certain degree of inconsistency between C_3 – C_4 grass abundance from floristic surveys and soil $\delta^{13}C$ values, as well as between temperature and $\delta^{13}C$ values, is expected. However, as mentioned earlier, in order to obtain a $\delta^{13}C$ signal representative mostly of the grass component of the vegetation, sample collection was deliberately conducted in open-canopy areas only. Thus, other factors must explain the poor match between $\delta^{13}C$, C_4/C_3 GSSCP proportion, and grass community composition. C_3 and C_4 plants display $\delta^{13}C$ spectra that are quite wide and result in a bimodal distribution of $\delta^{13}C$ values (Cerling et al. 1997b), which renders interpretations of C_4 grass abundance based on $\delta^{13}C$ values particularly complex. Additionally, time averaging of the soil might result in $\delta^{13}C$ values representing a mix of past and present vegetation, and less negative $\delta^{13}C$ values might record past vegetation shifts between

closed and open canopy in sites that today are dominated by C_4 grasses. However, biases related to time averaging also would have affected the phytolith record (e.g., Harris et al. 2017). We found no evidence for such bias in the phytolith assemblages. Nonetheless, it is possible that the shorter turnover time of phytoliths compared with the $\delta^{13}C$ counteracts the effect of time averaging on phytolith assemblages but not on the carbon isotopic ratio of the samples (Aleman et al. 2012). Pedogenic and diagenetic alteration could also bias $\delta^{13}C$ values (e.g., Chen et al. 2015) but should also result in phytolith preservation biases. In light of the good preservation of the phytolith samples, pedogenic and diagenetic alteration can be excluded. Finally, differences between the C_3 – C_4 grass signal of $\delta^{13}C$ and phytoliths may also be due to differential degradation of organic matter. Higher decomposition rates have been observed for C_4 organic material compared with C_3 organic material in mixed C_3 – C_4 soils (Wynn and Bird 2007) and could explain less negative $\delta^{13}C$ ratios, at least in mixed C_3 – C_4 vegetation areas (Cotton et al. 2012).

Applications to the Deep-Time Fossil Record.—The application of the new proxy proposed herein is of particular interest in tropical montane, high-elevation ecosystems where available paleotemperature proxies tend to be scarce—at lower elevations, the *Ic* index is less accurate due to the abundance of several C_3 tropical grasses in closed-canopy tropical ecosystems. The use of the *Ic* index to reconstruct paleotemperature is independent of whether the paleoelevation of a fossil site is known. However, paleoelevation data might improve paleoclimatic interpretations based on the *Ic* index. For instance, a change in temperature in the absence of a shift in elevation might be the result of global climate change as opposed to local temperature decrease, which might be driven by mountain building. In the Andean region, the application of the *Ic* index to the deep-time phytolith record has the potential to advance our understanding of the expansion of grasses and grass-dominated ecosystems across the region. However, there are some caveats. First, the application of our model to the deep-time record assumes that the relationship between C_3 – C_4 grass distribution and

MAT observed today also occurred in the past. Indeed, other factors, such as aridity and atmospheric $p\text{CO}_2$, are known to influence C_3 – C_4 grass distribution. While increasing evidence points to relatively stable $p\text{CO}_2$ level during at least the last 7 Ma, and possibly since 23 Ma (Cui et al. 2020), the uplift of the Andes is thought to have driven a progressive shift toward aridification, the timing and magnitude of which are still debated (Martínez et al. 2020). Caution is therefore required, and phytolith-based inferences should be coupled with other proxies. Second, the temporal range to which the proposed proxy can be applied is dependent on the timing of the expansion of C_4 grasses and C_4 -dominated habitats. Although the use of the C_4 photosynthetic pathway developed at least since the Oligocene (Edwards et al. 2010) or even earlier during the Paleocene (Gallaher et al. 2020), the expansion of C_4 grasses has not been documented until the late Miocene, starting at around 8 Ma (Cerling et al. 1997b; Strömberg 2011). In South America, the grass fossil record is scarce, and most of our knowledge on the expansion of grasslands comes from Patagonia (Strömberg 2011), where C_3 grasses remained dominant at least until the early Miocene (ca. 20 Ma), whereas grassland expansion did not take place until the early-late Miocene (ca. 10 Ma) (Barreda and Palazzesi 2007; Strömberg et al. 2013; Dunn et al. 2015). In northern South America, it has been hypothesized that the retreat of the Pebas system and the establishment of the Amazon drainage basin around 9 Ma (Hoorn et al. 2010) favored the expansion of grasses in the newly established ecosystems of the Andean slopes and floodplains. The expansion of grasses in Andean ecosystems would have been further favored by Andean orogenesis and subsequent climate change (Kirschner and Hoorn 2020) creating new ecological niches. By the Plio-Pleistocene, the Andes had reached near-modern elevations (Gregory-Wodzicki 2000; Garziona et al. 2017), and the first Paramo and Puna-like ecosystems are reported around 5 Ma (Bermúdez et al. 2017; Martínez et al. 2020). No data recording the occurrence of C_4 grasses in these ecosystems are currently available, although work is in progress in the Bolivian

