Phytoliths from modern and ancient habitats: toward a modern-based, quantitative approach to reconstruct vegetation change during the MMCO of Patagonia, Argentina

Camilla Crifo'

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Caroline A. E. Strömberg, Chair

Janneke Hille Ris Lambers

Richard Olmstead

Program Authorized to Offer Degree:

Department of Biology

University of Washington

Abstract

Phytolith from modern and ancient habitats: toward a modern-based, quantitative approach to reconstruct vegetation change during the MMCO of Patagonia, Argentina

Camilla Crifo'

Chair of the Supervisory Committee:
Professor Caroline A.E. Strömberg
Biology

Phytolith analysis has high potential for reconstructing past vegetation with higher spatial resolution compared other high-resolution proxies, such as pollen and spores. Phytolith assemblages are used in paleoecology to reconstruct changes in vegetation structure through time. In addition, spatial variability of the phytolith signal (across samples collected along a single stratigraphic level) is interpreted as indicative of habitat heterogeneity based on the notion that phytolith assemblages are derived from vegetation that died and decayed in place and therefore hold a local signal. However, this and other assumptions have not yet been tested directly in modern environments; current data are insufficient to establish modern calibrations for the deep time phytolith record, and thus understand the fossil phytolith records in different vegetation types.

In Chapter 1 and 2 of this dissertation I aim at helping bridging this gap, by 1) defining an appropriate methodology to sample phytolith for modern analogue studies that is applicable to

the deep-time phytolith record; 2) and by providing a modern reference study of soil phytolith along transects in two Neotropical vegetation types in Costa Rica: a rainforest and a dry forest. I investigate the following questions: 1) how many samples and from which part of the (phytolith-rich) soil A-horizon are needed to reflect accurately the standing vegetation? (Chapter 1); 2) are gradients in vegetation structure, composition, and diversity recorded in phytolith assemblages across transects in rainforest and dry forest soils? (Chapter 2); and 3) can we use one or more phytolith assemblages to characterize these two vegetation types, and distinguish them in the fossil record? (Chapter 2).

In Chapter 3, I apply the lessons learned from Chapter 1 and 2 to the study of vegetation heterogeneity and vegetation change in Patagonia, at the onset of the Middle Miocene Climatic Optimum (MMCO) –the last global warming event taking place on Earth before the current one, between ~17 and 14.5 Ma. The MMCO is poorly documented in the Southern Hemisphere and at high latitudes. The Santa Cruz Formation (SCF), in southern Patagonia, is an exception, preserving one of the most diverse and well-preserved fauna assemblages on Earth. Fauna and stable isotope data from the SCF suggest that global warming associated with increased aridity favored heterogeneous habitats characterized by many ecological niches which were able to support abnormally high fauna diversity. The phytolith record of SCF has been so far poorly studied but constitute the best line of evidence for high resolution reconstruction of vegetation change through time as well as of spatial patterns of vegetation variability (heterogeneity). Using phytolith assemblages from the SCF I investigate the following questions: 1) How did vegetation structure change in response to the initial warming pulse of the MMCO? 2) How did grass community composition change in response to warmer and drier conditions at the onset of the MMCO? 3)Was the remarkably high diversity of the Santa Cruz fauna supported by habitats

characterized by vegetation heterogeneity (i.e., a mix of forested and open vegetation areas) throughout the onset of the MMCO as would be predicted based on modern ecology and SCF faunal data?

In Chapter 1 phytolith from modern soil assemblages from two Neotropical forests in Costa Rica (a dry forest and a rainforest) are studied to determine a sample strategy for future modern analogue studies that is applicable to the phytolith deep-time record. Results suggest that the typical approach in deep-time paleoecology of taking point samples from the lower A-horizon of paleosols is justifiable (at least for paleosols reflecting rainforest and dry forest soils), and should therefore be implemented in future phytolith modern analogues studies that aim at improving interpretations of the deep-time phytolith record. Thus, the results of Chapter 1 constitute the basis upon which the modern analogue study described in Chapter 2 was conducted

In Chapter 2, additional soil phytolith assemblages collected along vegetation transects are used to investigate whether and how soil phytoliths reflect gradients in vegetation structure, composition and diversity across the two habitat types (dry forest and rainforest). In all, our results demonstrate that phytolith assemblages can definitely distinguish dry and wet forest habitats. In addition, our results also suggest that phytolith assemblage characteristics within vegetation types do not capture all aspect of environmental and plant community gradients. However, overall higher environmental heterogeneity of the dry forest results in higher heterogeneity of the phytolith assemblages. This result suggests that overall, spatial sampling (along a transect) and the analysis of phytolith assemblage composition allow to reconstruct some structural, and compositional aspects of habitat heterogeneity, and that that phytolith assemblage heterogeneity within a habitat might be indicative of habitat heterogeneity.

In Chapter 3, phytolith assemblages from The Santa Cruz Formation (Patagonia) spanning the onset of the Middle Miocene Climatic Optimum (MMCO) are analyzed to reconstruct vegetation response to the climatic event as well as to reconstruct vegetation heterogeneity across two stratigraphic layers, representing two snapshots of the SCF vegetation at two different times. Results show that before the onset of the MMCO southeastern Patagonia was characterized by heterogeneous habitats with abundant pooid C_3 grasses and a woody component represented by conifers, dicots, as well as palms in varying abundance. This habitat corresponded to woodland or open woodland/shrubland, including palm shrubland. In the upper SCF, at the onset of the MMCO (inferred from isotopic data to be drier), grass abundance decreased, and phytolith assemblages indicate that the landscape was dominated by a woody component of the vegetation. In addition, grass communities were dominated by C_3 pooid grasses whereas grasses of the tropical PACMAD clade (which includes both C_3 and C_4 grasses) were only a minor component of grass communities. We interpret these trends as reflecting the expansion of dry-adapted woody vegetation in response to MMCO climate change, and to the detriment of a C_3 grass community which was not adapted to dry conditions. Further, we suggest that PACAMD grasses at the SCF were likely primarily C_3, and the expansion of dry-adapted C_4 grasses and grass-dominated open habitats did not take place in Patagonia until after the early middle Miocene.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, I want to express my gratitude to Caroline Strömberg, for her intellectual enthusiasm, and support, her patience, and understanding throughout the course of my dissertation research. I deeply appreciate Caroline's commitment to be an excellent role model, and to help me growing as a scientist. I want to thank my committee members, Janneke Hille Ris Lambers, Dick Olmstead, and Gregory Wilson, and my GSR Roger Buick for their contribution to my dissertation with useful advice and insights. I want to thank Jennifer Ruesnik for accepting enthusiastically to be an *ad interim* member of my committee until my general exam. I want thank Dick Olmstead for kindly accepting to be a "jumping" member of my committee –initially on, then off when I decided to switch the focus of my research, and on again when I realized that, independently of the focus of my research, Dick's advice would have been precious.

For work in Argentina, I want to thank my collaborators: Susana Bargo, José Cuitiño, Matt Kohn, Robin Trayler, Richard Kay, Sergio Vizcaíno, Alejandro Zucol. I am especially grateful to José Cutiño and Nahuel Muñoz who generously and patiently thought me and help me during my first field experience collecting phytolith samples; without them I couldn't have survived the first (and following) day. I want to thank all other field companions: Juan Carlos Fernicola, Santiago Hernández Del Pino, Néstor Toledo for their help, interesting conversations, insights and generosity in the field. Thanks to Luciano Zapata for sharing his sample with us.

For work in Costa Rica, I want to thank the Organization for Tropical Studies. In particular I want to thank La Selva and Palo Verde Biological Stations and their staff, as well as Orlando Vargas, and Ulises Chavarria from La Selva and Palo Verde herbaria. I want to thank my field assistant Mireya Córdoba, and Leith Miller for help in the field. I am grateful to Francisco Campos

from MINAE and Davinia Beneyto who helped me with logistics, and collection and export permits.

I am grateful to the UW Biology staff including Mitchell Chen, Davies Chong, Alex Chun, Krista Clouser, Michele Conrad, Brianna Divine, Janet Gemeraad, David Giblin, Alex Hansen, Aaron Hernandez, Dave Hurley, Ronald Killman, Nile Kurashige, Jeanette Milne, Sarah O'Hara, John Parks, Eduardo Sabiniano, Christine Savolainen, who greatly facilitated working at UW Biology and in most of the cases, simply made it possible.

I want to thank Scott Freeman, John Parks, Dick Olmstead, Audrey Ragsac, Ana Bedoya, and all fellow TAs for their help and support during my teaching assistant experience.

I want to thank Katherine Anderson and Ron Eng of the Burke Museum, the Paleolunch group, as well as all previous and current members of the Wilson, and Sidor labs with whom I had the pleasure of spending some time. I want to thanks the Strömberg previous undergrads, lab grad, postdocs: Ryan Thummel, Regan Dunn, Elisha Harris, Will Brightly, Tim Gallaher, Alice Novello, Paige Wilson, and Elena Stiles. I am very grateful to Erick Frederickson teaching me phytolith processing and explaining me everything about the Strömberg lab when I first arrived.

I am deeply thankful to many undergraduate researchers whose help in the lab was made my research possible: Thy Huyhn, Erin Sofinoski, Sarah Larson, Kevin Jackson, Matt Butrim, Una O'Connell, Matt Bloch, Gabrielle Alampay, Kirsten Olson, Kristen Hamel, Alexander Arrendale, and Madeline Mamer.

I want to thank Ellen Currano, who inspired and supported me, and "pushed me into the harms" of the Strömberg lab. I want to thank my first mentor and academic father Carlos Jaramillo, who 10 years ago welcomed me to Panama, initiating my scientific journey.

I am grateful to all UW grads for interesting and fun conversations, and their commitment to improve grad student life. I want to thank my friends in graduate school: Johanna Cantillo for giving me the warmest welcome to the rainy Seattle; Alex T. Lowe the kindest, whitest, and taller male friend I have ever had; Dave Slager for laughing with me in unbearably boring situations; and Leith Miller for her unconditional love and support, and for sharing her laugh, cry, and many beers with me. I want to thank my friends in Seattle for all the fun, their support, and friendship; I could not imagine Seattle without them: Maria José Bravo, Piera Carpi and Callum, Mariona Claret, Lise Comte and Santiago Baez, the little Milo Baez-Comte, David Li, Maite Pons, Max Needle, Glwadys and Sholto Redford... I am grateful to all my high school friends who have been present in my life for more than 20 years; to my friend Tasneem Elzein for her light and love that I can feel from huge distances; and to my oldest friend Martina Giaccone, who is my North Star. Last, I want to thank my family for always being my home, even 9,110 kilometers away; and my partner, for loving me, for trusting me, and for always teasing, and supporting me.

If I forgot anyone, I apologize, and I want you to know that I appreciate your friendship and support.

DEDICATION

To my friends, and family.

"Trouver la joie dans le ciel, dans les arbres, dan les fleurs. Il y a des fleurs partout pour qui veut bien les voir".

Henri Matisse, 1947. "Jazz".

Chapter 1. Small-scale spatial resolution of the soil phytolith record in a Neotropical rainforest and a dry forest in Costa Rica: applications to the deep-time fossil phytolith record

ABSTRACT

Phytolith analysis, a well-established tool in archaeology and Quaternary paleoecology, has become a source of data for deep-time paleoecological studies only in the last 15 years. Recent years have also witnessed the publication of numerous soil phytolith inventories from extant vegetation types, representing modern analogues. However, this work suffers from several methodological shortcomings limiting inter-study comparisons, and the development of a single, and repeatable protocol for soil sample collection. In this paper, we focus on two fundamental methodological questions that must be answered before modern analogue studies can be meaningfully applied to deep-time paleoecology. (1) Do phytoliths from the lower portion of the A-horizon (more commonly preserved in paleosols) reflect standing vegetation equally well as the upper A-horizon (which is more commonly used in modern analogue studies but often truncated by erosion in paleosols)? (2) Does a point soil sample (usually collected from paleosols) capture standing vegetation as accurately as a composite soil sample (commonly used in modern analogue studies), which is made from subsamples taken from a small area? To address these questions, we collected soil from 10x10 m quadrats in a rainforest (La Selva) and a dry forest (Palo Verde) in Costa Rica and compared phytolith assemblages from upper vs. lower soil A-horizons, and composites vs. single phytolith assemblages. At both sites, phytolith assemblages from the lower vs. upper A-horizon, and from single vs. composite samples provide similar vegetation structure signal despite some compositional differences are found. These results suggest that the typical

approach in deep-time paleoecology of taking point samples from the lower A-horizon of paleosols is likely justifiable (at least for paleosols reflecting rainforest and dry forest soils). We discuss possible mechanisms driving differences between sites, and future approaches to further expand and test our results.

1.1 INTRODUCTION

Phytoliths represent an emerging tool in paleoecology with unique characteristics compared to other paleobotanical proxies (Strömberg et al., 2018). Unlike palynomorphs and plant macrofossils, phytolith are often found in sediments associated with vertebrate remains, allowing direct study of the relationships between past habitats and faunas (Strömberg, 2002). Moreover, phytolith assemblages represent a more local vegetation signal compared to the generally coarse (regional) spatial resolution of the pollen and spore record (Prentice, 1985; Piperno, 1988), making phytolith analysis a more suitable tool for the study of variability in local vegetation across the landscape.

Phytolith analysis has been extensively used in the last 40 years in Quaternary paleoecology and archaeobotany to study a wide range of topics, including plant domestication history (e.g., Iriarte, 2003; Piperno, 2009; Rosen & Weiner, 1994), early anthropogenic disturbance (e.g., Kealhofer, 1996; McMichael et al., 2012, Tromp & Dudgeon, 2015), early human diet (e.g., Ciochon et al, 1990; Henry et al, 2011, Tromp & Dudgeon, 2015), and past vegetation and climate (Kealhofer & Penny, 1998; Bremond et al, 2017). In contrast, concerted efforts to apply phytolith analysis in deep-time paleoecological research has occurred only in the last ~15 years (see discussion in Strömberg 2004, Strömberg et al., 2018), despite its potential and demonstrated utility in Quaternary research. So far, phytolith analysis from paleosols and lacustrine or marine sediments has helped elucidate the Cretaceous-Cenozoic origin and spread

of grasses and grasslands (e.g. Prasad et al., 2005; Strömberg et al., 2007a, 2013; Strömberg and McInerney, 2011; Cotton et al., 2014; Chen et al., 2015), and, more broadly, reconstruct past vegetation history across several continents (e.g. Thorn, 2001; Strömberg et al., 2007b; Zucol et al., 2007;Miller et al., 2012; Dunn et al., 2015). Phytolith assemblages from teeth or sediment associated with vertebrate fossils have further allowed inference of animal diet, and flora-fauna relationships through evolutionary time (e.g. Henry et al., 2012; Merceron et al., 2016; Strömberg, 2006).

Nevertheless, because phytolith science is a relatively young field, its limitations and biases are not all known in detail. One basic question that remains at least partially unanswered is, how well (and at what spatial scale) do soil phytolith assemblages represent (modern) standing vegetation? To address this question, phytolith modern analogue studies have been undertaken in different ecosystems (Piperno 1988; Fredlund and Tieszen 1994; Kerns et al 2001; Bremond et al., 2005; Mercader et al., 2011; Aleman et al., 2014), ultimately seeking to devise protocols for how to accurately and reliably interpret the phytolith paleosol record. However, the number of such studies is still limited, and restricted to a few geographic regions and vegetation types. Moreover, these studies use diverse sampling and analytical approaches, building on untested assumptions about the nature and importance of different taphonomical processes, which act to variously bias the phytolith record.

The lack of standardization remains a critical gap in phytolith research that complicates meaningful comparisons of vegetation inferences from modern-submodern and deep-time phytolith assemblages. Our work seeks to start filling this gap. In particular, we focus on issues concerning the construction of a *consistent methodology of sampling* for modern soil analogue studies designed to interpret the phytolith paleosol record. By defining basic guidelines for the

collection of modern soil samples we hope to build the foundation for future modern analogue studies that seek to refine and widen our understanding of the spatial resolution of the phytolith record in ancient soils.

1.2 PHYTOLITH ASSEMBLAGE FORMATION: CURRENT UNDERSTANDING

Below, we provide a overview of: 1) the main, known taphonomical processes and associated biases influencing the phytolith record (focusing mainly on soil phytoliths); 2) the core lessons from previous modern analogue studies of soil phytoliths; and 3) two important assumption made in these studies. We then test these assumptions in two Neotropical ecosystems, a rainforest, and a dry forest in Costa Rica.

1.2.1 *Taphonomy and associated biases of the phytolith record*

Taphonomic processes include phytolith production, dispersal, inheritance, and preservation, as well as phytolith intrinsic properties such as size, shape, and texture (see Table 1). These processes vary among vegetation types resulting in different biases.

1.2.1.1 Production

Phytoliths are produced by most vascular plants but at varying abundances. Among the biggest phytolith producers are monocotyledonous angiosperms (e.g., the Poaceae and Arecaceae families), ferns, horsetails, and some tropical non-monocotyledonous angiosperms (hereafter 'dicots'); by contrast, non-angiosperm seed plants and many dicots from temperate regions are among the lowest phytolith producers (Hodson et al., 2005; Piperno 2006; Katz, 2015; Trembath-Reichert et al., 2015, Strömberg et al., 2016, 2018). These differences might result in

overrepresentation of certain taxa (e.g., grasses and palms), especially in temperate regions where dicotyledonous trees produce less biosilica compared to their tropical relatives (Carnelli et al., 2001; Hyland et al., 2013; Strömberg, 2004).

1.2.1.2 Dispersal

Among the most common transporting agents of phytoliths are herbivores, fire, wind, and water (fluvial/colluvial input) (Piperno 1988; Fredlund and Tieszen, 1994; Wallis, 2001). In general, phytolith transport acts to homogenize phytolith assemblages across the landscape and lower the spatial resolution of the phytolith signal, although to what degree depends on the agent. For example, herbivores can transport phytoliths over distances corresponding to their habitat range. Fire can be enhanced by wind (Piperno 1988; Wallis 2001; Bremond et al., 2005), especially in open habitats, so that phytolith transport can take place over hundreds of kilometers (Folger 1970; Folger et al., 1967; Locker and Martini 1986; Wallis, 2001; Piperno, 2006; Osterrieth et al., 2009;). Whereas phytolith assemblages in hydrologically closed lakes (e.g., maar lakes) primarily reflect vegetation near the lake shore (Strömberg et al., 2007b), transport by water in large lakes with fluvial input might result in phytolith assemblages reflecting the regional vegetation rather than the vegetation surrounding the lake (Zhao and Piperno, 2000; Aleman et al., 2014). In addition, variation in shape, size, and specific gravity among phytolith morphotypes might result in preferential transport (by wind and water) of certain morphotypes, in particular palm, and grass phytoliths (Folger 1970; Strömberg et al., 2018).

1.2.1.3 Inheritance

Phytoliths are deposited *in situ* in the soil A-horizon when organic plant material decays on its surface (O-horizon) or belowground, although lateral transport also takes place at varying degrees under different canopy types (see above and Table 1). The distribution of phytoliths in

the soil horizon depends on the interplay between sedimentation rate and processes in the soil. Under certain circumstances (e.g., rapid sedimentation), soil phytolith assemblages can become stratified within a soil profile (Piperno, 1988, 2006; Kerns et al., 2001; White et al., 2012). On the other hand, phytolith vertical translocation, due to bioturbation and illuviation might lead to time averaging and size sorting along soil profiles (Fishkis et al., 2009, 2010a, 2010b; Madella and Lancelotti, 2012). Biosilica dissolution in the soil column both acts to bias the soil phytolith pool in various ways (see below) and to limit the average residence time of phytoliths; the latter in some sense counteracting the time-averaging resulting from processes such as bioturbation (Alexandre et al., 1997; Kerns et al., 2001; Fishkis et al 2010a; Song et al., 2016; Strömberg et al., 2018).

Environmental conditions influence the turnover time of the phytolith pool (shorter in wet and warm climates) (Song et al., 2016), as well as the degree to which phytoliths are translocated by bioturbation and illuviation.

The resulting distribution of phytoliths in modern soils, is one by which phytolith concentration is typically higher within the first few centimeters and decreases gradually with depth, with a drastic decline below 30-80 cm due to vertical translocation and dissolution processes (Beavers and Stephen, 1958; Kelly et al., 1991; Alexandre et al., 1997, 2011; Piperno and Becker, 1996; Meunier et al., 1999; Blecker et al 2006; White et al., 2012). In lacustrine settings, phytoliths, like other clastic and organic fossils, can be transported and preserved as part of non-bioturbated, laminated sediment, resulting in a typically very fine temporal resolution (Strömberg et al., 2018).

1.2.1.4 Preservation

Phytoliths in soils undergo different degrees of dissolution depending on soil conditions and phytolith intrinsic properties (Song et al., 2016; Strömberg et al., 2018). Dissolution rates are lower in acidic (pH<8.5) over alkaline soils (Alexandre et al., 1997; Fraysse et al., 2006; Ehrlich et al., 2010; Prentice and Webb, 2016), and lacustrine over marine settings (although they are known to preserve in marine sediments well; see Locker and Martini 1986; Thorn, 2001). Among phytolith properties, specific surface area (surface area per unit volume) seem to have the biggest effect on phytolith dissolution rates (Fraysse et al., 2006; Ehrlich et al., 2010), overall favoring the preservation of grass phytolith morphotypes over others.

1.2.2 *Modern analogue studies and phytolith spatial resolution*

Understanding phytolith spatial resolution no doubt has important implications for how we use modern phytolith assemblages to make inferences about vegetation in ancient soils. For instance, a 'super-local' phytolith signal would enable paleoecologists to reconstruct vegetation spatial heterogeneity within landscapes but prevent them from drawing broad conclusions about regional vegetation change from a small number of samples. Conversely, an extra-local signal limits inferences to more regional vegetation patterns, and precludes direct reconstruction of spatial heterogeneity.

Only a handful of studies has investigated this question thus far. They show that the spatial resolution of the phytolith record varies from the scale of tens of meters in tropical rainforests (Piperno, 1998), and African savannas (Aleman et al., 2012), hundreds of meters in a forest-savanna transition in Africa (Bremond et al., 2005), to tens of kilometers in a woodland-grassland transition in Mozambique (Mercader et al., 2011). In other ecosystems, only a week

relationship between soil phytolith assemblages and vegetation structure has been found (Kerns, 2011), resulting in a regional rather than local spatial signal of the phytolith record (Fredlund and Tieszen, 1994), likely as a result of spatial mixing. Furthermore, the substantial differences in resolution among these studies seem to be tied to taphonomic processes, unique to different vegetation types, and resulting in distinct biases in the phytolith record. Phytoliths assemblages preserved in lacustrine sediments show similar variability linked to lake taphonomy (Aleman et al., 2014; see also discussion in Zhao and Piperno, 2000, Strömberg et al., 2007a).