and Peruvian Altiplano (e.g., B. Saylor personal communication 2021; C.C. personal observations). Thus, we argue that phytolith based paleotemperature reconstructions in the region might be reasonably made onwards from at least the late-Miocene-early Pliocene. Finally, local calibrations of the *Ic* index, may allow its application to montane regions across a wider geographic range.

Conclusion

The relationship between C_3 – C_4 grass distribution and climate has long been studied. Likewise, in the last two decades, the notion that phytoliths provide accurate reconstruction of grass community composition has been widely recognized. However, only a handful of studies have focused on grass phytoliths as a tool for reconstructing paleoclimate quantitatively. Our study is the first to use the proportion of C_4/C_3 GSSCP in soil samples collected along an elevation gradient to develop a paleothermometer in the Andean region. Our data also support the notion that grass phytoliths represent grass community composition, and in particular the proportion of C_3 versus C_4 grasses, more accurately than indirect proxies such as the carbon isotope ratio of soil organic matter. Moreover, the use of phytoliths presents several advantages over other direct and indirect grass vegetation proxies, which are either rarely available in the fossil record (e.g., grass macrofossils, pollen grains, or leaf cuticles) or subject to many biases (e.g., leaf wax *n*-alkanes, $\delta^{13}\text{C}$ of soil carbonates, and herbivore tooth enamel).

Applications of this new tool include paleoclimatic and paleoenvironmental reconstructions in the Andean region from at least the late Miocene–early Pliocene on. Coupling this new method with the analysis of whole phytolith assemblages and with other proxies will improve paleoclimatic and paleoenvironmental reconstructions in the Andes, potentially contributing a better understanding on the complex dynamics involving tectonics, climate, and vegetation that shaped most modern Andean ecosystems.

Based on the strong, inverse relationship between temperature and elevation, future

work should focus on the development of the *Ic* index as a tool for paleoaltimetry. However, this requires accurate calibrations. The link between C_3/C_4 grass distribution and elevation is only indirect and relies on the relationship between temperature and elevation. Thus, changes in regional and temporal factors that might influence temperature must be taken into account. For instance, calibrations of the *Ic* index as a paleoaltimetry proxy should include modeling C_3/C_4 grass distribution along temperature and elevation gradients under varying pCO_2 .

Acknowledgments

We thank M.-P. Ledru for providing Paramo samples from Ecuador. We also want to thank C. L. Yost and two other anonymous reviewers for their thorough reviews, which contributed to improve the quality of this article. This work was initiated during the Intra-European Marie Curie Fellowships of L.B. (MEIF-CT-2005-024625).

Data Availability Statement

Data available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.5h5hqbzksk>.

Literature Cited

- Aleman, J. C., and A. C. Staver. 2018. Spatial patterns in the global distributions of savanna and forest. *Global Ecology and Biogeography* 27:792–803.
- Aleman, J. C., B. Leys, R. Apema, I. Bentaleb, M. A. Dubois, B. Lamba, J. Lebamba, C. Martin, A. Ngomanda, L. Truc, J.-M. Yangakola, C. Favier, L. Bremond, and K. Woods. 2012. Reconstructing savanna tree cover from pollen, phytoliths and stable carbon isotopes. *Journal of Vegetation Science* 23: 187–197.
- Aleman, J. C., A. Saint-Jean, B. Leys, C. Carcaillet, C. Favier, and L. Bremond. 2013. Estimating phytolith influx in lake sediments. *Quaternary Research* 80:341–347.
- Alexandre, A., J.-D. Meunier, A.-M. Lezine, A. Vincens, and D. Schwartz. 1997. Phytoliths: indicators of grassland dynamics during the late Holocene in intertropical Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology* 136:213–229.
- An, X., H. Lu, and G. Chu. 2015. Surface soil phytoliths as vegetation and altitude indicators: a study from the southern Himalaya. *Scientific Reports* 5:15523.
- Barboni, D., R. Bonnefille, A. Alexandre, and J.-D. Meunier. 1999. Phytoliths as paleoenvironmental indicators, West Side Middle Awash Valley, Ethiopia. *Palaeogeography Palaeoclimatology Palaeoecology* 152:87–100.
- Barreda, V., and L. Palazzesi. 2007. Patagonian vegetation turnovers during the Paleogene–early Neogene: origin of arid-adapted floras. *Botanical Review* 73:31–50.
- Benvenuto, M. L., M. Fernández Honaine, M. L. Osterrieth, and E. Morel. 2015. Differentiation of globular phytoliths in *Arecaceae* and other monocotyledons: morphological description for paleobotanical application. *Turkish Journal of Botany* 39:341–353.
- Bermúdez, M. A., C. Hoorn, M. Bernet, E. Carrillo, P. A. van der Beek, J. I. Garver, J. L. Mora, and K. Mehrkian. 2017. The detrital record of late-Miocene to Pliocene surface uplift and exhumation of the Venezuelan Andes in the Maracaibo and Barinas foreland basins. *Basin Research* 29(S1):370–395.
- Biswas, O., R. Ghosh, S. Agrawal, P. Morthekai, D. K. Paruya, B. Mukherjee, M. Bera, and S. Bera. 2021. A comprehensive calibrated phytolith based climatic index from the Himalaya and its application in palaeotemperature reconstruction. *Science of the Total Environment* 750:142280.
- Bond, W. J., F. I. Woodward, and G. F. Midgley. 2005. The global distribution of ecosystems in a world without fire. *New Phytologist* 165:525–538.
- Boom, A., R. Marchant, H. Hooghiemstra, and J. S. Sinninghe Damsté. 2002. CO_2 - and temperature-controlled altitudinal shifts of C_4 - and C_3 -dominated grasslands allow reconstruction of palaeoatmospheric pCO_2 . *Palaeogeography, Palaeoclimatology, Palaeoecology* 177:151–168.
- Boutton, T. W. 1996. Stable carbon isotope ratios of soil organic matter and their use as indicators of vegetation and climate change. Pp. 47–82 in T. W. Boutton and S.-i. Yamasaki, eds. *Mass spectrometry of soils*. Dekker, New York.
- Bremond, L., A. Alexandre, C. Hély, and J. Guiot. 2005a. A phytolith index as a proxy of tree cover density in tropical areas: calibration with Leaf Area Index along a forest–savanna transect in southeastern Cameroon. *Global and Planetary Change* 45:277–293.
- Bremond, L., A. Alexandre, O. Peyron, and J. Guiot. 2005b. Grass water stress estimated from phytoliths in West Africa. *Journal of Biogeography* 32:311–327.
- Bremond, L., A. Alexandre, O. Peyron, and J. Guiot. 2008a. Definition of grassland biomes from phytoliths in West Africa. *Journal of Biogeography* 35:2039–2048.
- Bremond, L., A. Alexandre, M. J. Wooller, C. Hély, D. Williamson, P. A. Schäfer, A. Majule, and J. Guiot. 2008b. Phytolith indices as proxies of grass subfamilies on East African tropical mountains. *Global and Planetary Change* 61:209–224.
- Bremond, L., A. Boom, and C. Favier. 2012. Neotropical C_3/C_4 grass distributions—present, past and future. *Global Change Biology* 18:2324–2334.
- Brightly, W. H., S. E. Hartley, C. P. Osborne, K. J. Simpson, and C. A. E. Strömberg. 2020. High silicon concentrations in grasses are linked to environmental conditions and not associated with C_4 photosynthesis. *Global Change Biology* 26:7128–7143.
- Castañeda, I. S., and S. Schouten. 2011. A review of molecular organic proxies for examining modern and ancient lacustrine environments. *Quaternary Science Reviews* 30:2851–2891.
- Cerling, T. E., J. Quade, Y. Wang, and J. R. Bowman. 1989. Carbon isotopes in soils and palaeosols as ecology and palaeoecology indicators. *Nature* 341:138–139.
- Cerling, T. E., J. M. Harris, S. H. Ambrose, M. G. Leakey, and N. Solounias. 1997a. Dietary and environmental reconstruction with stable isotope analyses of herbivore tooth enamel from the Miocene locality of Fort Ternan, Kenya. *Journal of Human Evolution* 33:635–650.
- Cerling, T. E., J. M. Harris, B. J. MacFadden, M. G. Leakey, J. Quade, V. Eisenmann, and J. R. Ehleringer. 1997b. Global vegetation change through the Miocene/Pliocene boundary. *Nature* 389:153–158.
- Chazdon, R. L. 1978. Ecological aspects of the distribution of C_4 grasses in selected habitats of Costa Rica. *Biotropica* 10:265–269.
- Chen, S. T., S. Y. Smith, N. D. Sheldon, and C. A. E. Strömberg. 2015. Regional-scale variability in the spread of grasslands in the late Miocene. *Palaeogeography, Palaeoclimatology, Palaeoecology* 437:42–52.