Further surveys across vegetation types are therefore needed to calibrate the phytolith method and, ultimately, to link the results in a comprehensive, quantitative, and predictive model for how phytoliths move through the environment, akin to what has been done in palynology (e.g., Prentice et al., 1996; Whitmore et al., 2005). The first step in this direction is to determine the appropriate methodology to sample soil phytolith assemblages for modern analogue studies applicable to deep-time paleoecology. Below, we highlight the two main assumptions on which current sampling strategies are based, and then describe tests of these assumptions.

1.2.3 *Common assumptions and hypotheses to be tested*

The first, common assumption that we test is that soil upper A-horizons are analogous to sampled paleosols. In fact, most previous modern analogue studies have used samples from the upper soil A-horizon where phytoliths are known to usually be more abundant, and phytolith assemblages are less prone to bias (Beavers and Stephen, 1958; Runge, 1999; Song et al., 2016). The vast majority of such surveys (e.g., Fredlund and Tieszen, 1994; Kerns et al 2001; Bremond et al., 2005; Mercader et al 2011; Aleman et al., 2012) have attested to the validity of upper A-horizon phytolith assemblages for reconstructing the vegetation cover. However, it is not clear that these results are directly applicable to deep-time studies. In paleosols, the upper portion of

the soil A-horizon is frequently partially or completely absent due to erosion (Tabor and Myers, 2015), such that the lower (rather than the upper) portion of the A-horizon, if not the B horizon, is predominantly sampled (e.g., Miller et al., 2012; Chen et al., 2015;). It remains to be tested whether the upper and lower A-horizon are comparable in terms of their phytolith signal (but see Kerns et al., 2001), thus, whether conclusions drawn from modern analogue studies can be used to interpret phytolith assemblages from more ancient paleosols.

We propose to tackle the issue of differential sampling depth by comparing modern soil phytolith assemblages corresponding to the upper soil A-horizon with phytolith assemblages from the lower part of the soil A-horizon, and with the standing vegetation in two different Neotropical vegetation types, namely, a rainforest and a dry forest in Costa Rica. We test the hypothesis ($H_{1,horizon}$) that phytoliths assemblages in the upper A-horizon represent the vegetation significantly better than phytolith assemblages in the lower A-horizon, which (a) might be phytolith-depleted and/or biased toward phytolith morphologies/sizes that preferentially accumulate in lower parts of the soil horizon, or (b) might record sub-modern vegetation that differed substantially from that growing today. If this hypothesis is valid, the two soil layers should yield statistically different phytolith assemblages. Our null hypothesis ($H_{0,horizon}$) is that phytolith assemblages from the upper and lower soil A-horizon are equally representative of the standing vegetation in a Neotropical rainforest and dry forest. If this latter hypothesis is valid, the two soil layers should yield statistically indistinguishable phytolith assemblages.

The second assumption we propose to test here is that spatially composite samples better record vegetation than single-point samples. Phytolith workers commonly solve the problem of studying soil phytolith assemblages that optimally represent standing vegetation by collecting a variable number of modern soil subsamples from a given area and then combining them into a

composite (mixed) sample (see Table 1). This practice seems to assume that there is more between-sample variation (in phytolith morphotypes) in soil phytolith assemblages collected within a small area than there is within a single sample. However, this assumption does not take into account the fact that time averaging of a single-point sample might produce greater or equal variability in the phytolith assemblage composition (e.g., by lateral vegetation shifts through time) than the "spatial averaging" created by mixing several samples. In other words, the phytolith composition of a mixed sample might result from both spatial variability between phytolith assemblages, and temporal variability within each single-point sample—two effects that are undistinguishable. This distinction matters, because in the pre-Quaternary paleosol record lateral samples can rarely be assumed to be exactly coeval (but see e.g., Strömberg et al., 2013); thus, most paleoecologists are limited to the use of point-samples (Strömberg et al., 2018).

We propose to address the question of mixed- versus point samples by comparing the phytolith assemblages of a single sample with a mix of several samples collected in a given area. We test the hypothesis ($H_{1,\text{sample type}}$) that a phytolith assemblage from a composite sample represents the vegetation better than a phytolith assemblage from a single sample, which might provide a 'super-local' signal. If this hypothesis is valid, the two sample types should yield statistically different phytolith assemblages. Alternatively, our null hypothesis ($H_{0,\text{sample type}}$) is that soil phytolith assemblages from a single vs. a composite sample are equally representative of the standing vegetation in a neotropical rainforest and dry forest.

1.3 MATERIALS AND METHODS

1.3.1 *Study sites*

We conducted fieldwork in Costa Rica at the Organization for Tropical Studies (OTS) La Selva, and Palo Verde biological stations during winter 2016 (Fig. 1.1). La Selva Biological Station (10° 26' N, 83° 59' W) is a protected area located at the confluence of two major rivers in the Caribbean lowland of northern Costa Rica. Part of the Cordillera Volcánica Central Biosphere Reserve, it is situated between premontane and wet forest, with primary tropical rainforest covering 73% of its area (Organization for Tropical Studies, 2017). Soils are derived from Andesitic lava flows (Alvarado Induni, 1990), and were classified as Oxisols (Kleber et al., 2007); sparse marshy depressions are also present (Organization for Tropical Studies, 2017). Palo Verde Biological Station (10° 21' N, 85° 21' W) is located within the Palo Verde National Park, on the Pacific slopes of Guanacaste province in northwestern Costa Rica, and borders the Tempisque river, which flows into the Gulf of Nicoya. It comprises several habitats, including mangrove and swampy forests, salty and freshwater marshes, evergreen forests, deciduous forests, lowland and limestone forests, and secondary forests. The dry forest, located on limestone outcrops, is a remnant of seasonal and transitional dry forests in the Neotropics (Organization for Tropical Studies, 2017). Soils range from Inceptisols (dominated by limestone) in the uplands, to Vertisols (older, deeper soils composed of a mix of limestone, and clay/silt deposits derived from the Tempisque River) in the lowlands (Powers et al., 2009). We chose these two sites (referred to as La Selva, and Palo Verde) because they are both undisturbed, protected areas with known land use history.

1.3.2 *Soil collection*

We collected and analyzed 22 soils samples from La Selva, and 62 soil samples from Palo Verde (Ministry of the Environment and Energy research permit # 006-2016-INV-ACAT; export permits # DGVS-076-2016, and DGVS-082-2016). In each site, we established one primary 10 x 10-meter quadrat (Q1). We chose quadrat location randomly but avoided disturbed patches, flooded areas, and steep surfaces, as well as zones adjacent to slopes, in order to reduce bias in the collected samples associated with eluviation by rainfall, and transportation of phytolith or other material from higher elevated soils or outcrops (Aleman et al., 2014).
In each of the primary quadrats we excavated ten randomly located, ~15 cm deep soil profiles. From each soil profile, after removing the superficial litter material and organic horizon (O; Fig. 1.2), we collected two samples (~100 g each) corresponding to the upper (3-5cm, $U_{1\text{-}10}$) and lower (5-10 cm, $L_{1\text{-}10}$) section of the soil A-horizon. Two additional samples, U_{mix} and L_{mix} were obtained for each primary quadrat by mixing fractions of 1) all ten U samples, and 2) all ten L samples, respectively (Figure 1.2 illustrates the sampling protocol). In ten (Q2-Q11) additional quadrats in Palo Verde we also excavated ten different soil profiles per quadrat. However, in each quadrat we only collected one single set of larger (100g/sample) U and L soil samples from one of the soil profiles; from the remaining nine soil profiles, we extracted smaller U and L subsamples (~10 grams each). These nine smaller samples, plus two 10-gram subsamples of the 100-gram U_1 and L_1 samples, respectively, were combined in two separate containers, producing a U_{mix} and an L_{mix} sample. We therefore ended up with four samples for each of the ten quadrats (Q2-Q11): one U ($U_{Q2\text{-}11}$) point sample, one L ($L_{Q2\text{-}11}$) point sample, one U_{mix} ($U_{mixQ2\text{-}11}$), and one L_{mix} ($L_{mixQ2\text{-}11}$); the ten quadrats represent replicas.

1.3.3 *Soil sample treatment*

After collection we placed the samples in an oven at 45 °C for 24-48 hours to dry and sterilize them. Soil samples were shipped to the University of Washington (USDA permit P330-14-00285), where they were again dried at 50 °C for three hours to complete sterilization, following USDA regulations. Soil samples and resulting extracted material are reposited in the Burke Museum of Natural History and Culture (UWBM).

1.3.4 *Phytolith extraction*

We homogenized soil samples by crushing them with a hammer and measured out 1 gram per sample for processing. After additional crushing with mortar and pestle, we used standard methods for phytolith extraction from modern soils (Piperno, 1988; Zhao and Pearsall, 1998), which include treatment with hydrochloric acid for carbon removal, potassium hydroxide for humic acid removal, strong acid solution (hydrochloric acid + nitric acid in equal amounts) for removal of oxides, and Schulze's solution for removal of organic material. We removed all particles with a diameter >250 μm by sieving, and disaggregated the remaining soil particles (usually containing a high proportion of clays which tend to adhere to phytolith particles, preventing them from passing through the sieve) by sieving through a smaller diameter (53 μm) sieve. We collected the sieved material (<53 μm in diameter) and recombined it with the material left on the sieve (which includes phytolith >53 μm and <250μm in diameter). We isolated the biogenic silica (sponge spicules, diatoms, chrysophyte cysts, and phytoliths) using a heavy liquid solution of zinc bromide, hydrochloric acid, and water prepared at a specific gravity of 2.38.

1.3.5 *Phytolith identification and counting*

A dried subsample of the extracted silica was mounted on a slide using Cargille MeltMount™ and viewed at 1,000x magnification using a Nikon i80 compound microscope with mounted Nikon DS-Fi1 camera. We counted at least 200 diagnostic phytoliths per sample for statistically significant results (Pearsall, 2000; Strömberg, 2009).

The classification scheme of phytolith morphotypes followed Strömberg et al., (2013), which builds on Strömberg (2003) and later publications, supplemented by other publications (Piperno, 2006; Chen and Smith, 2013). This classification scheme has the advantage of being based on a quantitative analysis of a modern phytolith reference collection, which uses the relative abundance distribution of morphotypes among plant taxa to interpret the plant functional group affinity of each morphotype. These inferred affinities (referred to as "compound variables"; Strömberg, 2003) are used to interpret the composition and structure of phytolith assemblages. For this study, we used the following plant functional group affinities (which combines some of the compound variables used in, e.g., Strömberg et al., 2013): 1) "Palms", 2) "Woody dicots," which includes phytoliths that are commonly produced by woody angiosperms, 3) "Other FI," pertaining to phytoliths produced by woody and herbaceous dicots and/or a broad set of so-called forest indicators (FI) (e.g., conifers, ferns, cycads; Strömberg, 2003, 2005), 4) "Zingiberales"; 5) "Aquatic plants, " including phytoliths typical of sedges, other aquatic monocots, and *Equisetum*; 6) "GSSCPs", corresponding to grass silica short cell phytoliths, which are diagnostic of grasses and commonly used as the only set of morphotypes representing grasses in deep-time phytolith analyses (e.g., Strömberg, 2004; Harris et al., 2017); 7) "GRASS", which includes phytoliths produced by both conifers and monocots and non-GSSCP phytoliths that are diagnostic of grasses; as well as 8) "Other", non-diagnostic or unknown phytoliths. An additional functional

group was created, which encompasses echinate spheres-like morphotypes found at Palo Verde that we were not able to assign to the "Palms" functional group. These phytolith morphologies are somewhat similar to the morphotype Clm-2 ("echinate sphere") described by Strömberg (2003, 2005), but differ either in their more irregular tubercle (echina) shape (including irregular tubercle-like projections), or their smaller size compared to the typical Clm-2 (≤ 5 μm). Identical forms have been described by Chen and Smith (2013) from Marantaceae, Heliconiaceae ("globular-rugulate"), and Zingiberaceae ("globular-microechinate", and "globular-rugulose-granulate") in Order Zingiberales; they were herein classified as "Palms/Zingiberales". Based on phytoliths observed in our La Selva samples, we also added several forest indicator (FI) phytolith morphotypes with diagnostic characters to our classification; these morphotypes were previously described from neotropical forests by other authors (e.g., Piperno, 2006).

To evaluate similarity in vegetation structure, we chose to first focus on the relationship between forest indicators (FI = Palms + Zingiberales + Woody dicots + Other FI + Palm/Zingiberales) and grass phytoliths (GSSCPs) referred to as the *FI-t* ratio (see Strömberg et al., 2007b), and calculated as follows:

$$FI - t = FI_{phytoliths} \Big/ \left(FI + GSSCP \right) \times 100 \qquad (2.1)$$

The *FI-t* ratio is similar to the D/P ratio (sum of the phytoliths produced by tropical ligneous dicotyledons over the sum of Poaceae phytoliths) developed by Alexandre et al., (1997), but differs in that it considers all phytolith morphotypes produced by taxa indicative of closed forest habitats, ("Palms", "Woody dicots", "Other FI", and "Zingiberales", "Palm/Zingiberales"), instead of phytoliths produced by ligneous dicotyledons only—and in that it is a percentage instead of a true ratio (see discussion in Strömberg et al., 2018). Phytolith belonging to the groups "Aquatic plants", "GRASS", and "Other" are not used in the *FI-t* ratio.

1.3.6 *Statistical analysis*

Robustness in the calculated *FI-t* ratio was evaluated by bootstrapping analysis of the phytolith counts, yielding 95% confidence intervals (CIs) (Strömberg, 2009). This technique allows estimation of the precision of the *FI-t* ratio obtained from each sample by resampling randomly and with replacement 10,000 times using the phytolith size and composition of that sample, and it has the advantage of being nearly free from assumptions about the underlying sample universe (Simon and Bruce, 1991). In order to compare samples with each other, we considered the overlap of CIs between samples (or lack thereof). If the CI of a given sample overlaps with that of another sample, we can consider the two samples as likely to be drawn from the same phytolith population (i.e., resulting from standing vegetation with similar structure/openness).

For La Selva, we first compared U and L samples from each soil profile within the quadrat (Q1) to test for differences between the phytolith *FI-t* ratio of the upper versus the lower soil A-horizons, and to test the hypothesis ($H_{1,horizon}$) that phytoliths assemblages in the upper A-horizon and phytolith assemblages in the lower A-horizon do not equally represent the structure of the standing vegetation. We also compared the *FI-t* ratio of all U samples (U_{1-10}) to the mixed U_{mix} sample, and the *FI-t* ratio of all L samples (L_{1-10}) to the mixed L_{mix} sample to test the hypothesis ($H_{1,sample\ type}$) that a soil phytolith assemblage from a single sample and a soil phytolith assemblage from a composite sample are not equally representative of the structure of the standing vegetation.

The same statistical analyses were used for Palo Verde samples. However, the ten U and the ten L soil samples from our first quadrant (Q1) did not yield any phytoliths. Therefore, in

order to test the same hypothesis in Palo Verde, we had to design a different protocol, using the additional samples collected from quadrats Q2-Q11 as replicas (one U, one L, one U_{mix}, and one L_{mix} sample per quadrat, for a total of 40 samples; see Fig. 1.2). Although we recognize that the use of two different protocols for our two study sites might not be ideal, this approach nevertheless provides statistically robust tests of our hypothesis. Specifically, the ten replicated quadrats in Palo Verde allowed us to repeat, ten times, the comparison between the *FI-t* ratio of an upper versus a lower A-horizon sample, and between the *FI-t* ratio of a single versus a composite sample (instead of comparing ten pairs of upper and lower A-horizon samples, and ten single samples with a composite sample from the same quadrat, as done for La Selva). Thus, to test our hypothesis $H_{1,horizon}$ at Palo Verde, we compared (separately) all twenty pairs of upper vs. lower samples (two pairs in each quadrat: U versus L, and U_{mix} versus L_{mix}). To test our hypothesis $H_{1,sample\ type}$ we compared (separately) all twenty pairs of single vs. composite samples (two pairs in each quadrat: U versus U_{mix}, and L versus L_{mix}).

In addition, to investigate differences in morphotype composition of the phytolith assemblages within the La Selva quadrat (Q1), we conducted Non-metric MultiDimensional Scaling (NMDS) on the phytolith counts of all assemblages within Q1. This technique has been shown to outperform other commonly used ordination methods for species abundance datasets (Gauch 1982; Minchin, 1987), and it is one of the most commonly used ordination techniques in community ecology. For example, one advantage of NMDS over other ordination methods is that it emphasizes relative differences (rank distances) between objects rather than absolute differences (actual distances themselves) (Kruskal, 1964). Furthermore, this technique does not require assumptions of linear relationships between variables.

For this analysis we decided to include all counted phytoliths, adding those belonging to the groups “Aquatic plants”, “GRASS”, and “Other” to the phytolith groups shown in Figure 1.3. We used squared-root transformation to reduce skewness in the count data of phytolith groups, and we standardized columns to their totals in order to adjust for large variance among groups. To perform NMDS we further transformed the data count into a Bray-Curtis dissimilarity matrix. In order to choose the number of dimensions to be considered, we examined the plot of stress values versus the number of dimensions obtained using the function *nmds.scree* in R. The stress value for two dimensions was 0.14, indicating a good fit of the ordination to the observed distances (Sturrock and Rocha, 2000). We then performed a Monte Carlo randomization test of the stress values using the *nmds.monte* function in R to further evaluate the accuracy of the ordination with the previously established number (two) of dimensions. This test compares the stress obtained using the original dataset with the stress obtained by reiteration of randomized versions of the original dataset, and calculates a *p-value*, which is the chance of randomly obtaining the given stress value if there was no fit between the ordination distance and the observed dissimilarity. In addition, we analyzed the relationship between the observed dissimilarity (Bray-Curtis) and the ordination space distances (and corresponding R^2 values) using the function *stressplot* in R. To calculate the loadings of the diagnostic groups on the first two axes of the ordination, and their statistical significance, we used the function *envfit* in R. All NMDS functions are contained in the package *vegan* (Oksanen et al., 2018).

For Palo Verde the extremely small sample size (4 samples per quadrat) did not allow us to perform NMDS. In addition, the presence of zeroes in the diagnostic morphotype counts, together with the general data over-dispersion (with some diagnostic groups being typically dominant or typically rare in many samples) did not allow further analysis at the quadrat level

(e.g., Pearson's X^2 test). Therefore, similarities between phytolith assemblages within individual quadrats in Palo Verde were only evaluated in term of the vegetation structure signal (*FI-t* ratio). R statistical software version 3.3.2 (R Core Team, 2017) was used for all ordinations and statistical analyses.

1.4 RESULTS

1.4.1 *Phytolith yield*

All the samples processed from the La Selva quadrat 1 (ten U, ten L, one $_{Umix}$ and one L_{mix}) yielded very abundant and well-preserved phytoliths. In contrast, as mentioned above, the samples processed from the Palo Verde quadrat 1 (ten U, ten L, one $_{Umix}$ and one L_{mix}) only yielded very scarce and poorly preserved phytoliths, preventing further analysis. In general, the quality of phytolith preservation was higher and homogeneous in La Selva, and lower and more variable across Palo Verde quadrats, with some quadrats displaying a much lower level of preservation (Q5, Q9, and Q10) and yield (Q9, and 10) than the others.

1.4.2 *Differences in habitat openness (FI-t ratio) between samples from the upper and lower soil A-horizon*

All upper and lower soil A-horizons samples collected from the single quadrat (Q1) in La Selva are strongly dominated by FI phytoliths, which make up on average 97.13±1.57% (93.61 to 99.76%) of diagnostic morphotypes (FI+GSSCPs) in the soil phytolith assemblages (Fig. 1.3, see A.2.1 and A.2.2 for raw data). The average within-quadrat difference in *FI-t* ratio between U and L samples is 0.96±0.76 percentage points (0.22 to 2.33 percentage points). The bootstrap

analysis produced 95% CI with a relatively narrow range of values (90.64% to 100%; Fig. 1.3) and largely overlapping (see A.2.2 for detailed % FI-t and CIs data).

At Palo Verde, FI phytoliths make up on average 58.07±21.08% (15.02 to 95.85%) of the phytolith assemblages, with some quadrats strongly dominated by FI phytoliths and others dominated by open habitat indicators (Fig. 1.4; see A.1.1 and A.1.3 for raw data). The within-quadrat difference in *FI-t* ratio between U and L samples is on average 6.71±7.12 percentage points, with a maximum difference of 21.32 percentage points (in Q10) and a minimum of 0.86 percentage points (in Q9). The within-quadrat difference in *FI-t* ratio between U_{mix} and L_{mix} samples is on average 6.51±6.89 percentage points, with a maximum of 21.27 percentage points (in Q10), and a minimum of 1.09 percentage points (in Q11). Overall, we found CIs overlap in eight of the ten U-L sample comparisons (quadrats 2, 3, 4, 6, 7, 8, 9, and 11), and in eight of the ten U_{mix}-L_{mix} sample comparisons (quadrats 2, 4, 5, 6, 7, 8, 9, 11) (see A.2.3 for detailed % FI-t and CI data).

1.4.3 *Differences in habitat openness (FI-t ratio) between single and composite samples*

At La Selva, the within-quadrat difference in *FI-t* ratio between individual samples ($U_{1\text{-}10}$, and $L_{1\text{-}10}$) and the composite samples (U_{mix}, and L_{mix}) is on average 1.39±1.23 percentage points (for the upper A-horizon) and 1.66±1.05 percentage points (for the lower A-horizon) with a maximum difference of 3.15 (upper A-horizon) and 3.42 (lower A-horizon) percentage points, and a minimum difference of 0.02 (upper A-horizon) and 0.38 (lower A-horizon) percentage points (Fig. 1.3; see also A.1.2 for detailed % FI-t and CIs data).

At Palo Verde, the within-quadrat difference in *FI-t* ratio between individual and composite samples is on average 9±7.37 ($U_{Q2\text{-}11}$ - $U_{mix\ Q2\text{-}11}$) and 7.08±6.83 ($L_{Q2\text{-}11}$ - $L_{mix\ Q2\text{-}11}$)

percentage points, with a maximum difference of, respectively, 21.46 (Q5) and 22.76 (Q10) percentage points, and a minimum difference of, respectively, 0.76 (Q6) and 1.07 (Q11) percentage points. Overall, we found CIs overlap in seven of the ten U-U_{mix} sample comparisons (quadrats 2, 3, 4, 6, 7 8, and 11), and in eight of the ten L-L_{mix} sample comparisons (quadrats 2, 3, 4, 5, 6, 7, 8, and 11) (Fig. 1.4, see also A.1.3 for detailed % FI-t and CIs data).