- Collatz, G. J., J. A. Berry, and J. S. Clark. 1998. Effects of climate and atmospheric CO₂ partial pressure on the global distribution of C₄ grasses: present, past, and future. *Oecologia* 114:441–454.
- Conley, D. J., and J. C. Carey. 2015. Silica cycling over geologic time. *Nature Geoscience* 8:431–432.
- Cotton, J. M., N. D. Sheldon, and C. A. E. Strömberg. 2012. High-resolution isotopic record of C₄ photosynthesis in a Miocene grassland. *Palaeogeography, Palaeoclimatology, Palaeoecology* 337–338:88–98.
- Cotton, J. M., T. E. Cerling, K. A. Hoppe, T. M. Mosier, and C. J. Still. 2016. Climate, CO₂, and the history of North American grasses since the Last Glacial Maximum. *Science Advances* 2:e1501346.
- Crepet, W. L., and G. D. Feldman. 1991. The earliest remains of grasses in the fossil record. *American Journal of Botany* 78:1010–1014.
- Crifò, C., and C. A. E. Strömberg. 2021. Spatial patterns of soil phytoliths in a wet vs. dry Neotropical forest: implications for paleoecology. *Palaeogeography Palaeoclimatology Palaeoecology* 526:110100.
- Cui, Y., B. A. Schubert, and A. H. Jahren. 2020. A 23 m.y. record of low atmospheric CO₂. *Geology* 48:888–892.
- Diefendorf, A. F., K. E. Mueller, S. L. Wing, P. L. Koch, and K. H. Freeman. 2010. Global patterns in leaf ¹³C discrimination and implications for studies of past and future climate. *Proceedings of the National Academy of Sciences USA* 107:5738.
- Dunn, R. E., C. A. E. Strömberg, R. H. Madden, M. J. Kohn, and A. A. Carlini. 2015. Linked canopy, climate, and faunal change in the Cenozoic of Patagonia. *Science* 347:258–261.
- Edwards, E. J., and C. J. Still. 2008. Climate, phylogeny and the ecological distribution of C₄ grasses. *Ecology Letters* 11:266–276.
- Edwards, E. J., C. P. Osborne, C. A. E. Strömberg, S. A. Smith, C. G. Consortium, W. J. Bond, P. A. Christin, A. B. Cousins, M. R. Duval, D. L. Fox, R. P. Freckleton, O. Ghannoum, J. Hartwell, Y. Huang, C. M. Janis, J. E. Keeley, E. A. Kellogg, A. K. Knapp, A. D. Leakey, D. M. Nelson, J. M. Saarela, R. F. Sage, O. E. Sala, N. Salamin, C. J. Still, and B. Tipler. 2010. The origins of C₄ grasslands: integrating evolutionary and ecosystem science. *Science* 328:587–591.
- Ehleringer, J. R., and T. E. Cerling. 2002. C₃ and C₄ photosynthesis. *Encyclopedia of Global Environmental Change* 2(4).
- Ehleringer, J. R., T. E. Cerling, and B. R. Helliker. 1997. C₄ photosynthesis, atmospheric CO₂, and climate. *Oecologia* 112:285–299.
- Farquhar, G. D. 1983. On the nature of carbon isotope discrimination in C₄ species. *Functional Plant Biology* 10:205–226.
- Favier, C., J. Aleman, L. Bremond, M. A. Dubois, V. Freycon, and J.-M. Yangakola. 2012. Abrupt shifts in African savanna tree cover along a climatic gradient. *Global Ecology and Biogeography* 21:787–797.
- Fox, D. L., S. Pau, L. Taylor, C. A. Strömberg, C. P. Osborne, C. Bradshaw, S. Conn, D. J. Beerling, and C. J. Still. 2018. Climatic controls on C₄ grassland distributions during the Neogene: a model-data comparison. *Frontiers in Ecology and Evolution* 6:147.
- Fredlund, G. G., and L. T. Tieszen. 1994. Modern phytolith assemblages from the North American Great Plains. *Journal of Biogeography* 21:321–335.
- Gallaher, T. J., S. Z. Akbar, P. C. Klahs, C. R. Marvet, A. M. Senses, L. G. Clark, and C. A. E. Stromberg. 2020. 3D shape analysis of grass silica short cell phytoliths: a new method for fossil classification and analysis of shape evolution. *New Phytol* 228:376–392.
- Garzzone, C. N., N. McQuarrie, N. D. Perez, T. A. Ehlers, S. L. Beck, N. Kar, N. Eichelberger, A. D. Chapman, K. M. Ward, M. N. Ducea, R. O. Lease, C. J. Poulsen, L. S. Wagner, J. E. Saylor, G. Zandt, and B. K. Horton. 2017. Tectonic evolution of the Central Andean Plateau and implications for the growth of plateaus. *Annual Review of Earth and Planetary Sciences* 45:529–559.