1.4.4 *More detailed variation in morphotype composition of phytolith assemblages*

The ordination diagram resulting from the NMDS (Fig. 1.5) illustrates that dissimilarities between the phytolith compositions of different samples within the La Selva rainforest quadrat are mainly driven by difference in the abundance of phytolith morphotypes belonging to "Palms", "Woody dicots", and "Other forest indicators" as well as "Other" (non-diagnostic phytoliths) ($p<0.001$). In Palo Verde, the most abundant phytolith morphotypes are either forest indicators or open habitat grasses. Phytolith morphotypes diagnostic of palms, Zingiberales, and other monocots (also part of the broader FI functional group) were generally scarce.

1.5 DISCUSSION

1.5.1 *Differences in phytolith yield between sites*

Phytoliths are overall better preserved at La Selva compared to Palo Verde, a difference that might be linked to soil conditions in the two different ecosystems. Furthermore, within Palo Verde dry forest, both the underlying rock type and the degree and type of soil development are variable. In particular, we observed the presence of calcium carbonate detritus likely coming from the rock matrix or the overlying limestone outcrops. This might result in overall more alkaline soil conditions, known to accelerate phytolith dissolution (e.g., Bartoli and Wilding,

1980; Fraysse et al., 2006). It is therefore possible that soil alkalinity has negatively affected the preservation of soil phytolith assemblages in some parts of the dry forest, such as the unproductive Palo Verde quadrat (in which $CaCO_3$ detritus was particularly abundant). In addition, we note that the assemblages with the highest proportion of GSSCPs were recovered from forest-covered quadrats (2-5) located on Eocene marine limestones characterized by poor soil development (Gillespie et al., 2000). This is consistent with studies showing that the preservational bias for GSSCPs and against the more fragile and easily dissolvable FI phytolith morphotypes is especially pronounced at high pH (Wilding and Drees, 1974; Bartoli and Wilding, 1980; Cabanes et al., 2011).

1.5.2 *Habitat openness signal: forest indicators (FI) to grass phytolith ratio*

1.5.2.1 Lower versus upper soil A-horizons

At La Selva, the lack of significant difference between the *FI-t* ratio of the phytolith assemblages from the upper and lower soil A-horizons leads to effectively identical interpretations of the standing vegetation (closed-canopy) for all samples (Fig. 1.3). Therefore, we cannot reject our null hypothesis $H_{0,horizon}$ that the two soil layers equally represent the structure of the standing vegetation. The implication of this outcome is that, in theory, samples from any part of the A horizon can be used and are comparable.

The similarity between the upper- and lower A-horizon in the studied soils can be explained by a) time-averaging (e.g., through bioturbation) acting to homogenize the upper and lower A-horizons; b) relative temporal stability and homogeneity of the rainforest ecosystem in term of floral composition and/or structure (at the level discernable in the phytolith analysis); or c) both. Vigorous bioturbation, characteristic of tropical Oxisols (Weil and Brady, 2016) can be

expected to be an important process in the soils at La Selva. Combined with vertical translocation (due to illuviation, and dissolution,) this mixing of soil layers is likely to lead to time-averaging. On the other hand, because of the short turn-over time of the soil phytolith pool in wet and warm environments, phytolith assemblages from La Selva should record only the last few hundred years of vegetation (Song et al., 2016; Strömberg et al., 2018). The age of the oldest tree in La Selva primary forest was estimated to be 309 years in 1982 (Lieberman and Lieberman, 1985), suggesting that this forest is at least a few hundred years old. The prevalence of FI phytolith morphotypes in all of the assemblages, together with the age of La Selva primary forest supports the idea that the soil A-horizon at La Selva reflects a relatively stable rainforest habitat. Further investigation to directly measure time averaging should include carbon dating of the soil samples (such as in McClaran and Umlauf, 2000); however, this is well beyond the scope of our study.

At Palo Verde, the *FI-t* ratio of the phytolith assemblages from upper and lower soil A-horizons are nearly identical in the majority of the quadrats, and in the quadrats with non-overlapping CIs the differences do not fundamentally change the interpretation of the vegetation structure within single quadrats (Fig. 1.4). For example, in quadrat 5, both U_{Q5} and L_{Q5} samples are dominated by forest indicators (representing, respectively, 66.54 and 84.45%) and can be interpreted as reflecting predominantly wooded vegetation with a relatively important grass component. Therefore, like at La Selva, we cannot reject $H_{0,horizon}$ that the upper and lower A-horizon similarly record the vegetation structure.

We further propose that within-quadrat differences in Palo Verde can be explained by biases related to a) phytolith production, and b) other taphonomic processes, perhaps combined with (c) non-negligible vegetation heterogeneity in time and space. Unlike at La Selva, the

woody species at Palo Verde are not all consistently high-silica producers (e.g., Fabaceae typically contains very little silica; Hodson et al., 2005; Strömberg et al., 2016). Because the *FI-t* ratio in effect measures the relative abundance of grass phytoliths, it can potentially underestimate woody cover in parts of the landscape where the woody component consists of poor silica accumulators (see discussion in Strömberg et al., 2018). Also, bioturbation is less likely to be important in the drier soils at Palo Verde (Powers et al., 2009), potentially resulting in less thorough mixing of the A-horizon. Palo Verde is known as a fire-prone habitat (Powers et al., 2009), making it likely that the distribution of trees has changed over time as an outcome of ecological succession following fire. Given the potential for less time-averaging (than in La Selva), and the aforementioned differences in silica production among trees contributing to the soil phytolith assemblages at Palo Verde, it is possible that such shifts in the distribution of trees could result in differences between L and U samples in some quadrats. As for La Selva, our data from the Palo Verde dry forest do not allow us to assess the presence and magnitude of time averaging (or controlling factors), but they indicate that soil phytolith assemblages from the upper and lower A-horizon within individual quadrats nonetheless represent substantially the same vegetation structure.

1.5.2.2 Single versus composite soil samples

At La Selva, the *FI-t* ratios of single and composite samples (from both upper and lower A-horizons), do not differ significantly (Fig. 1.3); thus, we cannot reject our null hypothesis $H_{0,\text{sample type}}$ that phytolith assemblages from single and composite samples equally represent the structure of the standing vegetation. That is, each soil sample provides a fair representation of the vegetation structure at the quadrat level in this rainforest environment, despite the fact that

minimal lateral mixing (Piperno, 1988) is commonly thought to result in a 'hyper-local' phytolith signal.

At Palo Verde, we similarly found no significant differences between composite and single samples in all but two U-L sample comparisons and three U_{mix}-L_{mix} comparisons (Fig. 1.4). In these latter cases, differences between composite and single samples are small enough (~10 to ~23% percent points) to have negligible effect on the interpretation of the structure of the standing vegetation. We suggest that a higher degree of lateral variation in phytolith transport and soil processes at Palo Verde dry forest compared to La Selva rainforest result in overall greater variation between soil phytolith assemblages (measured as the *FI-t* ratio) at the quadrat level. Thus, in this less dense forest environment, some local or even extra-local phytolith transport can be expected to occur (e.g., by wind, water, fire, or herbivores). The effect of phytolith lateral transport can be the homogenization of all assemblages within a quadrat, rendering them more similar to each other. However, lateral transport could also provide extra local input to some but not all parts of a quadrat, especially if there is variability in vegetation openness. Added to spatial differences in the degree of time averaging in the soil (see above), such variation might lead to within-quadrat disparity in phytolith assemblage composition.

1.5.3 *More detailed variation in morphotype composition of phytolith assemblages*

Within La Selva quadrat (Q1), phytolith morphotype composition differs between samples in term of relative abundaces of specific FI components, as shown by the NMDS (see axis loadings on Figure 1.1). However, variation at such a small spatial scale (10 x 10 meters quadrat) is unlikely linked to major vegetational or environmental gradients, and therefore cannot be considered informative. In addition, the presence of relatively less abundant

morphotypes (such as Zingiberales) is captured by some but not all samples (e.g., see Zingiberales in Figure 1.3). This suggests that 'hyper-local' components of the vegetation can be recorded by some phytolith assemblages in this rainforest soil, but that their presence can be obscured by the most abundant morphotypes. As this seems to be simply related to sampling error, we suggest that a greater number (>>200) of diagnostic phytolith morphotypes should be counted from an individual sample in order to account for rare morphotypes. Alternatively, the already common practice of additional scanning (Pearsall, 2000; Strömberg, 2009) of sample slides should be implemented.

1.5.4 *Future sampling strategy*

Our results allow us to define guidelines for future modern analogue studies based on the *FI-t* ratio in both neotropical rainforest and dry forest habitats. We propose that the most rigorous methodology for modern-analogue studies with application in the deep-time fossil record is to collect lower-A horizon samples, as this portion of the soil is more likely to be preserved in the fossil record than the upper portion. However, other factors, such as land use history, should be taken into account when making this choice. For instance, in a previously disturbed, recently restored habitat, the upper A-horizon is more suitable, as its phytoliths should represent a later successional stage compared to the lower A-horizon.

We argue further that combining multiple samples to obtain a phytolith assemblage that better represents the standing vegetation is unnecessary for the purpose of reconstructing vegetation structure (*FI-t* ratio), at least at the level of detail that is commonly used in deep-time work (e.g., Strömberg et al., 2013; Smiley et al., 2018). In these studies, the *FI-t* ratio is used primarily to infer relative changes in vegetation structure rather than precise, absolute

reconstructions of the vegetation (see discussion in Strömberg et al., 2018). Finally, we suggest additional scanning of sample slides or counting of a greater number (>>200) of phytoliths per slide to account for the presence of rare morphotypes, in particular those diagnostic of plant functional groups that provide important clues for paleoenvironmental reconstructions (e.g.,Zingiberales, aquatic plants).

1.5.5 *Differences between vegetation types*

Although soil sample depth (U or L) and type (single, or composite) do not affect the overall interpretation of the vegetation structure from phytolith assemblages from either vegetation type (rainforest and dry forest), rainforest soil phytoliths do seem to represent a better preserved and more accurate record of the standing vegetation than dry forest phytolith assemblages. For instance, at Palo Verde some morphotypes, namely GSSCPs, and to a lesser extent palm phytoliths, appear to be overrepresented overall, resulting in potentially inaccurate reconstructions of the surrounding vegetation structure.

Several factors may contribute to the overall poorer representation in Palo Verde phytolith assemblages. These factors include preservation, production, and dispersal biases (see discussion above) which may be augmented at Palo Verde because of the greater heterogeneity in soil and environmental conditions, the greater spatial and temporal heterogeneity in vegetation, and the prominence of fire and wind as dispersal agents in this ecosystem.

Further study at the transect level (including comparison of phytolith assemblages with the plant communities) is needed to tease apart the respective effects of taphonomy and spatial heterogeneity of vegetation on soil phytolith assemblage composition. In addition, the phytolith composition of soil assemblages should be analyzed using a higher taxonomic/morphotype resolution to improve vegetation inferences. For instance, GSSCPs and phytoliths produced by

either palms (family Arecaceae) or families of order Zingiberales (classified here as Palm/Zingiberales) warrant more detailed study to resolve their taxonomic affinities. Similarly, the FI functional group should be broken up into several sub-groups in order to better link phytolith assemblages composition to plant community composition along vegetation transects.

1.6 CONCLUSION

This study seeks to establish a standardized protocol for collecting soil samples in modern analogue phytolith work, to make it maximally comparable to sampling in the deep-time geologic record. To do so, we investigated in two different vegetation types (rainforest and dry forest) in Costa Rica whether to sample from the upper or lower soil A-horizon and whether to use single point samples or combining multiple point samples into a mixed sample. Our analysis showed that single samples from the lower A-horizon—most comparable to samples taken in pre-Quaternary paleosols—yield phytolith assemblages adequately representing the soil phytolith content at the quadrat scale. Further, our data suggest that mixing several soil samples from a small area is unnecessary, at least for reconstructing vegetation structure using the *FI-t* ratio.

These results not only provide useful guidelines for future modern analogue work, but also largely validate the sampling approach employed in many paleontological studies using phytoliths, which is to rely on point samples from the lower A-horizon (or at an unknown level in the A-horizon. Recent work comparing fossil phytolith assemblages from A- and B-horizons failed to reveal differences that can be solely linked to known taphonomical processes within the soil perhaps indicating that B-horizon assemblages may, at least in some cases also preserve meaningful paleo-vegetation information. However, these suggestions remain to be rigorously tested in modern soils.

These promising results notwithstanding, much work is needed to both improve the resolution of the phytolith signal at larger spatial scales (hundreds of meters) within rainforest and dry forest (studied herein), and test the applicability of our results in other habitat types, such as open and mixed vegetation (e.g., savannas and shrublands), in which the factors affecting the resolution of the soil phytolith record (e.g., climate, soil type, fire regime, herbivore behavior) are very different. It is worth noting that many of the Cenozoic phytolith assemblages studied so far are interpreted as representing grass-dominated habitats or non-analog shrublands, that is, vegetation types that have not yet been examined in detail. Thus, expanding the (bio)geographic and climatic scope of this modern analogue work is vital for our ability to interpret the deep-time fossil phytolith record both precisely and accurately, adding confidence to our reconstructions of past ecosystem history.

ACKNOWLEDGMENTS

We are grateful to the staff of the Organization for Tropical Studies, La Selva and Palo Verde Biological stations for their help with field work logistics; MINAE and Francisco Campos for their help with collection and export permits, the Burke Museum of Natural History and Culture (UWBM) and UWBM Botany Collections (WTU herbarium) at the University of Washington (Seattle), and the Herbario Nacional de Costa Rica (San José). We also thank Claire Grant and Kailyn Zard, who helped with sample processing, Alex T. Lowe, Jennifer Hsiao, Nicholas Waldo, and Janneke Hille Ris Lambers, and two anonymous reviewers for comments on an earlier version of this manuscript. Fieldwork and sample processing were supported by the Organization for Tropical Studies, WRF-Hall Fellowship (University of Washington), Philosophical Society of America, Geological Society of America, Paleontological Society, and

Quaternary Research Center (grants awarded to C.C.), and the National Science Foundation (NSF EAR-1253713 and EAR-1349530 awarded to C.A.E.S).

REFERENCES

Aleman, J., Leys, B., Apema, R., Bentaleb, I., Dubois, M.A., Lamba, B., Lebamba, J., Martin, C., Ngomanda, A., Truc, L., Yangakola, J.-M., Favier, C., Bremond, L., Woods, K., 2012. Reconstructing savanna tree cover from pollen, phytoliths and stable carbon isotopes. Journal of Vegetation Science 23, 187-197.

Aleman, J.C., Canal-Subitani, S., Favier, C., Bremond, L., 2014. Influence of the local environment on lacustrine sedimentary phytolith records. Palaeogeography, Palaeoclimatology, Palaeoecology 414, 273-283.

Alexandre, A., Meunier, J.-D., Colin, F., Koud, J.-M., 1997. Plant impact on the biogeochemical cycle of silicon and related weathering processes. Geochimica et Cosmochimica Acta 61, 677-682.

Alexandre, A., Bouvet, M., Abbadie, L., 2011. The role of savannas in the terrestrial Si cycle: A case-study from Lamto, Ivory Coast. Global and Planetary Change 78, 162-169.

Alvarado Induni, G.E., 1990. Características geólogicas de la Estación Biológica L Selva, Costa Rica. Tecnología en marcha 10, 11-22

Bartoli, F., Wilding, L.P., 1980. Dissolution of biogenic opal as a function of its physical and chemical properties. Soil Science Society of America Journal 44, 873-878.

Beavers, A., Stephen, I., 1958. Some features of the distribution of plant-opal in Illinois soils. Soil Science 86, 1-5.

Blecker, S.W., McCulley, R.L., Chadwick, O.A., Kelly, E.F., 2006. Biologic cycling of silica across a grassland bioclimosequence. Glob. Biogeochem. Cycle 20, 1-11.

Bremond, L., Alexandre, A., Hély, C., Guiot, J., 2005. 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., Bodin, S.C., Bentaleb, I., Favier, C., Canal, S., 2017. Past tree cover of the Congo Basin recovered by phytoliths and ä13C along soil profiles. Quaternary International 434, 91-101.

Cabanes, D., Weiner, S., Shahack-Gross, R., 2011. Stability of phytoliths in the archaeological record: a dissolution study of modern and fossil phytoliths. Journal of Archaeological Science 38, 2480-2490.

Carnelli, A., 2001. Biogenic Silica Production in Selected Alpine Plant Species and Plant Communities. Annals of Botany 87, 425-434.

Chen, S.T., Smith, S.Y., 2013. Phytolith variability in Zingiberales: A tool for the reconstruction of past tropical vegetation. Palaeogeography, Palaeoclimatology, Palaeoecology 370, 1-12.

Chen, S.T., Smith, S.Y., Sheldon, N.D., Strömberg, C.A.E., 2015. Regional-scale variability in the spread of grasslands in the late Miocene. Palaeogeography, Palaeoclimatology, Palaeoecology 437, 42-52.

Ciochon, R.L., Piperno, D.R., Thompson, R.G., 1990. Opal phytoliths found on the teeth of the extinct ape Gigantopithecus blacki: implications for paleodietary studies. Proceedings of the National Academy of Sciences 87, 8120-8124.

Cotton, J.M., Sheldon, N.D., Strömberg, C.A., 2012. High-resolution isotopic record of C4 photosynthesis in a Miocene grassland. Palaeogeography, Palaeoclimatology, Palaeoecology 337, 88-98.

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

Ehrlich, H., Demadis, K.D., Pokrovsky, O.S., Koutsoukos, P.G., 2010. Modern Views on Desilicification: Biosilica and Abiotic Silica Dissolution in Natural and Artificial Environments. Chemical Reviews 110, 4656-4689.

Fishkis, O., Ingwersen, J., Streck, T., 2009. Phytolith transport in sandy sediment: Experiments and modeling. Geoderma 151, 168-178.

Fishkis, O., Ingwersen, J., Lamers, M., Denysenko, D., Streck, T., 2010a. Phytolith transport in soil: A field study using fluorescent labelling. Geoderma 157, 27-36.

Fishkis, O., Ingwersen, J., Lamers, M., Denysenko, D., Streck, T., 2010b. Phytolith transport in soil: a laboratory study on intact soil cores. European Journal of Soil Science 61, 445-455.

Folger, D.W., 1970. Wind transport of land-derived mineral, biogenic, and industrial matter over the North Atlantic. Deep Sea Research and Oceanographic Abstracts 17, 337-352.

Folger, D.W., Burckle, L.H., Heezen, B.C., 1967. Opal Phytoliths in a North Atlantic Dust Fall. Science 155, 1243.

Fraysse, F., Pokrovsky, O.S., Schott, J., Meunier, J.-D., 2006. Surface properties, solubility and dissolution kinetics of bamboo phytoliths. Geochimica et Cosmochimica Acta 70, 1939-1951.

Fredlund, G.G., Tieszen, L.T., 1994. Modern Phytolith Assemblages from the North American Great Plains. Journal of Biogeography 21, 321-335. https://doi.org/10.2307/2845533

Gauch, H.G., Gauch Jr, H.G., 1982. Multivariate analysis in community ecology. Cambridge University Press.

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

Harris, E.B., Strömberg, C.A., Sheldon, N.D., Smith, S.Y., Vilhena, D.A., 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.

Henry, A.G., Brooks, A.S., Piperno, D.R., 2011. Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets (Shanidar III, Iraq; Spy I and II, Belgium). Proceedings of the National Academy of Sciences 108, 486-491.

Henry, A.G., Ungar, P.S., Passey, B.H., Sponheimer, M., Rossouw, L., Bamford, M., Sandberg, P., de Ruiter, D.J., Berger, L., 2012. The diet of Australopithecus sediba. Nature 487, 90-93.

Hodson, M.J., White, P.J., Mead, A., Broadley, M.R., 2005. Phylogenetic Variation in the Silicon Composition of Plants. Annals of Botany 96, 1027-1046.

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

Iriarte, J., 2003. Assessing the feasibility of identifying maize through the analysis of cross-shaped size and three-dimensional morphology of phytoliths in the grasslands of southeastern South America. Journal of Archaeological Science 30, 1085-1094.

Jones, L.H.P., Handreck, K.A., 1967. Silica In Soils, Plants, and Animals, in: Norman, A.G. (Ed.), Advances in Agronomy. Academic Press, pp. 107-149.

Katz, O., 2015. Silica phytoliths in angiosperms: phylogeny and early evolutionary history. New Phytologist 208, 642-646.

Kealhofer, L., 1996. The human environment during the terminal Pleistocene and Holocene in northeastern Thailand: phytolith evidence from Lake Kumphawapi. Asian Perspectives, 229-254.

Kealhofer, L., Penny, D., 1998. A combined pollen and phytolith record for fourteen thousand years of vegetation change in northeastern Thailand. Review of Palaeobotany and Palynology 103, 83-93.

Kelly, E.F., Amundson, R.G., Marino, B.D., Deniro, M.J., 1991. Stable isotope ratios of carbon in phytoliths as a quantitative method of monitoring vegetation and climate change. Quaternary Research 35, 222-233.

Kerns, B.K., Moore, M.M., Hart, S.C., 2001. Estimating forest-grassland dynamics using soil phytolith assemblages and d13C of soil organic matter. Écoscience 8, 478-488.

Kleber, M., Schwendenmann, L., Veldkamp, E., Rößner, J., Jahn, R., 2007. Halloysite versus gibbsite: Silicon cycling as a pedogenetic process in two lowland neotropical rain forest soils of La Selva, Costa Rica. Geoderma 138, 1-11.

Kohlmann, B., Roderus, D., Elle, O., Solís, Á., Soto, X., Russo, R., 2010. Biodiversity conservation in Costa Rica: a correspondence analysis between identified biodiversity hotspots (Araceae, Arecaceae, Bromeliaceae, and Scarabaeinae) and conservation priority lifezones. Revista Mexicana de Biodiversidad 81, 511-559.

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

Lieberman, D., Lieberman, M., Hartshorn, G., Peralta, R., 1985. Growth Rates and Age-Size Relationships of Tropical Wet Forest Trees in Costa Rica. Journal of Tropical Ecology 1, 97-109.

Locker, S., Martini, E., 1986. Phytoliths from the southwest Pacific Site 591. Initial reports of the Deep Sea Drilling Program 90, 1079-1084.

Madella, M., Lancelotti, C., 2012. Taphonomy and phytoliths: A user manual. Quaternary International 275, 76-83.

McClaran, M.P., Umlauf, M., 2000. Desert grassland dynamics estimated from carbon isotopes in grass phytoliths and soil organic matter. Journal of Vegetation Science 11, 71-76.

McMichael, C.H., Bush, M.B., Piperno, D.R., Silman, M.R., Zimmerman, A.R., Anderson, C., 2012. Spatial and temporal scales of pre-Columbian disturbance associated with western Amazonian lakes. The Holocene 22, 131-141.

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

Merceron, G., Novello, A., Scott, R.S., 2016. Paleoenvironments inferred from phytoliths and dental microwear texture analyses of meso-herbivores. Geobios 49, 135-146.

Meunier, J.D., Colin, F., Alarcon, C., 1999. Biogenic silica storage in soils. Geology 27, 835-838.

Miller, L.A., Smith, S.Y., Sheldon, N.D., Strömberg, C.A.E., 2012. Eocene vegetation and ecosystem fluctuations inferred from a high-resolution phytolith record. Geol. Soc. Am. Bull. 124, 1577-1589.

Minchin, P.R., 1987. An evaluation of the relative robustness of techniques for ecological ordination, Theory and models in vegetation science. Springer, pp. 89-107.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2018. Package 'vegan'. Community ecology package, version 2.5-3.

Osterrieth, M., Madella, M., Zurro, D., Fernanda Alvarez, M., 2009. Taphonomical aspects of silica phytoliths in the loess sediments of the Argentinean Pampas. Quaternary International 193, 70-79.

Pearsall, D., 2000. Paleoethnobotany: A Handbook of Procedures Academic Press. San Diego, Calif.

Piperno, D.R., 1988. Phytolyth analysis: an archaeological and geological perspective. Academic Press, San Diego.

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

Piperno, D.R., 2009. Identifying crop plants with phytoliths (and starch grains) in Central and South America: a review and an update of the evidence. Quaternary international 193, 146-159.

Piperno, D.R., Becker, P., 1996. Vegetational History of a Site in the Central Amazon Basin Derived from Phytolith and Charcoal Records from Natural Soils. Quaternary Research 45, 202-209.

Postek, M.T., 1981. The occurrence of silica in the leaves of Magnolia grandiflora L. Botanical Gazette 142, 124-134.

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

Prasad, V., Strömberg, C.A.E., Alimohammadian, H., Sahni, A., 2005. Dinosaur coprolites and the early evolution of grasses and grazers. Science 310, 1177-1180.

Prentice, A.J., Webb, E.A., 2016. The effect of progressive dissolution on the oxygen and silicon

isotope composition of opal-A phytoliths: Implications for palaeoenvironmental reconstruction. Palaeogeography, Palaeoclimatology, Palaeoecology 453, 42-51.

Prentice, I.C., 1985. Pollen Representation, Source Area, and Basin Size: Toward a Unified Theory of Pollen Analysis. Quaternary Research 23, 76-86.

Prentice, I,C., Guiot, J., Huntley, B., Jolly, D., Cheddadi, R., 1996. Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka. Climate Dynamics 12, 185-194.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rosen, A.M., Weiner, S., 1994. Identifying ancient irrigation: a new method using opaline phytoliths from emmer wheat. Journal of archaeological Science 21, 125-132.

Runge, F., 1999. The opal phytolith inventory of soils in central Africa —quantities, shapes, classification, and spectra. Review of Palaeobotany and Palynology 107, 23-53.

Simon, J.L., Bruce, P., 1991. Resampling: A tool for everyday statistical work. Chance 4, 22-32.

Smiley, T.M., Hyland, E.G., Cotton, J.M., Reynolds, R.E., 2018. Evidence of early C4 grasses, habitat heterogeneity, and faunal response during the Miocene Climatic Optimum in the Mojave Region. Palaeogeography, Palaeoclimatology, Palaeoecology 490, 415-430.

Song, Z., McGrouther, K., Wang, H., 2016. Occurrence, turnover and carbon sequestration potential of phytoliths in terrestrial ecosystems. Earth-Science Reviews 158, 19-30.

Strömberg, C.A.E., 2002. The origin and spread of grass-dominated ecosystems in the late Tertiary of North America: preliminary results concerning the evolution of hypsodonty. Palaeogeography Palaeoclimatology Palaeoecology 177, 59-75.

Strömberg, C.A.E., 2003. The origin and spread of grass-dominated ecosystems during the

Tertiary of North America and how it relates to the evolution of hypsodonty in equids. University of California, Berkeley.

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

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 of the United States of America 102, 11980-11984.

Strömberg, C.A.E., 2006. Evolution of hypsodonty in equids: testing a hypothesis of adaptation. Paleobiology 32, 236-258.

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., Werdelin, L., Friis, E.M., Sarac, G., 2007a. 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., Friis, E.M., Liang, M.-M., Werdelin, L., Zhang, Y.-L., 2007. 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., McInerney, F.A., 2011. The Neogene transition from C-3 to C-4 grasslands in North America: assemblage analysis of fossil phytoliths. Paleobiology 37, 50-71.

Strömberg, C.A.E., Dunn, R.E., Madden, R.H., Kohn, M.J., Carlini, A.A., 2013. Decoupling the spread of grasslands from the evolution of grazer-type herbivores in South America. Nature communications 4, 1478-1478.

Strömberg, C.A.E., Di Stilio, V.S., Song, Z., 2016. Functions of phytoliths in vascular plants: an evolutionary perspective. Functional Ecology 30, 1286-1297.

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

Sturrock, K., Rocha, J., 2000. A multidimensional scaling stress evaluation table. Field methods 12, 49-60.

Tabor, N.J., Myers, T.S., 2015. Paleosols as indicators of paleoenvironment and paleoclimate. Annual Review of Earth and Planetary Sciences 43, 333-361.

Thorn, V., 2001. Oligocene and early Miocene phytolits from CRP-2/2A and CRP-3, Victoria Land Basin, Antarctica. Terra Antartica 8, 407-422.

Trembath-Reichert, E., Wilson, J.P., McGlynn, S.E., Fischer, W.W., 2015. Four hundred million years of silica biomineralization in land plants. Proceedings of the National Academy of Sciences 112, 5449-5454.

Tromp, M., Dudgeon, J.V., 2015. Differentiating dietary and non-dietary microfossils extracted from human dental calculus: the importance of sweet potato to ancient diet on Rapa Nui. Journal of Archaeological Science 54, 54-63.

Vivanco, L., Austin, A.T., 2006. Intrinsic effects of species on leaf litter and root decomposition: a comparison of temperate grasses from North and South America. Oecologia 150, 97-107.

Wallis, L.A., 2001. Environmental history of northwest Australia based on phytolith analysis at Carpenter's Gap 1. Quaternary International 83-85, 103-117.

Weil, R., Brady, N.C., 2016. The nature and properties of soils: Pearson new international edition. Pearson Higher Ed.

White, A.F., Vivit, D.V., Schulz, M.S., Bullen, T.D., Evett, R.R., Aagarwal, J., 2012. Biogenic and pedogenic controls on Si distributions and cycling in grasslands of the Santa Cruz soil chronosequence, California. Geochimica et Cosmochimica Acta 94, 72-94.

Whitmore, J., Gajewski, K., Sawada, M., Williams, J.W., Shuman, B., Bartlein, P.J., Minckley, T., Viau, A.E., Webb, T., Shafer, S., Anderson, P., Brubaker, L., 2005. Modern pollen data from North America and Greenland for multi-scale paleoenvironmental applications. Quaternary Science Reviews 24, 1828-1848.

Wilding, L.P., Drees, L.R., 1973. Scanning Electron Microscopy of Opaque Opaline Forms Isolated from Forest Soils in Ohio1. Soil Science Society of America Journal 37, 647-650.

Wilding, L.P., Drees, L.R., 1974. Contributions of forest opal and associated crystalline phases to fine silt and clay fractions of soils. Clays and Clay Minerals 22, 295-306.

Table 1.1. Potential biases in phytolith assemblages at la Selva and Palo Verde.

Factors (column 1) of potential bias (column 2) in modern soil phytolith assemblages, and expected effects on the soil phytolith assemblages of La Selva rainforest (column 3) and Palo Verde dry forest (column 4). References are listed in column 5.

Factors	Potential bias	Predicted effect at La Selva rainforest	Predicted effect at Palo Verde dry forest	References
Production	a) Higher in monocotyledons (in particular grasses and palms), ferns, and horsetails	a) Palms may be significantly overrepresented	a) Both open- and closed-habitat grasses may be overrepresented; palms (rare in the vegetation) may be overrepresented (without altering the overall FI-t proportion)	Hodson et al., 2005; Katz, 2015; Piperno, 2006; Strömberg et al., 2016; Trembath-Reichert et al., 2015
	b) Higher in some angiosperm clades (especially tropical trees and herbs) than others	b) Some tropical trees may be significantly underrepresented	b) Some tropical trees (dominant in the vegetation) may be significantly underrepresented (affecting the overall FI-t proportion)	
	* OVERALL BIAS: Open habitat indicators over FI, as well as palms over other FI	* OVERALL EFFECT: no effect (in term of *FI-t* ratio) but FI may be dominated by palms	* OVERALL EFFECT: Soil phytolith assemblages may appear to reflect a more open habitat than actually present (in term of *FI-t* ratio)	
Transport	a) Eolian transport is enhanced by fire and can result in phytolith transport over hundreds of kilometers	a) Little transport by wind in closed forest habitats	a) Stronger eolian transport in more open habitats and more frequent fires in seasonally dry habitats might result in a more mixed vegetation signal in soil phytolith assemblages	Aleman et al., 2012; Bremond et al., 2005; Folger et al., 1967, 1970; Fredlund and Tieszen, 1994; Locker and Martini, 1986; Osterrieth et al.,

	b) Transport by herbivores, and fluvial/colluvial transport can occur	b) Fluvial/lacustrine settings, and surfaces exposed to enhanced erosion, as well as pastures were avoided in this study. Limited influence on soil phytolith assemblages is expected	b) Fluvial/lacustrine settings, as well as pastures were avoided in this study.	2009; Piperno, 1988, 2006; Wallis, 2001
	* OVERALL BIAS: In open habitats where wind, fire, herbivores, and erosion are more important, phytolith assemblages are more (regionally) mixed	* OVERALL EFFECT: Little bias due to transport is expected	* OVERALL EFFECT: Some bias is expected, primarily due to eolian transport facilitated by fire	
Inheritance & preservation	a) Shorter turnover time of the phytolith stable pool in the soil in wetter and warmer climates.	a) High turnover time	a) Unknown turnover time	Alexandre et al., 1997; Blecker et al., 2006; Ehrlich et al., 2010; Fraysse et al., 2006; Prentice and Webb, 2016; Song et al., 2016; Vivanco and Austin, 2006; Weil and Brady, 2016 White et al., 2012
	b) Acidic conditions (pH<8.5) favor phytolith preservation	b) High precipitations, microbiological activity, plant growth, and root respiration decrease soil pH	b) Weathering of calcium carbonate transported from nearby limestone outcrops, and seasonal aridity increase soil pH, accelerating phytolith dissolution	
	c) Warm and wet climates favor bioturbation	c) Bioturbation results in well-mixed A-horizon	c) Strong precipitation seasonality slows down bioturbation; as a result, illuviation processes dominate	

	OVERALL BIAS: Higher time averaging and preferential preservation of resistant phytolith morphotypes in warm and wet climates, and acidic soils	OVERALL EFFECT: Little time averaging, phytoliths preservation should be optimal, well-mixed A-horizon	OVERALL EFFECT: Unknown time averaging, poor preservation of fragile (FI) morphotypes due to erosion and dissolution, overrepresentation of resistant morphotypes (GSSCPs, palms); upper and lower A-horizons less well mixed	
Size/ shape/ texture	a) Smaller morphotypes (such as GSSCPs, and globular echinates) might be preferentially transported by wind	a) Wind transport is minimal; GSSCPs are unlikely to be transported over long distances	a) GSSCPs and palms might be preferentially transported by wind and fire from surrounding open areas and within the forest	Bartoli and Wilding, 1980; Cabanes et al., 2011; Jones and Handreck, 1967; Piperno, 1988; Wilding and Drees, 1973, 1974
	b) Thin, porous, flattened, and large phytolith morphotypes are more susceptible to dissolution compared to GSSCPs	b) Dissolution mainly affects non-palm FI morphotypes, whereas GSSCPs, palm phytoliths, and globular FI morphotypes are less affected	b) Dissolution mainly affects non-palm FI morphotypes, whereas GSSCPs, palm phytoliths, and globular FI morphotypes are less affected	
	* OVERALL BIAS: GSSCPs, palm phytoliths, and globular FI morphotypes are more likely incorporated and preserved in the soil	* OVERALL EFFECT: Potential overrepresentation of palm and other globular FI morphotypes, but no overrepresentation of GSSCPs due to low relative abundance)	* OVERALL EFFECT: Potential overrepresentation of GSSCPs (and palm phytoliths and globular FI morphotypes) in phytolith assemblages	

Table 1.2. Selected modern analogue phytolith studies showing existing difference in sampling methodology.

Reference	Vegetation type	Region	Sample depth	Samples	Plots	Transects	Phytoliths-vegetation relationship
Piperno, 1988	Semideciduous tropical forest	Tropical, (Panama)	Surface (1 cm)	25 samples, each one consisting of 5-7 trough-fills from a 5x5 m area next to a transect	25 25 m^2 plots	2 transects (320 and 180 m)	Strong local signal of the phytolith assemblages
Fredlund & Tieszen, 1994	Grassland	Temperate (North American Great Plains, USA)	Surface (3-5 cm)	50 samples, each one consisting of 4 subsamples from the four corners of each plot	50 1 m^2 plots	NA	Lack of sensitivity to small-scale local vegetation variability
Kerns *et al*., 2001	*Pinus ponderosa* forest	Temperate (Arizona, USA)	Surface (0-2 cm) subsurface (2-7 cm)	30 samples, 1 composite surface sample, and 1 composite subsurface sample from ten cores in each plot	15 circular 40 m^2 plots	NA	Surface samples weakly reflect modern vegetation; subsurface samples do not reflect modern vegetation
Bremond *et al*., 2005	Savanna/forest transition	Tropical (Cameroon)	Litter or soil (NA)	26 samples, each one consisting of 20-30 randomly collected subsamples	26, 1, and 1 plots of 20x10, 10x1, and 15x10 m^2 respectively	1 transect (750 m)	Non-linear relationship between D/P (ligneous dicots to Poaceae phytoliths ratio) and LAI (leaf area index)
Mercader *et al*., 2011	Miombo woodlands	Tropical (Mozambique)	Surface (1-2 cm)	25 samples	25 1 m^2 quadrats	1 transect (50 km)	Phytolith analysis is able to detect shift from woodland to grassland vegetation
Aleman *et al*., 2012	Savanna	Tropical (Senegal, Congo)	Surface (NA)	17 samples consisting of 20 subsamples randomly collected along transects (10 meters from transect center)	NA	3 transects (120, 120, and 140 m)	Strong relationship between D/P index and LAI

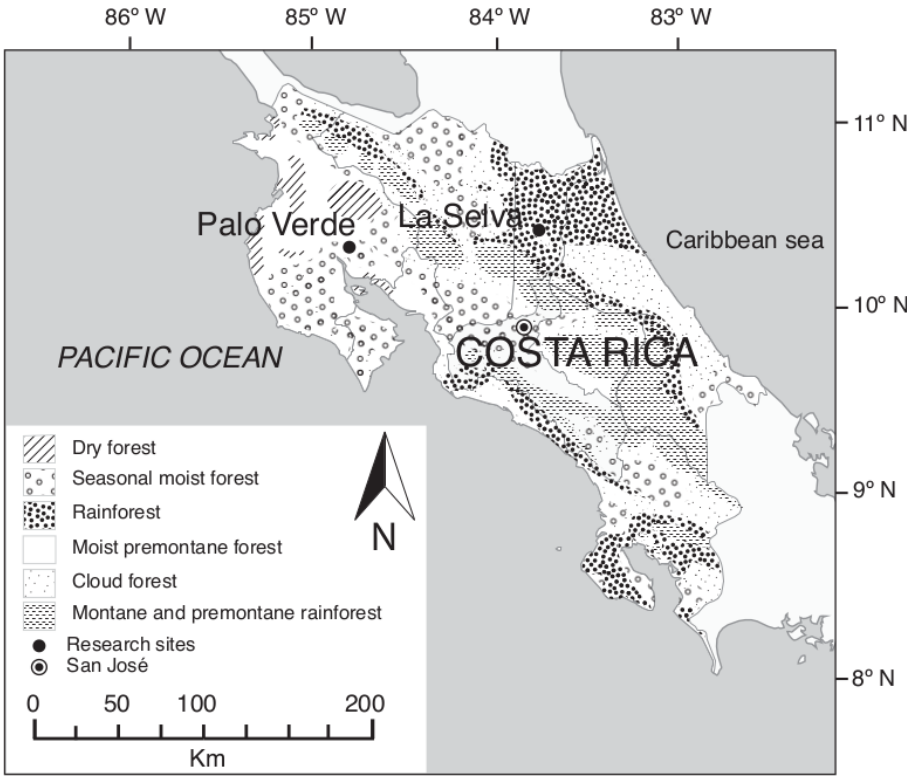


Figure 1.1. Map of the study site.

Vegetation zones were simplified from Kohlmann et al., 2010.

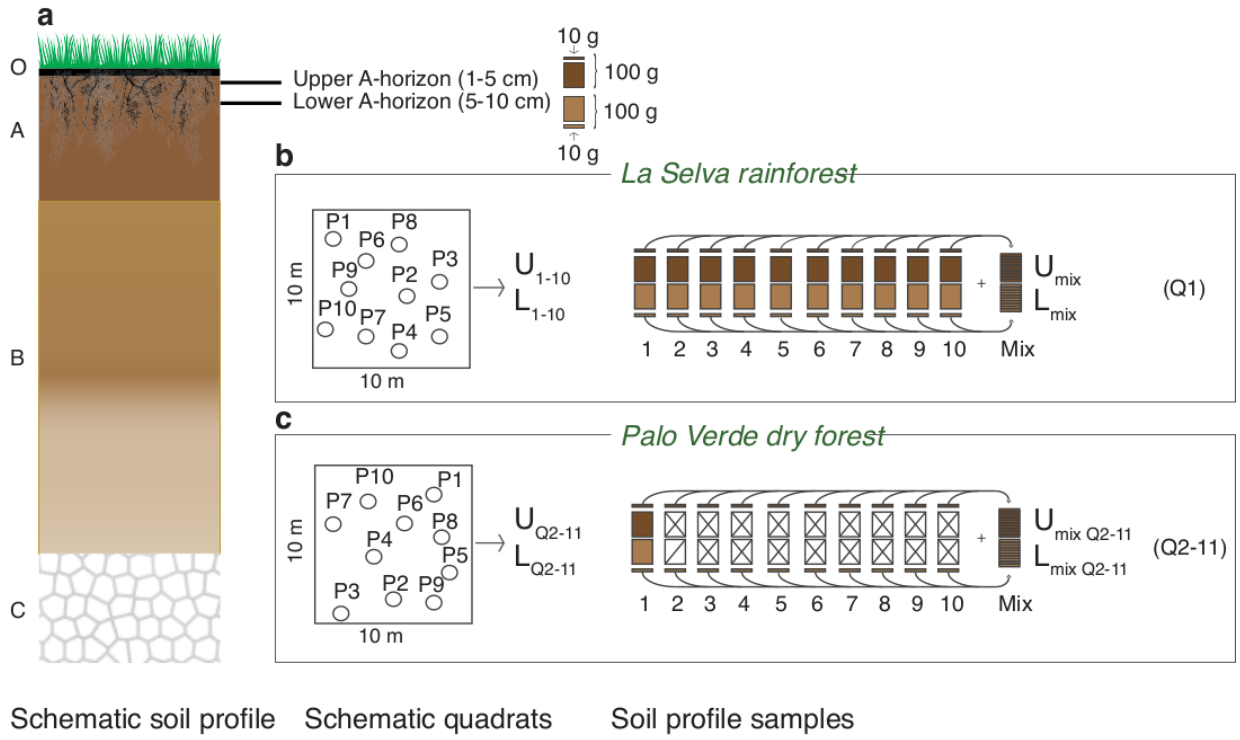


Figure 1.2. Sampling methodology.

a. Schematic soil profile showing the location of the samples collected. O = organic horizon, A = surface horizon, B = subsoil horizon, C = parent rock material (Weil and Darby, 2016). Samples (100 g) collected within the A horizon are represented by the light and dark brown rectangles which correspond, respectively, to the upper A-horizon (U) and the lower A-horizon (U); smaller subsamples (10 g) are represented by the smaller light and dark brown rectangles; **b.** Sampling schemes at La Selva; **c.** Sampling scheme at Palo Verde. Within a 10 x 10 m quadrat ten randomly located soil profiles (P1 to P10) were excavated. At La Selva, ten U and ten L samples (100 g) were collected ($U_{1\text{-}10}$ and $L_{1\text{-}10}$); a U mix (mU) and a L mix (mL) were obtained by combining subsamples (10 g) of, respectively, all ten U and all ten L samples. At Palo Verde, one U and one L sample (100 g each) were collected in one soil profile (P1) in each of the ten quadrats ($U_{Q2\text{-}11}$ and L_{Q211}); in addition, in each quadrat, a subsample (10 g) was collected from U and L in all ten soil profiles and combined to obtain one mix of all U and one mix of all L per

quadrat ($mU_{Q2\text{-}11}$ and $mL_{Q2\text{-}11}$). See text for further explanation for the difference in sampling scheme.