- Gibson, D. J. 2009. *Grasses and grassland ecology*. Oxford University Press, New York.
- Gregory-Wodzicki, K. M. 2000. Uplift history of the Central and Northern Andes: a review. *GSA Bulletin* 112:1091–1105.
- Harris, E. B., C. A. E. Strömberg, N. D. Sheldon, S. Y. Smith, and D. A. Vilhena. 2017. Vegetation response during the lead-up to the middle Miocene warming event in the northern Rocky Mountains, USA. *Palaeogeography, Palaeoclimatology, Palaeoecology* 485:401–415.
- Hooghiemstra, H., and T. Van der Hammen. 2004. Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. *Philosophical Transactions of the Royal Society of London B* 359:173–181.
- Hoorn, C., F. P. Wesselingh, H. ter Steege, M. A. Bermudez, A. Mora, J. Sevink, I. Sanmartin, A. Sanchez-Meseguer, C. L. Anderson, J. P. Figueiredo, C. Jaramillo, D. Riff, F. R. Negri, H. Hooghiemstra, J. Lundberg, T. Stadler, T. Sarkinen, and A. Antonelli. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330:927–931.
- Jacobs, B. F., J. D. Kingston, and L. L. Jacobs. 1999. The origin of grass-dominated ecosystems. *Annals of the Missouri Botanical Garden* 86:590–643.
- Kerns, B. K., M. M. Moore, and S. C. Hart. 2001. Estimating forest-grassland dynamics using soil phytolith assemblages and δ¹³C of soil organic matter. *Écoscience* 8:478–488.
- Kirschner, J. A., and C. Hoorn. 2020. The onset of grasses in the Amazon drainage basin, evidence from the fossil record. *Frontiers of Biogeography* 12(2):e44827.
- Ledru, M. P., V. Jomelli, P. Samaniego, M. Vuille, S. Hidalgo, M. Herrera, and C. Ceron. 2013. The Medieval Climate Anomaly and the Little Ice Age in the eastern Ecuadorian Andes. *Clim. Past* 9:307–321.
- Livingstone, D., and W. Clayton. 1980. An altitudinal cline in tropical African grass floras and its paleoecological significance. *Quaternary Research* 13:392–402.
- Lloyd, J., and G. D. Farquhar. 1994. ¹³C discrimination during CO₂ assimilation by the terrestrial biosphere. *Oecologia* 99:201–215.
- Loughney, K. M., M. T. Hren, S. Y. Smith, and J. L. Pappas. 2019. Vegetation and habitat change in southern California through the Middle Miocene Climatic Optimum: Paleoenvironmental records from the Barstow Formation, Mojave Desert, USA. *GSA Bulletin* 132:113–129.
- Lupien, R. L., J. M. Russell, C. L. Yost, J. D. Kingston, A. L. Deino, J. Logan, A. Schuh, and A. S. Cohen. 2021. Vegetation change in the Baringo Basin, East Africa across the onset of Northern Hemisphere glaciation 3.3–2.6 Ma. *Palaeogeography, Palaeoclimatology, Palaeoecology* 570:109426.
- Madella, M., M. Jones, P. Echlin, A. Powers-Jones, and M. Moore. 2009. Plant water availability and analytical microscopy of phytoliths: implications for ancient irrigation in arid zones. *Quaternary International* 193:32–40.
- Mander, L., M. Li, W. Mio, C. C. Fowlkes, and S. W. Punyasena. 2013. Classification of grass pollen through the quantitative analysis of surface ornamentation and texture. *Proceedings of the Royal Society of London B* 280:20131905.
- Martínez, C., C. Jaramillo, A. Correa-Metrio, W. Crepet, J. E. Moreno, A. Aliaga, F. Moreno, M. Ibañez-Mejía, and M. B. Bush. 2020. Neogene precipitation, vegetation, and elevation history of the Central Andean Plateau. *Science Advances* 6:eaa24724.
- McInerney, F. A., C. A. E. Strömberg, and J. W. C. White. 2016. The Neogene transition from C₃ to C₄ grasslands in North America: stable carbon isotope ratios of fossil phytoliths. *Paleobiology* 37:23–49.
- Meyers, P. A., and R. Ishiwatari. 1993. Lacustrine organic geochemistry—an overview of indicators of organic matter sources and diagenesis in lake sediments. *Organic Geochemistry* 20:867–900.
- Miller, L. A., S. Y. Smith, N. D. Sheldon, and C. A. E. Stromberg. 2012. Eocene vegetation and ecosystem fluctuations inferred from a high-resolution phytolith record. *Geological Society of America Bulletin* 124:1577–1589.