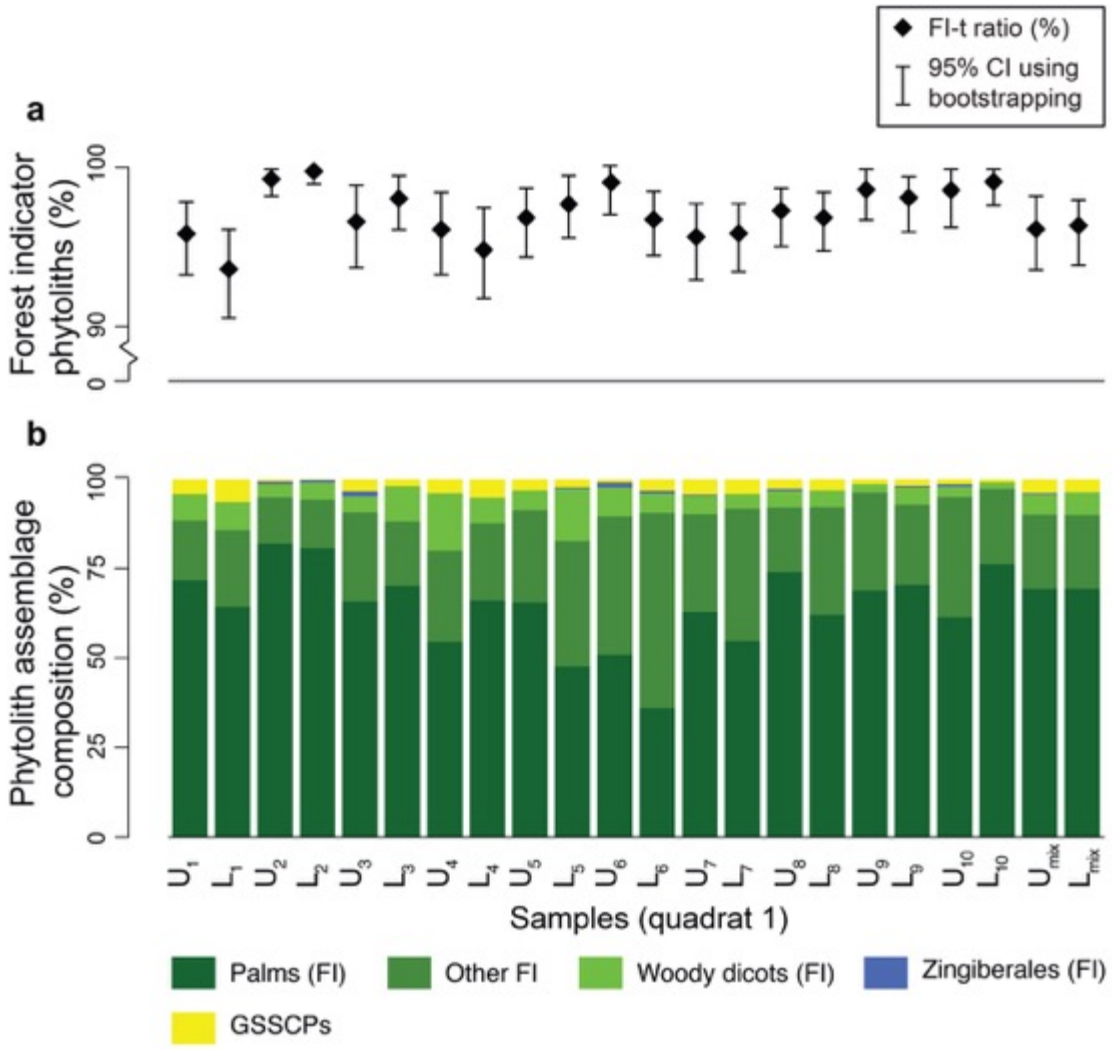


Figure 1.3. Phytolith assemblage *FI-t* and counts for La Selva

a. *FI-t* ratio expressed as the percentage of forest indicator (FI) phytoliths in a sum of FI and grass phytolith morphotypes (%FI/(FI+GRASS)) in the upper and lower A-horizons (U and L) of the ten soil profiles (1-10), and the two mixed profiles (mU_{Q1}, and mLA_{Q1}). Diamonds represent estimated *FI-t* ratios (from counts); vertical bars represent 95% confidence intervals based on bootstrap analysis of each sample; **b.** Phytolith assemblage composition of the samples indicating percentages of phytoliths in each of the six defined phytolith functional groups (in legend). Percent forest indicators in **a.** refers to the sum of all forest indicators (green gradient + blue in **b.**

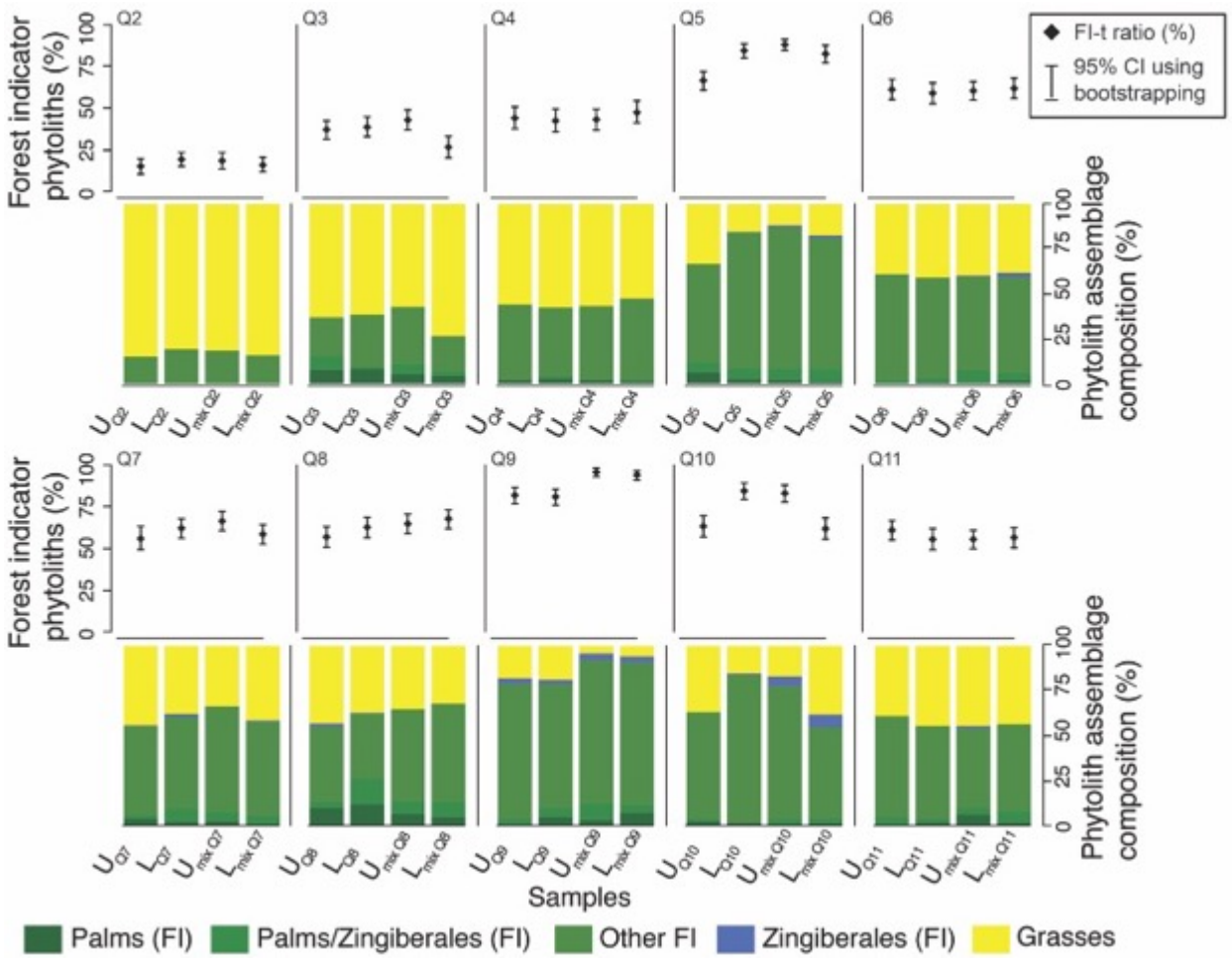


Figure 1.4. Phytolith assemblage *FI-t* and counts for Palo Verde.

FI-t ratio expressed as the percentage of forest indicator (FI) phytoliths in a sum of FI and grass phytolith morphotypes (%FI/(FI+GRASS)) in the upper and lower A-horizons (UA_{QxP1} and LA_{QxP1}) and the two mixed profiles (mU_{AQx} and mLA_{Qx}) for each one of the studied quadrats (Q2-11). Diamonds represent estimated *FI-t* ratios (from counts); vertical bars represent 95% confidence intervals based on bootstrap analysis. The bar plots represent phytolith composition of the samples indicating percentages of phytoliths in each one of the five defined phytolith functional groups (in legend). Percent forest indicators in upper graphs refers to the sum of all forest indicators (green gradient + blue in bar plots).

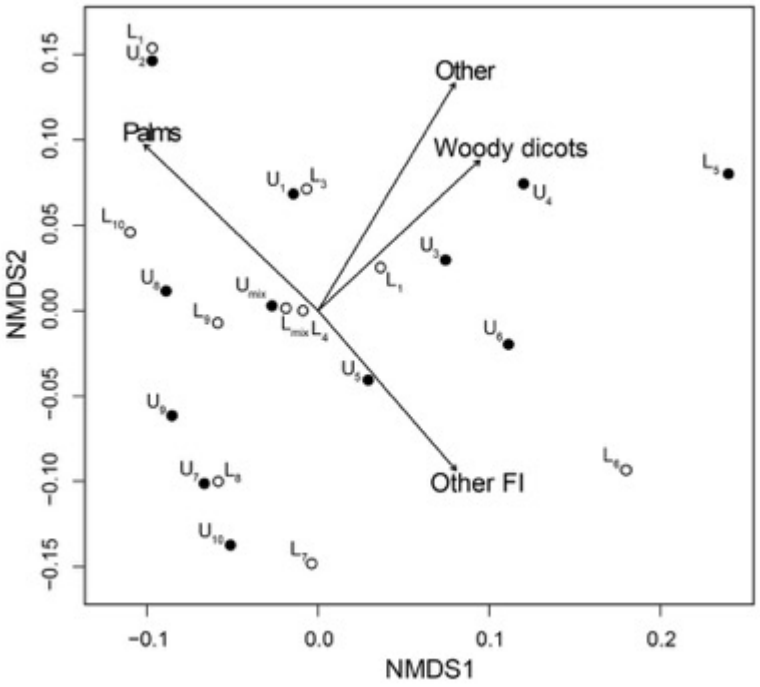


Figure 1.5. Ordination diagram of phytolith assemblages from La Selva.

The NMDS ordination diagram shows the soil profiles in the primary quadrat (Q1) of La Selva rainforest, and significant ($p<0.01$) phytolith classes (black vectors); stress value = 0.14. The direction of the vectors indicates maximum correlation between the scores of the samples (soil assemblages) and the phytolith classes. Assemblages from both lower and upper A horizons from the ten soil profiles in Figure 1.3 are included. Filled and empty circles indicate respectively upper and lower A-horizons samples.

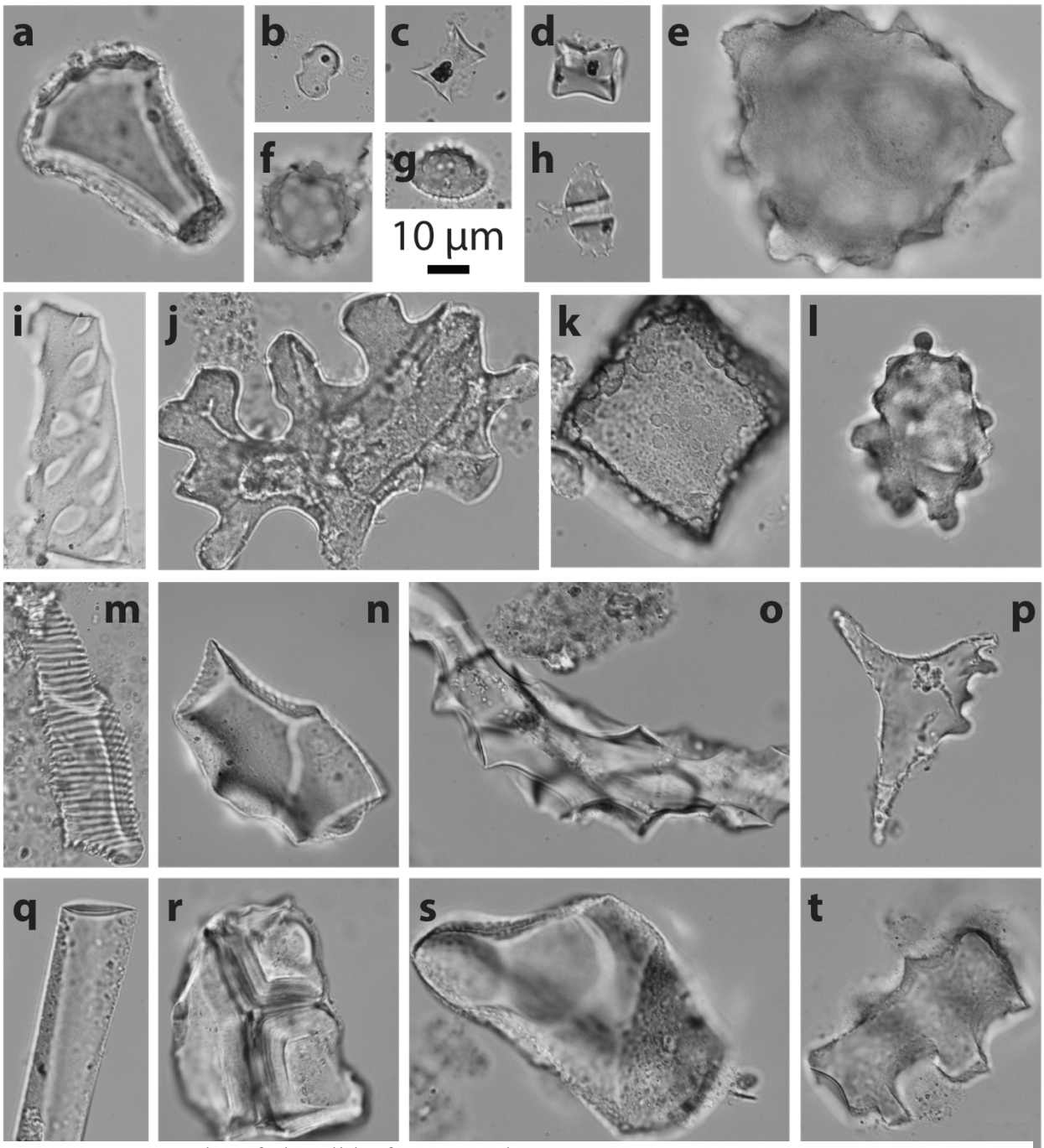


Figure 1.6. Examples of phytoliths from La Selva.

Phytolith morphotypes from La Selva tropical rainforest diagnostic of Poaceae (a, b, c, d), palms (f, g), Zingiberales (e h, l), dicots (j, r), monocots (k,q), forest indicators (m, p), and woody dicots (n, o, s, t), as well as non-diagnostic or produced by unknown plant taxa (i): **a)** Blo-10, bulliform keystone-shaped plate (GRASS-D); **b)** PO-1, polylobate with symmetry A GSSCP (Pooids); **c)** CO-1, generic conical rondel GSSCP (Pooids); **d)** SA-3, collapsed saddle GSSCP

(Bambusoideae); **e)** “Kn2 knobby” body from Heliconiaceae (Chen & Smith, 2013) (Zingiberales); **f)** Clm-2, echinate sphere (Palms); **g)** Clm-1, *Nypa*-type body (Palms); **h)** “T1 papillate trough” from *Heliconia* (Chen & Smith, 2013) (ZINGI); **i)** vessel element (Woody dicots?); **j)** Epi-2 Anticlinal epidermal (Other FI); **k)** Blo-1, rectangular plate (GRASS); **l)** “subglobular papillate” from Marantaceae (Chen & Smith, 2013) (Zingiberales); **m)** Tra-1, helical tracheary element (FI); **n)** Scl-3, multifaceted S-body (Woody dicots); **o)** Undetermined multidimensionally faceted S-body (cf. Postek, 1981; Piperno, 1988; Runge, 1999) (Woody dicots); **p)** Scl-6, spongy mesophyll body (FI); **q)** Elo-3, faceted elongate body (GRASS); **r)** VI-2, non spherical VI (laminated) body (Other FI); **s-t)** unclassified sclereid bodies (Woody dicots?).

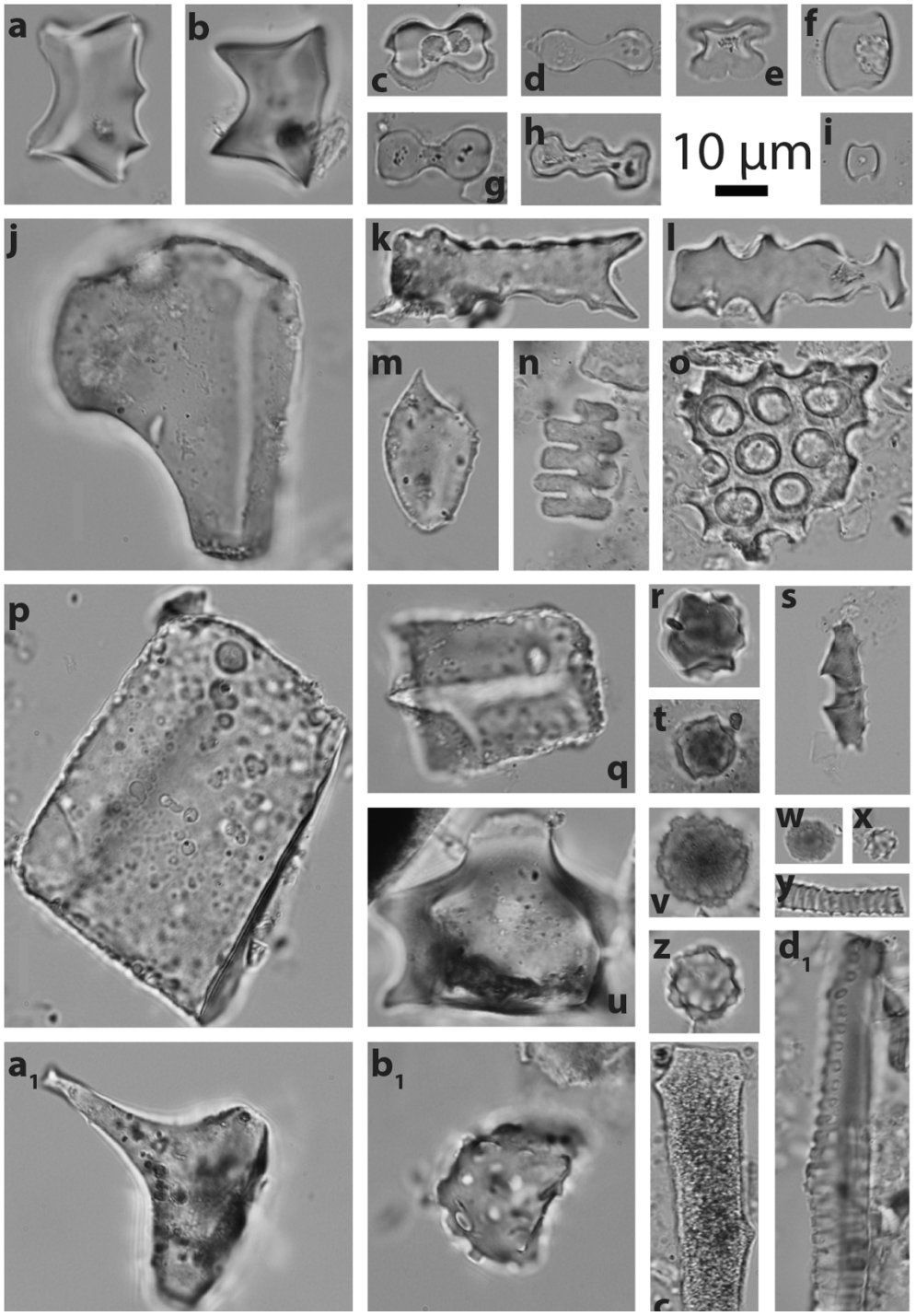


Figure 1.7. Examples of phytoliths from Palo Verde.

Phytolith morphotypes from Palo Verde tropical dry forest diagnostic of Poaceae (a-n), grass/monocots/conifers (o-q), Zingiberales (r-v), Palms (w), Palm/Zingiberales (x, z), general forest indicators (Other FI) (y, a_1, d_1), and Woody dicots (b_1), as well as non-diagnostic or produced by unknown plant taxa (Other) (c_1): **a)** SA-3, collapsed saddle GSSCP (Bambusoideae/Oryzoideae/Early diverging grasses); **b)** SA-3, collapsed saddle GSSCP (Bambusoideae/Oryzoideae/Early diverging grasses); **c)** BI-7, symmetry D bilobate GSSCP

(PACMAD clade); **d)** BI-5, symmetry B bilobate GSSCP (PACMAD clade); **e)** CR4-2, four lobed cross with near cross-shaped to cross-shaped top GSSCP (Panicoideae);); **f), i)** SA-1, true saddle GSSCP (Chloridoideae); **g)** BI-8, Panicoid-type bilobates GSSCP; **h)** PO-3, symmetry C polylobate GSSCP (PACMAD clade); **j)** Blo-10, bulliform keystone-shaped plate (GRASS); **k)** Epi-11, elongated epidermal with indented end (GRASS); **l)** Epi-12, papillate epidermal with large/elongate papilla (GRASS); **m)** Tri-8, spindle or teardrop-shaped trichome (GRASS); **n)** M-7, vertebral column mesophyll body (GRASS); **o)** M-6, crystal-like/well silicified straight file parenchyma (GRASS); **p)** Blo-2, thickened rectangular plate (Other); **q)** Blo-3, faceted rectangular plate (GRASS); **r), t), v)** "D1- and D2-type druses" from Heliconiaceae (Chen & Smith, 2013) (Zingiberales); **s)** "T1 papillate trough" from Heliconiaceae (Chen & Smith, 2013) (Zingiberales); **u)** "Kn2 knobby" body from *Maranta* seed (Chen & Smith, 2013), (Zingiberales); **w)** Clm-1, *Nypa*-type body from palms or alternatively "rugose hat-shaped" from Marantaceae (Chen & Smith, 2013), (Palm/Zingiberales); **z)** Clm-2, echinate sphere (Palm); **x)** Clm-2, echinate sphere from palms or alternatively, globular microechinate from Zingiberaceae (Chen & Smith, 2013), (Palm/Zingiberales); **y, d_1)** Tra-1, helical tracheary element (Other FI); **a_1)** Scl-4, spongy mesophyll body (Other FI); **b_1)** Scl-8, compact irregular S-body type A (Woody dicots); **c_1)** Elo-1, smooth elongate (Other).

Chapter 2. Modern phytolith assemblages across two neotropical forests: implications for interpretation of fossil assemblages

ABSTRACT

Phytolith analysis has emerged as a promising tool in deep time paleoecology, thanks to its potential for reconstructing past vegetation with higher spatial resolution compared to more traditional high-resolution proxies, such as pollen and spores. However, current data are insufficient to establish modern calibrations for the deep time fossil record and thus understand the fossil phytolith records in different vegetation types, as only a handful of phytolith modern analogue studies have been conducted so far. We seek to bridge this gap, by providing a modern reference study of soil phytolith along transects in two Neotropical vegetation types in Costa Rica: a rainforest and a dry forest. We characterize vegetation structure, composition, and diversity across environmental gradients, and compare this data to soil phytolith assemblages along yhe same gradients. We investigate two main questions: 1) is heterogeneity in vegetation structure, composition, and diversity recorded in phytolith assemblages in rainforest and a dry forest soils? 2) Can we use one or more phytolith assemblages to characterize these two vegetation types, and distinguish them in the fossil record? Our results suggest that overall, some structural, and compositional aspects of habitat heterogeneity can be reconstructed through spatial sampling (along a transect) and the analysis of phytolith assemblage composition. In addition, we argue that phytolith assemblage heterogeneity within a habitat is indicative of habitat heterogeneity. Our methodological approach therefore suggests that rainforest and dry forest vegetation can be distinguished in the fossil phytolith record. We provide suggestions for future modern analogue studies that could be conducted in other Neotropical vegetation types,

using the same methodology, in order to increase the range of habitats that can be reconstructed in the fossil record using phytolith analysis., including non-modern analogues.

2.1 INTRODUCTION

A central problem in paleoecological reconstruction is understanding the spatial scale that is captured in the fossil record and how and whether this reflects to the biological process or pattern of interest. For example, spatial environmental heterogeneity has been proposed as a possible driver of species richness in geological times by several authors (e.g., Janis et al., 2004; Kohn and Fremd, 2008; Finarelli and Badgley, 2010; Badgley et al., 2017). While the relationship between environmental heterogeneity and species richness has been definitively documented by modern community ecology studies (e.g., MacArthur and Mac Arthur, 1961; Auerbach and Shmida, 1987; Fraser, 1998; Willig et al., 2003; Tews et al., 2004; Cramer and Willig, 2005; Stein et al., 2015), it is harder to establish such relationship in the fossil record, where measuring environmental heterogeneity is not straightforward.

Among the factors known to cause such heterogeneity, paleotopography, primarily inferred from the record of tectonic activity has received a lot of attention in recent years (e.g., Kohn and Fremd, 2008; Badgley et al., 2017). Yet, topographic complexity is not the sole source of environmental heterogeneity because environmental gradients are not merely altitudinal, and act at many different spatial scales (within and between habitats).

One of the most important components of environmental heterogeneity is vegetation type. However, inferring spatial variation in vegetation from plant fossils is traditionally fraught with problems relating to resolution and scale of the fossil record. Thus, plant macrofossil assemblages provide instantaneous snapshots of the local vegetation but are usually restricted to lakes and river back-levee lagoons. In addition, they only allow for limited inferences about vegetation architecture (but see Burnham, et al., 1992) because they usually consist of isolated plant organs (such as leaves, trunks, flowers, fruits, and seeds) and tend to have an

overrepresentation of woody plants (but see, e.g., Wing et al., 1993, 2012). In contrast, although pollen and spores tend to be preserved in lakes or wetland deposit, and provide reconstruction of arboreal cover, they generally integrate across many local environments (pollen and spores can be dispersed over long distances), providing a more regional signal.

Phytolith analysis has emerged in the last two decades as a powerful tool in paleoecology, not least because phytolith assemblages have long been thought to often reflect relatively local plant communities, thus allowing for reconstruction of vegetational gradients (Piperno, 1988; Strömberg et al., 2018). Phytoliths are microscopic silica bodies produced by most vascular plants by absorption of monosilicic acid (H_4O_4Si, in solution in water) and subsequent precipitation in the form of silicates or opal-A in the intercellular, and intracellular space of plant tissues (Piperno, 1988). Phytoliths are deposited in situ after a plant decays (Piperno, 1988, 2006), thus arguably providing a more local vegetation signal than palynomorphs, which tend to be released from the aboveground shoots of plants (Prentice, 1985).

Moreover, phytoliths are preserved in ancient sediments and paleosols where they are often found in association with vertebrate remains (Strömberg, 2002). While the taxonomic resolution of phytoliths is limited (except for grasses, family Poaceae), the classification of phytolith morphotypes into plant functional groups allows reconstruction of relative changes in vegetation cover through time with accuracy. A prime example of the application of phytolith analysis to the deep time fossil record is the study of mid-Cenozoic phytolith assemblages from the North American continental interior, which tracked regional patterns of grass evolution, diversification and expansion through the Late Cenozoic of North America (e.g., Strömberg, 2005; Harris et al., 2017).

Despite the recent surge in phytolith-based studies in paleoecology, much work still remains to standardize the use of phytoliths in deep time and understand what phytoliths can and cannot tell us about vegetation. For example, the assumption that phytolith assemblages generally record local plant communities remains largely untested despite numerous modern-analog studies have emerged in recent years (e.g., Piperno, 1988; Fredlund and Tieszen, 1994; Kerns et al., 2001; Bremond et al., 2005; Mercader et al., 2011; Aleman et al., 2012, 2014; Crifò and Strömberg, in press). Thus, landscape-level patterns in phytolith assemblages are poorly understood, as are potential differences in the spatial resolution of phytolith assemblages from different environments. So far, these topics have been investigated in only a handful of modern analogue studies (Piperno, 1988; Fredlund and Tieszen, 1994; Kerns et al., 2001; Bremond et al., 2005; Mercader et al., 2011; Aleman et al., 2012, 2014). Overall, these studies suggest that some spatial sampling of phytoliths within a habitat (the equivalent of taking multiple samples along one stratigraphic level) rather than using a single point sample is a preferred approach to reconstruct local past vegetation cover, because using multiple samples allows for documentation of local spatial variation, while a single sample might only provide a "super-local" signal (e.g., Aleman et al., 2014). However, there is no consensus on the appropriate scale of such spatial sampling mainly due to the choice of different vegetation types, sampling methodologies, and analytical approaches by different phytolith researchers (see Crifò and Strömberg, in press). Indeed, some of these studies show a good correspondence between phytolith assemblages and vegetation cover on a scale varying from tens of meters in tropical rainforests and African savannas (Piperno 1989; Aleman et al., 2012) to hundreds of meters in a forest-savanna transition in Africa (Bremond et al., 2005), or tens of kilometers in a woodland-grassland transition in Mozambique (Mercader et al., 2011). Other studies found only a weak relationship

between phytolith assemblages and vegetation (Kerns et al., 2001) or point to a regional rather than local phytolith signal (Fredlund and Tieszen, 1993).

Another limitation of previous phytolith modern analogue work is that it has not been designed to be applied directly to the interpretation of the deep time fossil record (see chapter 2). However, with the increased interest in using phytoliths in pre-Quaternary paleoecology, studies with deep-time application are necessary (see Strömberg et al., 2018). One such recent study sought to determine how modern soil should be sampled for phytolith analysis over the scale of a small quadrat (Crifò and Strömberg, in press) in a Costa Rican rainforest and dry forest. This work indicates that in order to capture vegetation signal within a 10x10 meter quadrat, a single point sample can be collected (rather than multiple samples) from the lower part of the soil A-horizon, which corresponds more closely to paleosol layers commonly available in the fossil record (compared to surface or subsurface layers traditionally used in previous modern analogue studies). However, this study is not informative about larger spatial patterns of soil phytoliths, and therefore does not provide appropriate guidelines for sampling along vegetation transects within a landscape or a fossil locality.

Finally, modern analogue studies of phytolith assemblages typically reconstruct broad patterns of vegetation cover without focusing on linking phytolith assemblages with multiple ecological or functional ecology aspects of the vegetation (e.g. Fredlund and Tieszen, 1994; Alexandre et al., 1997; Kerns et al., 2001 Bremond et al., 2005; Barboni et al., 2007; but see Mercader et al., 2011). Hence, these studies consider phytolith assemblage composition primarily in term of relative proportion of forest versus open habitat indicators, whereas relative proportions of different phytolith morphotypes (phytolith types produced by different plants) are disregarded because little is known about their ecological and functional significance. Yet, plant

functional type affinity of most phytolith morphotypes is often known based on reference collections of modern plants (e.g. Piperno, 1985; Strömberg, 2003), and has the potential to improve paleoenvironmental reconstructions. For example, phytolith morphotypes indicative of forest vegetation are produced by woody and herbaceous dicotyledonous species, palms, and zingibers –groups that are not present in the same proportions in all vegetation types, and whose presence and abundance might be related to important aspects of the environment and climate such as, for example, water availability, disturbance, and seasonality.

Our work seeks to contribute to the knowledge of the spatial resolution of the phytolith fossil record by analyzing phytolith soil assemblages across vegetation transects in two functionally different forests, a dry semi-deciduous forest, and a rainforest, in Costa Rica. These two forest types differ in their structure, composition, diversity, and climate. For example, wet forests are more species rich, have denser and more homogeneous canopy cover, and present weaker seasonality in phenology as well as precipitation than dry forests (Holdridge,1967; Murphy and Lugo, 1986). We analyze the relationship between soil phytolith composition, vegetation structure and composition within and across these two forest types in a way that will make the results maximally applicable to deep-time paleoecology. To do so, we sampled along elevation and moisture gradients (at varying distance from the river) over an area that is comparable to a typical fossil locality (1-2 km), and used lower A-horizon soil samples, which have been shown to be most similar to the samples taken in paleosols (Crifò and Strömberg, in press); second, we attempt to focus on quantifying aspects of the vegetation (degree of tree cover/structure, composition, species richness) that are central for understanding the functioning of plant communities; and third, we consider parameters that are known to affect phytolith assemblages (phytolith production, preservation) to interpret potential differences between

observed vegetation and phytolith assemblages. Our two central questions are: 1) is variation in (standing) vegetation structure, composition, and diversity within a habitat recorded in phytolith assemblages and, if so, at what scale? 2) Can we characterize different vegetation types (i.e., dry forest and rainforest) using (one or more) phytolith assemblages, such that we could tell them apart in the fossil record?

2.2 POSSIBLE VEGETATION PREDICTORS AND BIASES OF THE PHYTOLITH SIGNAL IN THE SOIL

To better understand the links between vegetation type and phytolith assemblages we compare (standing) vegetation structure, composition, and function with phytolith assemblages at multiple sites within two distinct vegetation types. We studied two vegetation types–a tropical rainforest (La Selva Biological Station; hereinafter referred to as La Selva), and a tropical dry forest (Palo Verde National Park; hereinafter referred to as Palo Verde)–that differ in climate conditions, and therefore in forest structure, composition, and function. Overall, with regard to forest structure and composition, wet forests are characterized by high species richness, high complexity, high mean canopy height and number of canopy strata, high and homogeneous leaf area index (LAI), and high tree basal area and biomass (Holdridge, 1967). With respect to functional traits, wet forests are highly productive (i.e., have high net primary productivity, fine litter production, and annual tree diameter growth). In addition, wet forests exhibit continuous or intermittent growth, and are composed of mainly evergreen species with aseasonal phenology (Holdridge 1967, Murphy and Lugo, 1986). Costa Rican wet forests overall receive between 2,500 and 5,000 mm of annual precipitations, and have no pronounced rainy season (Quírico Jiménez et al., 2016); at La Selva Biological Station mean annual precipitation (MAP) is 4,000

mm, and temperature varies daily between 19 and 31˚C (Organization for Tropical Studies, 2019).

Tropical dry forests differ from wet forests in term of structure, composition, phenology and climate. Hence, dry forests have lower species richness, complexity, mean canopy height and number of canopy strata, lower and more heterogeneous (patchy) leaf are index (LAI), as well as lower tree basal area and biomass than wet forests (Holdridge, 1967). In addition, they exhibit growing periodicity, and are composed of deciduous and evergreen species with both seasonal and aseasonal phenology. Climate is highly seasonal in Costa Rican dry forests, with 95% of precipitations falling during the rainy season (May to December), (Quírico Jiménez et al., 2016); at Palo Verde National Park, MAP is 1,500 mm (Organization for Tropical Studies, 2019), and mean annual temperature (MAT) is 25˚C (Gillespie et al., 2000).

These differences in climate, vegetation type, and environmental factors might influence the soil phytolith record in several ways. Below, we give a brief overview of possible drivers and biases of phytolith composition in the soils of La Selva rainforest and Palo Verde dry forest.

2.2.1 *Drivers and biases linked to forest structure and composition*

The phytolith signal in rainforest soil should mainly be driven by their denser canopy (and less dense understory vegetation) as opposed to a patchier canopy (and denser herb and grass understory vegetation) in dry forests. At La Selva, we expect phytolith assemblages to be homogeneous across different quadrats, and dominated by phytoliths produced by dicotyledonous trees and palms (the latter being high phytolith producers, which should tend to be overrepresented), with a minor component represented by phytoliths produced by other plant groups, such as members of the order Zingiberales, and herbaceous dicots. In contrast, at Palo

Verde we expect more heterogeneous phytolith assemblages with varying proportions of phytolith produced by plants indicative of forest habitats, and grass phytoliths (open habitat indicators). In addition, among the phytoliths produced by forest plants, we expect the relative abundances of different types to vary across quadrats, a pattern related to vegetation patches of distinct species and density. Similarly, we expect the abundance of grass phytoliths to vary accordingly to forest patchiness, and to be higher in areas where the canopy cover is less dense.

These fundamental differences in the phytolith assemblages between La Selva and Palo Verde should correspond to differences in the *FI-t* ratio, a coarse measurement of grass-woody dynamics, with higher and more homogeneous *FI-t* ratio at La Selva than at Palo Verde. In contrast, phytolith assemblage diversity (number of phytolith morphotypes) is unlikely to reflect plant taxonomic diversity except in a very broad sense, due to the phytolith morphotype properties of *multiplicity* (a single species produces many different morphotypes) and *redundancy* (multiple species produce the same morphotype), (Rovner, 1971). Thus, we predict that we will not see a relationship between species richness and the number of different phytolith morphotypes encountered within a quadrat. We also expect that differential specific silica yield at the two sites might bias phytolith assemblages in ways that can be, in part, predicted by investigating the silica production of dominant taxa at each site. For instance, high abundance of big producer taxa such as palms (forest indicators) at La Selva, and grasses (open habitat indicators) at Palo Verde, likely result in more extreme *FI-t* ratio and overrepresentation of certain taxa in the phytolith assemblages. Furthermore, at Palo Verde this effect might be exacerbated by species of dominant families in dry forests (e.g., Fabaceae) that are known to be non-producers (Hodson et al., 2005; Strömberg et al., 2016), or by trees species characterized by low litter biomass production, such as evergreen species.

2.2.2 *Drivers and biases linked to environmental factors*

Environmental and climatic factors are likely to influence the soil phytolith signal through both plant silica accumulation and phytolith taphonomy. Silica uptake by plants occurs both passively and actively, with one of the major mechanisms for passive silica uptake in plant tissues being transpiration (Mitani and Ma, 2005). Lower transpiration rates in dry forests (compared to rainforests) might result in overall lower levels of silica accumulation at Palo Verde compared to La Selva. Further, it is possible that because active transport does not depend on transpiration, this effect would be diminished in species that are thought to rely largely on active (in addition to passive) forms of silica uptake and accumulation, such as grasses (Hodson et al., 2005; Piperno, 2006; Trembath-Reichert, et al., 2015; Strömberg et al, 2016). Such tendency would enhance the overrepresentation of open habitat indicator at Palo Verde, hence a bias toward lower *FI-t* ratios (more open habitat signal). In addition, phytolith incorporation into the soil record requires soil formation to take place. This process is favored by wet and warm conditions (Weil and Brady, 2016), which are constant at La Selva rainforest but not at the more seasonal Palo Verde dry forest. Slower soil formation rates (thus phytolith incorporation into the soil) and possible higher erosion at Palo Verde might result in an overall less abundant, less preserved and more "seasonal" soil phytolith assemblages at Palo Verde compared to La Selva.

Second, one of the main processes inhibiting phytolith preservation is chemical dissolution, which is favored by high soil pH (>8.5) and acts faster on phytolith morphotypes with high specific surface area (surface area per unit volume; i.e., forest indicator phytoliths) (Wilding and Drees, 1974; Bartoli and Wilding, 1980; Cabanes et al., 2011). Poor phytolith preservation has

been documented at Palo Verde (Crifò and Strömberg, in press), possibly due to alkaline soil pH at Palo Verde which is also suggested by the abundant presence of calcium carbonate detritus in the soil. In addition to overall low preservation, these chemical conditions might also result in overrepresentation of open habitat phytoliths (grasses; characterized by low surface area per unit volume).

Third, although phytoliths are thought to be deposited in situ in the soil (Piperno, 1988, 2006), dispersion might take place to some degree, depending of the presence and intensity of transporting agents, such as alluvial, and eolian factors, fire, and herbivory (e.g. Fredlund and Tieszen, 1997; Alexandre et al, 1997; Strömberg et al, 2018). While we avoided flooded areas and pastures, fires and winds are frequent at Palo Verde (Quírico Jiménez et al., 2016), and it is likely that they have a stronger effect on the phytolith signal at Palo Verde than at La Selva because of the more common presence of open patches in the vegetation. In particular, the smallest phytolith morphotypes (i.e., grass short cells, and palm phytoliths), might be overrepresented in the soil assemblages of Palo Verde dry forest because they tend to be preferentially transported by both wind and fire, for example the adjacent grass dominated wetland.

The take-home message is that various confounding factors are at play when comparing local vegetation type and phytolith assemblages. In this study we focus primarily on the relationship between vegetation type and phytolith assemblages, but we take into account some environmental factors when interpreting our results. We aim to determine the relationship between vegetation and soil phytolith assemblages within and across two different habitats in order to build a modern reference allowing us to characterize these two vegetation types in terms of structure, composition, function, and spatial heterogeneity, and to distinguish them in the

fossil record. Hence, we hypothesize that 1) despite possible biases, at each site spatial sampling is effective for reconstructing structural, compositional, and functional aspects of the vegetation and how they might vary across space, and that 2) when comparing the two sites, phytolith assemblages should differ in a way that would allow distinction of these vegetation types in the fossil record. To investigate what factors might be driving differences within and between sites, we also measured several vegetation variables such as canopy cover and phytolith productivity of dominant species, as well as environmental variables such as altitude, proximity to water, and geographic adjacency of quadrats along the transect (measured as longitude). We then compare our phytolith count data, vegetation structure data, and other measured environmental variables statistically to evaluate whether these factors covary with observed patterns in vegetation and/or soil phytolith assemblages.

2.3 MATERIALS AND METHODS

2.3.1 *Study sites*

Soil and plant material were collected in Costa Rica at the Organization for Tropical Studies (OTS) La Selva (10° 26’ N, 83° 59’ W) and Palo Verde (10° 21’ N, 85° 21’ W) biological stations during a single field season in February/March of 2016 (Fig. 2.1). La Selva Biological Station is located in the lowland of northern Costa Rica, and it includes a large area covered by primary tropical rainforest. Palo Verde Biological Station is located on the Pacific slopes of northwestern Costa Rica, and includes areas covered by the remnant of seasonal and transitional dry forests in the Neotropics (Organization for Tropical Studies, 2019). These two sites are protected areas, with known disturbance history, and therefore ideal for our study.

2.3.2 *Sample and data collection*

2.3.2.1 LAI and floristic data

At La Selva and Palo Verde, respectively, we established fourteen and eleven 10 x 10-meter quadrats along vegetation transects (see Crifò and Strömberg, in press). In La Selva we traced one ~1.5 km transect; at Palo Verde we traced one ~1 km and one 0.5 km transect (T1 and T2, respectively, Fig. 2.1). Quadrats were established along transects randomly, except that we avoided disturbed areas and flooded surfaces, resulting in an uneven distribution.

Within each quadrat, we measured LAI from hemispherical photographs following the protocol in Dunn et al., (2015) to estimate canopy openness. We took 10 hemispherical photos per quadrat using a Nikon CoolPix 4500 camera with attached Nikon Fisheye Converter FC-E8 0.21x lens mounted on a tripod at 1 m from the ground. Photographs were taken at different exposures and the optimal exposure was chosen for each photo following the protocol described by Zhang et al., (2005). Effective Leaf Area Index (L_e) and % canopy openness were measured from the photographs using Gap Light Analyzer (GLA) Version 2.0 imaging software. For our sites we used the LAI Ring 5 output from GLA which is the effective leaf area index integrated over the zenith angle of 0-75° (Dunn et al., 2015). We calculated a LAI value for each quadrat by averaging the LAI Ring 5 output from the ten photographs per quadrat. We used fewer than ten photographs when image quality was poor (blurry, or too bright due to sunny conditions). To quantify geographic relatedness of quadrats, we measured their planar coordinates using a GPS (Garmin GPSMAP® 62sc). To estimate species richness within sites and quadrats, as well as tree biomass, for each tree species present in a quadrat we identified all the plant species present with the exception of leafless deciduous species (we counted less than one per quadrat), and grasses; we recorded the number of individuals per each species present in a quadrat, and measured their

dbh (diameter at breast height, i.e., at 1.4 meters). To estimate the individual contribution of each species to the canopy cover of a quadrat (hence to its phytolith pool) we use their relative basal area obtained from the dbh. A positive relationship between tree basal area and LAI (canopy cover) has been previously found in rainforests (e.g. Clark et al., 2008) and dry forests (e.g., Kalácska et al., 2004).

Species relative basal area (rBA_{sp}) in each quadrat was calculated as follow:

$$rBA_{sp} = \frac{BA_{sp}}{A} \times 100 \tag{2.1}$$

Where BA_{sp} is the total basal area of a species calculated as the sum of the basal areas of all individuals obtained from their trunk dbh:

$$BA_{sp} = \sum_{i=1}^{n} \pi \left(dbh_i / 2 \right)^2 \tag{2.2}$$

The total tree basal area of a quadrat (A) was calculated as the sum of all species basal areas (BA_{sp}) as follows:

$$A = \sum_{i=1}^{n} BA_{sp_i} \tag{2.3}$$

While vegetation surveys investigating species diversity typically take into account only individual trees with $dbh \geq 5$ cm, here we consider all individuals with $dbh \geq 1$ cm as these plants might contribute considerably to the soil phytolith pool (especially when multiple individuals per species and/or high phytolith producers are represented). Hence, because of the way it is calculated, we used tree basal area exclusively for the purpose of estimating the relative biomass (hence, phytolith) contribution of each species to a quadrat, rather than using it as a measurement to characterize forest type as is typically done in ecology Grass and herbaceous plant *rBA* were not estimated.

We also estimated the distance between each quadrat and the closest water basin (as a proxy for moisture) by measuring the shortest distance to the Río Puerto Viejo (La Selva) and Río Tempisque (Palo Verde) using the line measurement tool in Google Earth Pro 7.3.4.5776. To asses elevation gradients, we recorded the elevation of each quadrat as reported by Google Earth.

2.3.2.2 Herbarium and soil sample collection

In each quadrat at La Selva and Palo Verde, we collected soil for phytolith extraction and plant samples to quantify vegetation structure and composition. We collected one soil sample per quadrat from the lower portion (5-10 cm deep) of the soil A-horizon after removing the overlaying litter and organic material (O horizon) (see Crifò and Strömberg, in press for more details about sample collection). We analyzed a total of fourteen (LS_{Q1-14}) and eleven (PV_{Q15-25}) lower A-horizon samples from fourteen and ten quadrats respectively at La Selva and Palo Verde; for a justification for the choice of sample depth analyzed, see Crifò and Strömberg (in press).