- Neumann, K., C. A. E. Strömberg, T. Ball, R. M. Albert, L. Vrydaghs, and L. S. Cummings. 2019. International code for phytolith nomenclature (ICPN) 2.0. *Annals of Botany* 124:189–199.
- Palmer, P. G. 1976. Grass cuticles: a new paleoecological tool for East African lake sediments. *Canadian Journal of Botany* 54:1725–1734.
- Parmenter, C., and D. W. Folger. 1974. Eolian biogenic detritus in deep sea sediments: a possible index of equatorial ice age aridity. *Science* 185:695–698.
- Pau, S., E. J. Edwards, and C. J. Still. 2013. Improving our understanding of environmental controls on the distribution of C₃ and C₄ grasses. *Global Change Biology* 19:184–196.
- Pearsall, D. M. 2000. *Paleoethnobotany: a handbook of procedures*. Academic Press, San Diego.
- Piperno, D. R. 1988. *Phytolith analysis: an archaeological and geological perspective*. Academic Press, San Diego.
- Prentice, I. C. 1985. Pollen representation, source area, and basin size: toward a unified theory of pollen analysis. *Quaternary Research* 23:76–86.
- Quade, J., N. Solounias, and T. E. Cerling. 1994. Stable isotopic evidence from paleosol carbonates and fossil teeth in Greece for forest or woodlands over the past 11 Ma. *Palaeogeography, Palaeoclimatology, Palaeoecology* 108:41–53.
- R Development Core Team. 2019. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rieley, G., R. J. Collier, D. M. Jones, G. Eglinton, P. A. Eakin, and A. E. Fallick. 1991. Sources of sedimentary lipids deduced from stable carbon-isotope analyses of individual compounds. *Nature* 352:425–427.
- Rundel, P. W. 1980. The ecological distribution of C₄ and C₃ grasses in the Hawaiian Islands. *Oecologia* 45:354–359.
- Sage, R. F., P. A. Christin, and E. J. Edwards. 2011. The C₄ plant lineages of planet Earth. *Journal of Experimental Botany* 62:3155–3169.
- Sage, R. F., R. K. Monson, J. R. Ehleringer, S. Adachi, and R. W. Pearcy. 2018. Some like it hot: the physiological ecology of C₄ plant evolution. *Oecologia* 187:941–966.
- Simpson, B. B., and C. A. Todzia. 1990. Patterns and processes in the development of the high Andean flora. *American Journal of Botany* 77:1419–1432.
- Smith, B. N., and S. Epstein. 1971. Two categories of ¹³C/¹²C ratios for higher plants. *Plant Physiology* 47:380–384.
- Song, Z., K. McGrouther, and H. Wang. 2016. Occurrence, turnover and carbon sequestration potential of phytoliths in terrestrial ecosystems. *Earth-Science Reviews* 158:19–30.
- Still, C. J., J. A. Berry, G. J. Collatz, and R. S. DeFries. 2003. Global distribution of C₃ and C₄ vegetation: carbon cycle implications. *Global Biogeochemical Cycles* 17(1):6-1-6-14.
- Strömberg, C. A. E. 2005. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy of Sciences USA* 102:11980–11984.
- Strömberg, C. A. E. 2009. Methodological concerns for analysis of phytolith assemblages: does count size matter? *Quaternary International* 193:124–140.