In addition, for each newly observed plant species, we collected herbarium material so that we have at least one herbarium reference sample per species present at the two sites for a total of 85 and 53 herbarium sheets from La Selva and Palo Verde respectively. At Palo Verde, we identified 55 plant species (49 dicot trees, 1 palm, 4 grasses, and 1 herbaceous dicot species). We collected material from 43 of the 49 tree species (for the remaining 6 species we were not able to collect any leaf material). Among these, 23 species are deciduous, 7 are semideciduous, and 13 are evergreen. At La Selva we identified 92 species (75 dicot trees, 8 palms, 2 grasses, and 7 other monocot species), and we collected material from all 75 tree and 8 palm species as well as from 3 monocots).

Soil samples were imported under USDA permit P330–14-00285. After sterilization, soil samples and resulting extracted material were reposited in the Burke Museum of Natural History and Culture (UWBM). Herbarium specimens were initially treated in a drying cabinet at the herbaria of La Selva and Palo Verde biological stations, and then deposited at the UWBM Botany Collections (WTU herbarium), University of Washington, Seattle; a second set of vouchers were deposited at the Herbario Nacional de Costa Rica, San José.

2.3.2.3 Phytolith processing and analysis of soil samples and herbarium specimens

To study the phytolith composition of soil assemblages, we extracted phytolith following standard methods for phytolith extraction from modern soils (Piperno, 1988; Zhao and Pearsall, 1998; see Crifò and Strömberg, in press). We used 1 gram from each soil sample, which had been previously homogenized. To estimate plant specific phytolith yield, we extracted phytoliths from dry plant material (2.5 grams per species) using standard protocols (Piperno, 1988). Dried subsamples of the extracted biogenic silica (i.e., sponge spicules, diatoms, chrysophyte cysts, and phytoliths from soils samples and phytoliths from plant samples) were mounted on slides and viewed at 1,000x magnification using a Nikon i80 compound microscope with mounted Nikon DS-Fi1 camera for phytolith identification and counting. We mounted the silica extracted from the soil samples on two sets of slides using, respectively, a solid mounting medium (Meltmount®) and a liquid mounting medium (immersion oil); the latter allowed for phytolith rotation and observation of their three-dimensional morphology. Silica extracted from plant material was mounted on slides using Meltmount® only.

We followed the classification scheme and analytical methods in Strömberg (2003, 2013), which use the relative abundance distribution of morphotypes among plant functional types to assign diagnostic morphotypes to plant functional groups. The proportion of different

functional groups ("compound variables") in a phytolith assemblage is used to interpret vegetation structure and composition. The plant functional types included in this study are 1) "Palm" (phytoliths produced by palms) ; 2) "Woody dicots" (including phytoliths produced by woody angiosperms; 3) "Other FI" (= other forest indicators, including phytolith produced by both woody and herbaceous dicots, and/or a group of plants called forest indicators comprising conifers, ferns and cycads; Strömberg 2003, 2005); 4) "Zingiberales" (phytolith produced by plants in the order Zingiberales; 5) "Palm/Zingi," which includes echinate sphere-like morphotypes found at Palo Verde that we were not able to assign to the "Palms" functional group (see Crifò and Strömber, in press); 6) "Palm/Zingi/Brom"(including <5 µm echinate sphere-like morhoptypes found at Palo Verde which, because of their small size, we could not assign to either of the groups that potentially produce them: palms, Zingiberales, Bromelids); 7) "Other", which are non-diagnostic and unknown phytolith morphotypes; 8) "BOP," including phytoliths produced by closed-habitat grasses of the Poaceae subfamilies Bambusoideae and Oryzoideae, and the early-diverging grasses in the subfamilies Anomochlooideae, Puelioideae, and Pharoideae; 9) "POOID-D," which includes phytoliths produced and diagnostic of the grass subfamily Pooideae; 10) "POOID-ND," including phytoliths produced by (but not exclusively) the subfamily Pooideae, and therefore non strictly diagnostic of this tribe; 11) "Chloridoideae," including phytoliths produced by the subfamily Chloridoideae, which present a C_4 metabolic path; 12) "Panicoideae," including phytoliths typical of the subfamily Panicoideae, which present mostly a C_4 metabolic path; 13) "PACMAD," including phytolith produced by the grasses in the PACMAD clade, which comprises C_3 and C_4 grass subfamilies Aristidoideae, Panicoideae, Chloridoideae, Danthonioideae, Arundinoideae, and Micrairoideae; and 14)"OTHG," which are "non-diagnostic and unknown grass phytolith morphotypes. We used

published literature to assign morphotypes that had not been previously described by Strömberg (2003, Strömberg et al., 2013, or thereafter) to the functional groups used herein.

The slides with solid mounting medium were used for general counting (all morphotypes), whereas the slides with liquid mounting medium were used to identify and count grass silica short cell phytoliths (GSSCP) whose three-dimensional morphology allows for taxonomic classification at the subfamily level within the Poaceae. We counted a minimum of 200 diagnostic morphotypes (i.e., belonging to our functional groups 1-5, 7-12 above) per slide (for both sets) for statistical significance (Pearsall, 2000; Strömberg, 2009) unless GSSCP represented <5% of the assemblages, in which case we counted all GSSCP present on one slide, after careful scanning.

Due to the potential effect of differential phytolith yield on soil phytolith assemblages especially in dry habitats (where low producers such as the family Fabaceae, are extremely abundant), we chose to estimate this factor for all collected Palo Verde tree species. To obtain semi-quantitative measurements of phytolith productivity by the plant species present at Palo Verde, we combined estimations of the relative phytolith yield (rY_{phyto}), and silica yield (rY_{Si}) of each species. To calculate (rY_{Si}), we estimated qualitatively the total amount of silica contained in the vials after extraction and desiccation (note that the same amount (2.5 grams) of plant material was used for extraction for all species), and ranked the species into six categories to which we assigned rY_{Si} values from 0 to 5 corresponding to 0) non-producers (no visible silica on the bottom of the vial), 1) very low producers (a faint, <1 mm thick, discontinuous film of silica on the bottom of the vial,), 2) low producers (a <1 mm thick, continuous film of silica at the bottom of the vial), 3) moderate producers (a ~ 1 mm thick, continuous film of silica at the bottom of the vial), 4) high producers (a >1 mm thick, continuous film of silica at the bottom of

the vial), and 5) very high producers (a thick ≥ 2 mm film of silica at the bottom of the vial). To calculate species relative phytolith yield (rY_{phyto}), we identified and counted all the diagnostic phytoliths present (n_{phyto}) in fifty fields of view randomly selected on each slide.

The relative phytolith yield for each species was calculated as follow:

$$rY_{phyto} = {}^{n_{phyto}}\!/_{50} \tag{2.4}$$

Central to this calculation is the assumptions that (a) our methods of preparing slides ensured that roughly the same volume/weight of dried silica is contained on each slide, and b) phytoliths were (on average) uniformly distributed across slides of different species, so that 50 fields together give a reliable estimate of phytolith abundance (which may vary due to differences in phytolith size produced by different species) . We then calculated the standard diagnostic phytolith yield per species (Y_{phyto}; yield per unit of extracted silica) as the product of the relative diagnostic phytolith yield (rY_{phyto}) and relative silica yield (rY_{Si}):

$$Y_{phyto} = rY_{Si} \times rY_{phyto} \tag{2.5}$$

This metric ranges from minimum values for species with little silica and few, large diagnostic phytoliths to maximum values for species with high silica yield and many, small diagnostic phytoliths. We ranked species based on their standard diagnostic phytolith yield using the quantiles of the Y_{phyto} distribution. Thus, we assigned the rank value of 0 to non-productive species, 1 to species in the first quantile (very low producers), 2 to species in the second quantile (low producers), 3 to species in the third quantile (moderate producers), and 4 to species in the 4th quantile (high producers).

We used the total basal area of each species (BA_{sp}) to estimate the relative contribution of each species, in each quadrat, to the total biomass (and soil phytolith pool). We compared the

relative contribution of each species to the quadrat total tree basal area to their phytolith yield rank. We interpret the phytolith assemblages of quadrats with a high proportion (>50%) of the total basal area represented by non-producer, very low producer, or low producer species, as potentially biased toward lower *FI-t* ratio values. Conversely, in quadrats with moderate or high proportion of high producers (>50%), *FI-t* ratio values may be inflated relative to their biomass contribution.

2.3.3 *Statistical analyses*

All statistical analyses were performed using R (R Core Team, 2017), and functions contained in R package *bootstrap* (Tibshirani and Leisch, 2017), *geosphere* (Hijmans, 2017), *MASS* (Venables and Ripley, 2002), *vegan*, (Oksanen et al., 2018), and *vegetarian* (Charney and Record, 2012).

2.3.3.1 *FI-t* ratio and robustness of *FI-t* ratio differences between soil phytolith assemblages.

For each sample the *FI-t* ratio ($FI\text{-}t_{quadrat}$) was calculated using the equation by Strömberg (2005):

$$FI - t_{quadrat}(\%) = \left(\frac{FI_{phytoliths}}{FI + GSSCP} \right) \times 100 \qquad (2.6)$$

Where FI represents phytolith morphotypes produced by plants belonging to functional groups typically indicative of closed forest habitats (palms, Zingiberales, woody dicots, and other forest indicators), and GSSCP stands for grass silica short cell phytoliths, diagnostic of the Poaceae family (including both open- and closed-habitat grasses).

To evaluate robustness of each *FI-t*$_{quadrat}$, we used bootstrap analysis (10,000 resampling with replacement, using the R function *bootstrap*) to assess 95% confidence intervals (CIs) for the phytolith proportions (Strömberg, 2009). We also determined the average *FI-t* ratio per site (*FI-t*$_{site}$) by calculating the mean of all *FI-t*$_{quadrat}$. In order to evaluate the similarity between each *FI-t*$_{quadrat}$ and the *FI-t*$_{site}$ we conducted pairwise comparisons of each *FI-t*$_{quadrat}$ with the confidence interval (CI) of the *FI-t*$_{site}$. The CI of the *FI-t*$_{site}$ was calculated by bootstrap analysis; note that to take into account differences in sample size between quadrats, for each pairwise comparison we calculated the CI of the *FI-t*$_{site}$ based on a population of size equal to the size of the compared *FI-t*$_{quadrat}$. We considered that *FI-t*$_{quadrat}$ is significantly different from *FI-t*$_{site}$ if it does not overlap with the CI of *FI-t*$_{site}$ (see Strömberg, 2009 for a discussion of this method). Likewise, we consider that *FI-t*$_{quadrat}$ values that do not differ significantly from *FI-t*$_{site}$ do not differ significantly from each other.

2.3.3.2 Gradients in vegetation structure

To test for a significant relationship between vegetation structure –as inferred from phytolith assemblages (*FI-t* ratio), and as measured in term of canopy openness (LAI) –and environmental gradients, we performed Spearman rank-order test of correlation between *FI-t* ratio and environmental variables (altitude, distance form water, and longitude), and between LAI and environmental variables. We used the function *cor.test* in the R package *stats*.

2.3.3.3 Gradients in vegetation composition

First, to test for a relationship between species/phytolith morphotype occurrence and spatial distribution of the quadrats we performed Mantel tests using the function *mantel* in the R package *vegan*. Matrices of species/phytolith morphotype abundance where transformed into matrices of occurrence using the R function *normalize.rows.* We then used the R function *vegdist*

to create matrices of dissimilarity (using the Jaccard's dissimilarity index) between quadrats. Matrices of distance between quadrats for both sites were obtained from planar coordinates using the R functions *distm* (R package *geosphere*) and *as.dist.*

To visualize trends in species composition along transects at each site, we performed Non-Metric Multidimensional Scaling (NMDS) on the species abundance datasets. NMDS is a commonly used ordination technique in community ecology that has several advantages, including that (1) it does not require an assumption of linear relationships between variables; and (2) it emphasizes relative, rather than absolute, differences between samples (quadrats, in this case) (Kruskal, 1964). As a result of these advantages, NMDS is often preferred to other ordination techniques (Gauch and Gauch Jr, 1982; Minchin, 1987). We first performed log transformation of the species abundance data to reduce data skewness. These data were transformed into a Bray-Curtis dissimilarity matrix (default settings) in R, meaning that no further data transformation was required. To determine the appropriate number of dimensions to consider for NMDS we use the R function *nmds.scree* which allows visualization of the plot of stress values versus the number of dimensions. Stress values for two dimensions were 0.07 and 0.03 for La Selva and Palo Verde respectively, low enough to suggest a good fit between the ordinations and the observed distances between quadrats. This was further confirmed by Monte Carlo randomization tests of stress values performed using the function *nmds.monte* ($p<0.05$), and by visualization of the relationship between the observed dissimilarities (Bray-Curtis) and the distances on the ordination space (and associated R^2 values) using the R function *stressplot*.

To determine whether species composition within sites was driven by environmental gradients in altitude, moisture, or canopy openness, we used linear regression analysis to test for

significant relationships between the quadrat scores on the first two axes of the NMDS and our environmental variables (elevation, distance from river, and LAI).

To visualize trends in phytolith morphotype composition along transects we performed NMDS on the phytolith assemblage counts (using functional groups, and morphotype counts), as well as cluster analysis (for Palo Verde only) to further test for the presence of significantly different groups of quadrats. We first performed a square root transformation of phytolith groups, and phytolith morphotypes count datasets to reduce data skewness, and obtained abundance matrices by calculating relative proportion of each phytolith functional group/morphotype in each assemblage. Then NDMS was performed following the same procedure as for the species abundance matrices (see above). The stress values for two dimensions were 0.078 and 0.03 for La Selva and Palo Verde respectively ($p>0.05$, and $p<0.01$, Monte Carlo randomization test of stress value) for NMDS on phytolith morphotype data; and 0.019 and 0.12 for La Selva and Palo Verde respectively ($p<0.01$, Monte Carlo randomization test of stress value) for NMDS on phytolith functional group data. We calculated the loading of the diagnostic phytolith functional groups/phytolith morphotype on the first two axes of the NMDS as well as their statistical significance using the R function *envfit* in R. In addition, for Palo Verde, we performed a cluster analysis using the Bray-Curtis dissimilarity matrix based on phytolith functional groups with the function *hclust* and using Ward's minimum variance method. We plotted the results of the cluster analysis on the NMDS ordination graph for better visualization. To determine whether phytolith composition within sites was driven by environmental gradients in altitude, moisture, or canopy openness, we used linear regression analysis to test for significant relationships between the quadrat scores on the first two axes of the NMDS and environmental variables (elevation, distance from river, and LAI). The influence

of environmental variables on phytolith assemblages in different quadrats at each of our sites (La Selva and Palo Verde) was further tested using additional multivariate ordination methods. To choose the appropriate ordination method (unimodal vs. linear response model) we first performed Detrended Correspondence Analysis (DCA) to evaluate the gradient length of the first DCA axis. The gradient length is a measure of "species" turnover, that is, the higher the gradient, the fewer the morphotypes/phytolith functional groups that are shared at both ends of the gradient. Based on the small axis length of 0.67 and 0.78 (for La Selva, and Palo Verde respectively), indicating low species turnover among quadrats, and therefore a linear model for species distribution along gradients, we chose Redundancy Analysis (RDA). RDA is an alternative to Canonical Correspondence Analysis (CCA) which uses a unimodal model of species distribution along gradients (Legendre and Anderson, 1999). RDA is basically a constrained version of PCA that allows regression of multiple response variables (phytolith functional groups in this case) on multiple explanatory variables (elevation, distance from the water, and LAI in this case). Through multiple linear regressions, a matrix of the fitted values of all response variables is generated. PCA is then performed on this matrix resulting in an ordination of the sites (quadrats) based on phytolith abundance data on axes representing combination of the environmental variables. We used both the matrix of raw phytolith morphotype counts as well as of counts of phytolith morphotypes classified into phytolith functional groups. In order to avoid giving a higher weight to variables with high magnitude values (e.g., distance from water) compared to variables with low magnitude values (e.g., LAI), we transformed environmental variable values so that they all ranged from 0 to 1, using the R function *data.trans*. We transformed the morphotype and phytolith functional group abundance data into matrices of morphotype relative abundances and functional group relative abundance

respectively, which we then square root transformed to correct for data skewness. Redundancy Analysis (RDA) with Monte Carlo permutation test of statistical significance (Legendre and Anderson, 1999) were conducted using the *vegan* package in R.

2.3.3.4 Gradients in diversity

To estimate per-quadrat and per-site diversity of plant communities at La Selva and Paleo Verde, we calculated α , β, and γ diversity using the *specnumber* and *d* functions in the R package *vegan*. To investigate the relationship between floristic α diversity and phytolith diversity (number of morphotype per sample) we performed linear regression analysis using the *lm* function of the R package *stats*.

2.3.3.5 Difference in phytolith assemblage composition between the two sites

To test whether the rainforest and dry forest at La Selva and Palo Verde, respectively, can be distinguished based on their phytolith composition we performed NMDS on the phytolith functional group counts of all quadrats in the two sites, as well as cluster analysis. For both NMDS and cluster analysis we followed the same procedure as described in section 3.4.3. For the NMDS analysis we used 2 dimensions given a stress value of 0.039.

2.4 RESULTS

2.4.1 *Soil and plant phytolith yield*

At La Selva, all soil assemblages yielded well preserved phytolith assemblages. At Palo Verde, all soil assemblages except for quadrat 18 (which was unproductive and therefore excluded from this study) yielded phytoliths with varying degree of preservation. In particular, phytolith preservation was poor in quadrats 19, 22, and 23.

At Palo Verde, plant phytolith yield varies widely among different species. In some of the quadrats (e.g., 15, 16, 17, and 21) a high percentage of the total tree biomass is represented by species that are either non-producers or low phytolith producers (table 1). In addition, in some quadrats (e.g., 17, 20, 25), a high proportion of tree biomass corresponds to evergreen species (table 1), which produce a lower amount of leaf litter than deciduous species. The phytolith yield of these species is potentially lower compared to deciduous species.

2.4.2 *Phytolith and vegetation data*

2.4.2.1 *FI-t* ratio and robustness of FI/t ratio differences between soil phytolith

At La Selva, $FI\text{-}t_{site}$ (average *FI-t* ratio) is 98.5±1.4%. All confidence intervals are narrow, (Fig. 2.2a) and all $FI\text{-}t_{quadrat}$ fall within the CI of $FI\text{-}t_{site}$, except for one quadrat. At Palo Verde, the $FI\text{-}t_{site}$ is 58.8±21.5%, with a minimum FI-t ratio of 19.2% (Q15) and a maximum *FI-t* ratio of 84.6% (Q23). Low *FI-t* ratios are also found in quadrats 16, and 17 (37.8%, and 41.5%), intermediate FI-t ratios are found in quadrats 20, 21, 22, and 25 (59.1%, 62.1%, 62.75%, and 55.5%), and the highest FI-t ratios are found in quadrats 19, 23, and 24 (84.4%, 81.1%, and 84.6% respectively) (Fig. 2.2b). Only the $FI\text{-}t_{quadrat}$ values of four quadrats (quadrats 20, 21, 22, and 25) falls within the CI of $FI\text{-}t_{site}$.

2.4.2.2 Gradients in vegetation structure

At La Selva, LAI varies along the transect with values ranging from 4.6 to 5.7 (average = 5.1±0.4). LAI variation does not seem to follow any gradient in elevation, distance from the water, or longitude ($p>0.05$, Spearman rank-order correlation test). Similarly, FI-t variation does not follow any gradient ($p>0.05$, Spearman rank-order correlation test), (Fig. 2.3a). At Palo Verde, LAI (average = 2.9±0.8) is lower than at La Selva, and ranges from 1.8 to 3.9. Along the

first transect (T1), LAI decreases with distance from the water, and longitude ($p<0.10$, Spearman rank-order correlation test), but not with elevation ($p>0.05$, Spearman rank-order correlation test). Along T2 LAI shows no clear pattern with distance from the water, longitude, or elevation ($p>0.05$, Spearman rank-order correlation test), (Fig. 2.4a). The FI-t ratio shows a similar pattern to LAI along T1 –specifically it decreases with distance from the water ($p=0.05$, Spearman rank-order correlation test), longitude ($p=0.05$, Spearman rank-order correlation test), and elevation ($p<0.01$, Spearman rank-order correlation test) –whereas it does not seem to follow a clear pattern along T2 ($p=0.05$, Spearman rank-order correlation test), (Fig. 2.4a).

While the *FI-t* ratio and LAI trends at LA Selva are negatively correlated ($p<0.01$, Spearman rank-order correlation test), (Fig. 2.3a), no statistically significant correlation was observed at Palo Verde ($p=0.05$, Spearman rank-order correlation test), although they both decrease along T1 (Fig. 2.4a).

2.4.2.3 Gradients in vegetation composition

Floristic and phytolith compositional gradients (or lack thereof) at La Selva are illustrated in Figure 2.3b. We found no correlation between plant species composition of the quadrats and their inter-site distances (Mantel test, $p > 0.05$). This is confirmed by the NMDS ordination of La Selva quadrats according to plant species composition ($p<0.01$), which shows no clear clusters of quadrats (ordination plot not shown here), as well as no significant relationship between the first two NMDS axes and environmental variables (elevation, distance from the water, longitude, and LAI), (see also Fig. 2.3b). Similarly, we found no correlation between phytolith composition of the quadrats and their inter-site distances (Mantel test, $p > 0.05$). NMDS ordination of La Selva quadrats according to phytolith morphotype composition ($p<0.01$), produced similar results (ordination plot not shown here) to NMDS ordination based on plant species composition, with

not clear clusters of quadrats, and no relationship between NMDS axis 1 and 2, and environmental variables ($p>0.05$), (see also Fig. 2.3b). Similarly, NMDS ordination of la Selva quadrats based on phytolith functional group composition (Fig. 2.5a) shows that quadrats distribute randomly across the ordination space. Phytolith assemblage differences among quadrats are mainly driven by forest indicator ("Other FI", significant, positive loading on NMDS 1, $p<0.01$, Fig. 2.5a), and palm ("Palm", significant, negative loading on NMDS axis 2, $p<0.01$, Fig. 2.5a) phytoliths, as well as PACMAD and other closed habitat indicator grass (significant, negative loadings on NMDS axis 2, $p<0.01$, Fig. 2.5a). However, these compositional differences do not follow environmental gradients. Indeed, regression analyses of sample (quadrat) scores along the first two axes of the NMDS ordination with each one of the environmental variables (altitude, distance from the water, LAI) show no significant relationships between NMDS axes 1and 2 and environmental gradients ($p>0.05$). In addition, RDA analysis shows no significant relationships between the composition of phytolith assemblages and environmental variables ($p>0.05$, see supplementary material for RDA ordination plot).