- Strömberg, C. A. E. 2011. Evolution of grasses and grassland ecosystems. *Annual Review of Earth and Planetary Sciences* 39:517–544.
- Strömberg, C. A. E., and F. A. McInerney. 2016. The Neogene transition from C₃ to C₄ grasslands in North America: assemblage analysis of fossil phytoliths. *Paleobiology* 37:50–71.
- Strömberg, C. A. E., E. M. Friis, M.-M. Liang, L. Werdelin, and Y.-L. Zhang. 2007a. Palaeoecology of an Early-Middle Miocene lake in China: preliminary interpretations based on phytoliths from the Shanwang Basin. *Vertebrata Palasiatica* 45:145–160.
- Strömberg, C. A. E., L. Werdelin, E. M. Friis, and G. Saraç. 2007b. The spread of grass-dominated habitats in Turkey and surrounding areas during the Cenozoic: phytolith evidence. *Palaeogeography, Palaeoclimatology, Palaeoecology* 250:18–49.
- Strömberg, C. A. E., R. E. Dunn, R. H. Madden, M. J. Kohn, and A. A. Carlini. 2013. Decoupling the spread of grasslands from the evolution of grazer-type herbivores in South America. *Nature Communications* 4:1478.
- Strömberg, C. A. E., V. S. Di Stilio, and Z. Song. 2016. Functions of phytoliths in vascular plants: an evolutionary perspective. *Functional Ecology* 30:1286–1297.
- Teeri, J., and L. Stowe. 1976. Climatic patterns and the distribution of C₄ grasses in North America. *Oecologia* 23:1–12.
- Thorn, V. 2001. Oligocene and early Miocene phytoliths from CRP-2/2A and CRP-3, Victoria Land Basin, Antarctica. *Terra Antarctica* 8:407–422.
- Tieszen, L. L., M. M. Senyimba, S. K. Imbamba, and J. H. Troughton. 1979. The distribution of C₃ and C₄ grasses and carbon isotope discrimination along an altitudinal and moisture gradient in Kenya. *Oecologia* 37:337–350.
- Trembath-Reichert, E., J. P. Wilson, S. E. McGlynn, and W. W. Fischer. 2015. Four hundred million years of silica biomineralization in land plants. *Proceedings of the National Academy of Sciences USA* 112:5449–5454.
- Twiss, P. C. 1992. Predicted world distribution of C₃ and C₄ grass phytoliths. Pp. 113–128 in G. Rapp and S. C. Mulholland, eds. *Phytolith systematics*. Springer, New York.
- Urban, M. A., D. M. Nelson, G. Jiménez-Moreno, J.-J. Châteauneuf, A. Pearson, and F. S. Hu. 2010. Isotopic evidence of C₄ grasses in southwestern Europe during the Early Oligocene–Middle Miocene. *Geology* 38:1091–1094.
- Winslow, J. C., E. R. Hunt Jr., and S. C. Piper. 2003. The influence of seasonal water availability on global C₃ versus C₄ grassland biomass and its implications for climate change research. *Ecological Modelling* 163:153–173.
- Wynn, J. G., and M. I. Bird. 2007. C₄-derived soil organic carbon decomposes faster than its C₃ counterpart in mixed C₃/C₄ soils. *Global Change Biology* 13:2206–2217.
- Young, H. J., and T. P. Young. 1983. Local distribution of C₃ and C₄ grasses in sites of overlap on Mount Kenya. *Oecologia* 58:373–377.
- Zhou, H., B. R. Helliker, M. Huber, A. Dicks, and E. Akcay. 2018. C₄ photosynthesis and climate through the lens of optimality. *Proceedings of the National Academy of Sciences USA* 115:12057–12062.