In contrast, at Palo Verde, plant species composition (occurrence) of quadrats is significantly related to their inter-site distances (geographic position) (Mantel test, $p < 0.01$), meaning that quadrats that cluster together in the geographic space are significantly more similar to each other than to those that are further away. This is confirmed by NMDS ordination, and cluster analysis of Palo Verde quadrats according to their phytolith functional group composition. Both analyses allow us to distinguish two groups Palo Verde quadrats that are located towards opposite ends of axis 1 of the NMDS ordination (Fig. 2.5b). The first group includes quadrats along T1 that are located on a rocky hill slope; the second group includes

quadrats from both T1 (19-20) and T2 (21-25). The main phytolith functional groups driving difference between assemblages are forest indicators ("Other FI", significant positive loading on NMDS axis 1, $p<0.01$), grasses of the tribes Pooideae and Panicoideae, and of the PACMAD clade (respectively "POOID-D", "Panicoid", and "PACMAD"; significant negative loadings on NMDS axis 1, $p<0.01$), as well as of grasses of the tribe Chloridoideae ("Chloridoid", significant negative loading on NMDS axis 2, $p<0.01$). Further, linear regression of axis 1 and 2 of the NMDS ordination ($p<0.01$) with environmental variables show a positive relationship between NMDS axis 1 and elevation ($p<0.05$), and NMDS axis 2 and longitude ($p<0.01$). Similar results were obtained using phytolith morphotype composition for the NMDS ordination.
Trends in floristic composition as estimated by values on NDMS axis 1 and 2, do not correspond to trend in phytolith composition of quadrats along environmental gradients at either site (Fig. 2.3b, and 2.4b).

2.4.2.4 Gradients in diversity

Species richness at La Selva is high, with more than 2,077 plant species in total; Palo Verde hosts about 619 plant species (Organization for Tropical Studies, 2019). In the La Selva and Palo Verde quadrats, we identified respectively 95 (78 dicot trees, 8 palms, 2 grasses, and 7 other monocot species) and 55 (49 dicot trees, 1 palm, 4 grasses, and herbaceous dicot species) plant species. At both the quadrat and site level species diversity is higher at La Selva (α=18.71, β= 5.08, γ= 95) than at Palo Verde (α= 11.4, β=4.03 γ = 55). Phytolith morphotype diversity is lower at La Selva (108 morphotypes of which 43 are GSSCP) than at Palo Verde (127 morphotypes of which 55 are GSSCP).

At both sites, floristic diversity (α) and phytolith diversity (number of morphotypes) do not follow any environmental gradient (fig 2.2c and 2.3c). Furthermore, linear regression

analysis suggests that there is no relationship between plant and phytolith diversity, with an increase or decrease in α not necessarily corresponding to an increase or decrease in the number of phytolith morphotypes ($p>0.05$).

2.4.2.5 Overall differences between vegetation types reflected in phytolith assemblages

Both cluster analysis and NMDS separate La Selva and Palo Verde phytolith assemblages well, based on the relative abundance of different phytolith functional groups (Fig. 2.6a and 2.6b). All the functional groups (except for Zingiberales, $p>0.05$) have significant loadings on the NMDS ordination axes 1 and 2 ($p<0.01$) (Fig. 2.6a). La Selva phytolith assemblages are separated from those of Palo Verde mainly along axis 1 of the NMDS ordination; on this axis the highest loadings correspond to "Palm" morphotypes ($p<0.01$, negative loading on NMDS axis 1), "Chloridoid" and "Closed habitat grasses" ($p<0.01$, positive loading on NMDS axis 1). On axis 2 the highest loadings correspond to "Other FI" (p,0.010, negative loading), and grass morphotypes ("PACMAD", "Panicoid", "POOID-ND", and "POOID-D", $p<0.01$, positive loadings on NMDS axis 2); these latter morphotypes seem to characterize differences in phytolith assemblages within Palo Verde quadrats (specifically between the two quadrat groups highlighted by cluster analysis in Fig. 2.5b) rather than between the two sites. In fact, in the NMDS ordination space (Fig. 2.6a), while La Selva phytolith assemblages from a compact group (toward the negative end of NMDS 1), Palo Verde quadrats are more spread (along NMDS axis 2), indicating more variation between quadrats within La Selva than within Palo Verde.

2.5 DISCUSSION

2.5.1 *Soil and plant phytolith yield and their impact on production and taphonomic biases*

Differences in soil phytolith assemblage preservation between La Selva and Palo Verde might be explained by differences in soil type, as well as environmental and climatic conditions. At Palo Verde, quadrats that yielded poorly preserved phytolith assemblages are located at the bottom of a porous limestone hill slope, where soils (Inceptisols) are poorly developed (Gillespie et al., 2000; Powers et al., 2009). It is possible that the presence of small limestone detritus in the soil coming from uphill results in high pH, which in turn could negatively affect phytolith preservation (e.g., Bartoli and Wilding, 1980; Fraysse et al., 2006). The occurrence of conditions unfavorable to phytolith preservation might also contribute to the higher percentage of GSSCP found in these assemblages, with preferential dissolution of less compact and well silicified FI phytoliths (Wilding and Drees, 1974; Bartoli and Wilding, 1980; Cabanes et al, 2011). In contrast, La Selva quadrats are located on well-developed Oxisols (Kleber et al., 2007) derived from Andesitic lava flows (Alvarado Induni, 1990), which tend to have a pH at or below neutral (Weil and Brady, 2016), which is consistent with the optimal phytolith preservation encountered at this site.

2.5.2 *Phytolith and vegetation data*

2.5.2.1 Gradients in vegetation structure

At La Selva, the *FI-t* ratio and the LAI of quadrats along the forest transect vary in a nonlinear way, and not according to variations in altitude, distance from the water, or longitude. Further, observed canopy openness (LAI) along the transect does not match estimated canopy openness (*FI-t*). Several factors might explain these mismatches. First, it is possible that the

narrow range of variation of both *FI-t* ratio and LAI is likely due to the lack of substantial and detectable structural variation within La Selva rainforest. Indeed, all measured *FI-t* values are indicative of a dense closed-canopy forest, and all LAI measured values fall within the range of values reported by other authors for late successional stage rainforests (e.g., Murphy and Lugo, 1986; Kalácska et al., 2004). Second, we hypothesize that soil phytolith assemblages record an extra-local and/or long-term vegetation signal (as opposed to a super-local, short term signal) and thus, at the scale used in this study, the *FI-t* ratio is not sensitive to variations in canopy openness. An extra-local signal might be the result of lateral transport of phytoliths (spatial mixing); a long-term signal might be due to vertical translocation and mixing of older and younger soil (temporal averaging). Consequently, the phytolith signal would represent a larger forest area, a longer amount of time, or both. The main factor favoring temporal mixing of the phytolith signal is bioturbation, which is extremely high in tropical rainforest soils (Weil and Brady, 2016). Lateral mixing by environmental factors such as fire and wind should be limited in this dense wet forest environment, but alluvial transport, and transport by herbivores are possible (Fredlund and Tieszen, 1992).

At Palo Verde, both the *FI-t* ratio and the LAI of quadrats decrease with increasing distance to the water (which increases northward), and from west to east along the first transect (T1) but not along the second transect (T2, Fig. 2.4a). It is possible that, along T1, higher moisture in the quadrats located closer to the water (Quírico Jímenez et al., 2016) favors denser vegetation (thus higher LAI and FI-t), whereas quadrats located at greater distance from the water experience drier conditions (Quírico Jímenez et al., 2016) resulting in less dense vegetation cover (thus lower LAI and *FI-t* ratio). Although *FI-t* ratio and LAI follow a similar trend along T1, LAI varies much less than the in *FI-t* ratio. Thus, whereas the range of measured LAI values (1.8 to

3.9) overlaps with previously measured values in Palo Verde dry forest (Kalácska et al., 2005), the range of variation of the *FI-t* ratio (19.16 to 84.45%) encompasses different vegetation types, from savanna to closed forest. An even more important mismatch between LAI and *FI-t* ratio is observed along T2, where LAI and *FI-t* show opposite trends. Temporal averaging of the soil phytolith assemblages is unlikely to be the cause of this mismatch, because bioturbation (and vertical translocation of the soil) is limited in seasonally dry conditions such as at Palo Verde dry forest. On the other hand, lateral phytolith transport is likely to take place at Palo Verde and would be favored by strong winds and seasonal fires (Powers et al., 2009). Such transport would result in spatially mixed phytolith assemblages which might provide an extra-local vegetation signal, biasing the *FI-t* ratio in some quadrats. In addition, other factors might contribute to the mismatch between phytolith *FI-t* ratio and LAI, such as 1) differential phytolith production by tree species; 2) differential litter production (leading to differential contribution to the soil phytolith pool) by deciduous and evergreen tree species; and 3) variation in grass cover among quadrats. All these factors could act to bias the *FI-t* ratio in several ways. For example, overall FI phytolith production might be lower in quadrats where non-producer or low producer tree species are abundant (e.g. quadrat 15, 16, 17, and 21), resulting in underestimation of the *FI-t* ratio. Evergreen species (which are dominant in quadrat 16, 20, and 25), might contribute less biomass to the leaf litter than deciduous species, resulting in lower accumulation of phytolith in the soils, and consequent underestimation of the *FI-t* ratio (especially in quadrats with high proportion of evergreen species). Although grass cover was overall very scarce at Palo Verde at the time of the field season (dry season), the rocky limestone slopes of Palo Verde hills have been reported as a suitable habitat for grasses (Quírico Jiménez et al., 2016) are located on these slopes. The quadrats with the lower FI-t values but high LAI (quadrat 15, 16, and 17) are located

on these slopes. It is therefore possible that the presence of a dense dry-adapted grass understorey in these quadrats (rather than open canopy structure) is responsible for the low *FI-t* ratio values measured from phytolith assemblages. Quadrat 23 presents an opposite pattern, with one of the highest *FI-t* values and one of the lowest LAI values. Most tree species in this quadrat are deciduous. Therefore, it is likely the LAI in this quadrat has been biased toward low values that by the presence of deciduous species

Overall, despite a mismatch in detailed pattern, trends in *FI-t* ratio and LAI along transects at La Selva and Palo Verde, suggest that spatial sampling for soil phytolith samples allows detection of variation (or lack thereof) in forest structure across rainforest and dry forest habitats. However, temporal averaging and possibly some lateral mixing of phytolith assemblages in rainforest soils appear to result in lower spatial resolution of the phytolith record compared to dry forest soils. Therefore, in this forest type, a single sample is representative of the overall vegetation structure, as laterals samples should provide a very similar signal. In the dry forest, while spatial heterogeneity in forest structure is captured by the *FI-t* signal, differences in *FI-t* ratio between samples appear to exaggerate differences in canopy openness. Furthermore, the *FI-t* ratio is influenced by other factors that are not related to canopy openness, such as phytolith production biases, and grass cover in the forest understory (see also Dunn et al., 2015). Nevertheless, extreme variations in the *FI-t* ratio, such as across Palo Verde transects, are indicative of some degree of heterogeneity within the dry forest habitat and suggests that spatial sampling is necessary to reconstruct such heterogeneity.

2.5.2.2 Gradients in vegetation composition

Floristic composition of La Selva quadrats is not related to geographic distance between quadrats or to any of the measured environmental variables. This apparently random pattern of

species abundance and distribution might be due to limitations of our sampling scheme (e.g., size and length of quadrats and transect), which might be insufficient for capturing the relevant scale of species distribution patterns. Similarly, phytolith composition at La Selva is not related to geographic distance between quadrats or environmental variables. This is not surprising, because even in the presence of a floristic gradient, we would expect any taxonomic signal in the phytolith assemblages to be at least obscured to some degree by multiplicity and redundancy in phytolith production by different species (Rovner, 1971). Notwithstanding, multiplicity and redundancy alone cannot explain the lack of correspondence between floristic and phytolith composition of the quadrats at La Selva, specifically the variation in abundance of particular, taxonomically diagnostic phytolith morphotypes among quadrats in La Selva. For example, we observed a high number of palm phytoliths ("echinate spheres", and "hats") in all quadrats, but with significant differences in relative abundance between quadrats. Palm tree abundance is also variable across quadrats. However, we found no relationship between palm phytolith relative abundance and the number of palm trees present in the quadrat. We hypothesize that temporal averaging and some spatial mixing of soil phytolith assemblages at La Selva explain the mismatch between observed vegetation and phytolith assemblage composition.

At Palo Verde, quadrats differ in measured floristic composition along a longitudinal and an altitudinal gradient which result in a mosaic of plant species that differ both within and between transects. Elevation explains similarities in floristic composition between most quadrats, such as: 1) between quadrats along T2 located on the higher part of the hills, on flat ground (quadrat 21, 22, and 23); 2) quadrats along T1 on the slopes of the limestone hills (quadrat 15, and 16); and 3) quadrats along T1 (quadrat 19, and 20), and T2 (quadrat 25) located on the foothill plain. The two remaining quadrats in T1 (quadrat 17) and T2 (quadrat 24) have

similar floristic composition but are not geographically related. Quadrat 17 is located along T1 at intermediate elevation and distance from the remaining quadrats, which could explain its "intermediate" floristic composition. Quadrat 24, although only a few hundreds of meters from quadrat 25, is the only quadrat that forms a regeneration plot located on seasonally flooded ground. it is likely that this quadrat represents an outlier, but its floristic similarity with quadrat 17 is surprising. In this context, it is important to note, that our floristic census only takes into account tree species; hence, other floristic differences (such as in grass and other monocots species) are nor captured by our analysis. Therefore, floristic similarities such as the one observed between quadrat 17 and 24 need to be interpreted with caution.

Phytolith composition of Palo Verde quadrats does not follow environmental gradients. However, phytolith assemblages of three quadrats (15, 16, and 17) differ from all other assemblages in that they are dominated by GSSCP. As previously discussed, in these quadrats non-producers or low phytolith producer species are particularly abundant. Further, alkaline soils potentially prevent optimal preservation of FI phytoliths. These two factors are likely to bias the FI-t ratio toward low values. In addition, grasses are likely seasonally abundant in the understorey of these quadrats, and the fact that our floristic census, unlike phytolith counts, does not take include grass taxa might also explain the lack of correspondence between floristic and phytolith compositional trends along Palo Verde quadrats.

Overall, trends in floristic and phytolith assemblage composition along transects at La Selva and Palo Verde suggest that the spatial resolution of the phytolith record varies according to the vegetation type. In rainforest habitats, lateral sampling is not effective to reconstruct compositional gradients, because the spatial resolution of the phytolith signal is limited by temporal averaging and possible lateral mixing of phytolith assemblages. In dry forest habitats,

spatial sampling can help detect overall heterogeneity. However, it is not effective for reconstructing detailed gradients in vegetation composition, unless it is coupled with other proxies for canopy openness, (e.g. stable carbon isotopes, phytolith undulation index; e.g., Dunn et al., 2015) which would allow separation of the compositional (grass cover) and structural (canopy openness) components of the *FI-t* ratio.

2.5.2.3 Gradients in diversity

The higher floristic diversity measured at La Selva, compared to Palo Verde is consistent with published surveys (Murphy and Lugo, 1986). In contrast, phytolith morphotype follows the opposite trend, with higher morphotype diversity at Palo Verde, suggesting that phytolith morphotype diversity is not indicative of taxonomic diversity. This is not surprising, given the strong limitation on taxonomic resolution of phytoliths imposed by multiplicity and redundancy in phytolith production (e.g. Piperno, 1988). Therefore, we argue that the absolute number of phytolith morphotype present in an assemblage should not be used directly to infer taxonomic diversity (e.g., Hyland et al., 2015).

2.5.2.4 Overall differences between vegetation types reflected in phytolith assemblages

Overall, La Selva rainforest is characterized by lower canopy openness (higher LAI), higher floristic diversity, more homogeneous structure (LAI variation, edaphic, and topographic conditions) and floristic composition (random versus geographic pattern of variation in floristic composition) than Palo Verde dry forest. These structural and compositional differences between La Selva and Palo Verde are reflected in soil phytolith assemblages.

At La Selva, the presence of a more closed canopy compared to Palo Verde is reflected in overall higher *FI-t* ratios of phytolith assemblages. In addition, the *FI-t* ratio is homogeneous across La

Selva quadrats, while it varies significantly across quadrats at Palo Verde. This pattern broadly reflects the more homogeneous canopy structure (LAI) at La Selva compared to Palo Verde. However, structural heterogeneity at Palo Verde is exaggerated by the *FI-t* ratio, which not only reflects heterogeneity in canopy openness but might also indicate compositional differences in the forest understory of different quadrats –in particular, very low *FI-t* ratios might be encountered when understorey grass cover is dense. Variations in the FI-t across the Palo Verde landscape can also be explained by taphonomic biases (i.e. low phytolith preservation in alkaline soils), and heterogeneity in soil type. In addition, the presence of a mix of deciduous and evergreen tree species (likely differing in biomass contribution and total phytolith production) is also potentially reflected in a more heterogeneous *FI-t* at Palo Verde, compared to La Selva.

Compositional differences between La Selva and Palo Verde vegetation are also reflected in phytolith assemblages. First, palm phytolith abundance is overall high at La Selva, while palm morphotypes were found in significantly lower abundance and only in some of Palo Verde quadrats. Second, grass phytoliths are overall very abundant at Palo Verde whereas they are rare (<5%) in all La Selva phytolith assemblages. Third, at Palo Verde, we recorded the unique presence of morphotypes that we assigned to the group Palm/Zingiberales. Although we could not determine with certainty the functional group affinity of these morphotypes (see section 3.3.3), they were extremely rare in La Selva assemblages (such that they were only observed by additional slide scanning). Lastly, phytolith assemblages are compositionally more heterogeneous across Palo Verde quadrats than across La Selva, as shown by the wide variation observed in GSSCP abundance across Palo Verde quadrats, as well as by the greater spread of Palo Verde quadrats in the NMDS ordination space compared to for La Selva.

Our results show that, in addition to the overall *FI-t* ratio, and relative abundances of phytolith functional groups (corresponding to plant functional types), spatial variation in vegetation structure and composition differs between La Selva and Palo Verde. Lateral sampling is therefore essential to capture such variation and to differentiate these two vegetation types.

2.5.3 *Guidelines for interpreting the fossil record*

For ideal characterization of the past vegetation, and comparisons of two or more coeval fossil localities we suggest that several phytolith samples should be collected horizontally along a stratigraphic level in each locality. In addition, the study of overall phytolith assemblage composition and *FI-t* ratio should be complemented by the analysis of spatial patterns of variation in phytolith composition and *FI-t* ratio along vertical transects. Our study points to a complex interplay between phytolith production and taphonomy in shaping the phytolith assemblage signal; nevertheless, our results indicate that the more heterogenous habitats are likely to also produce a heterogeneous phytolith signal (varying proportion of different functional groups and *FI-t* ratio values) along a stratigraphic level. Conversely, homogeneous habitats (i.e., rainforests) are likely to produce a homogeneous phytolith signal, at least over relatively short distances (<2 km). Therefore, spatial variation in the phytolith signal in the fossil record can be reasonably interpreted as representing a heterogenous habitat, and spatial sampling is instrumental for the purpose of recording habitat heterogeneity.

2.5.4 *Future directions*

Our study is the first to characterize and compare phytolith assemblage across vegetation transects in different Neotropical vegetation types in a way that is applicable to the deep time phytolith record. Our results suggest that phytolith assemblages have a good potential for paleoenvironmental reconstructions of rainforests and dry forests in deep time, and that these two vegetation types can be distinguished based on phytolith analysis. However, it remains to be tested to what degree modern phytolith assemblages from other Neotropical vegetation types overlap with phytolith assemblages from rainforest and dry forest. In particular, although no modern soil phytolith assemblages from Neotropical savannas have been studied, phytolith assemblages from a range of Central African grassland vegetation types with 30-40% to 60-70% tree cover (Novello et al., 2012) show GSSCP percentages within the range observed at Palo Verde. These data are not directly comparable to our data, because they were collected and analyzed using different methodologies, and in a different geographic region. However, they indicate that detailed study of phytolith assemblages from Neotropical savanna soils is needed in order to establish guidelines for distinguishing this vegetation type from dry forest in the fossil record. To allow inter-study comparison, future studies should be conducted using the same sampling methodology and classification system described herein. Furthermore, multiple sites along transect within and among vegetation types should be studied to account for a wider range of possible variation in structure and composition of observed vegetation and phytolith assemblages. Additional environmental and climatic variables could be taken into account such as depositional environment, soil type, mean annual precipitation and precipitation seasonality, mean annual temperature, as well remote sensing data such as MODIS (Moderate Resolution imaging Spectroradiometer) data to better describe inter-site difference in term of potential

taphonomic biases, climate, and vegetation. In addition, a larger dataset including many vegetation types would allow building predictive models for phytolith assemblages including ancient non-analogue systems. Our study, and future effort to study the modern phytolith record therefore has the potential of greatly improving the resolution and accuracy of paleoenvironmental reconstructions using phytoliths.

ACKNOWLEDGMENTS

We are grateful to the OTS staff of La Selva and Palo Verde Biological stations for helping with field work logistics; assistance with research and export permits (Ministry of the Environment and Energy research permit # 006-2016-INV-ACAT; export permits # DGVS-076-2016, and DGVS-082-2016) was provided by MINAE and Francisco Campos. We thank the Burke Museum of Natural History and Culture (UWBM), Ron Eng, Katie Anderson, UWBM Botany Collections (WTU herbarium), David Giblin and Soojeong Shin, at the University of Washington (Seattle) for helping with sample cataloguing. We also thank Claire Grant and Kailyn Zard, who helped with plant sample processing; and x, y, and z for comments on an earlier version of this manuscript. Fieldwork and sample processing were supported by the Organization for Tropical Studies, WRF-Hall Foundation (University of Washington), Philosophical Society of America, Geological Society of America, Paleontological Society, and Quaternary Research Center (grants awarded to C.C.), and the National Science Foundation (NSF EAR-1253713 and EAR-1349530 awarded to C.A.E.S).

REFERENCES

Aleman, J., Leys, B., Apema, R., Bentaleb, I., Dubois, M.A., Lamba, B., Lebamba, J., Martin, C., Ngomanda, A., Truc, L., Yangakola, J.-M., Favier, C., Bremond, L., Woods, K., 2012. Reconstructing savanna tree cover from pollen, phytoliths and stable carbon isotopes. Journal of Vegetation Science 23, 187-197.

Aleman, J.C., Canal-Subitani, S., Favier, C., Bremond, L., 2014. Influence of the local environment on lacustrine sedimentary phytolith records. Palaeogeography, Palaeoclimatology, Palaeoecology 414, 273-283.

Alexandre, A., Meunier, J.-D., Colin, F., Koud, J.-M., 1997. Plant impact on the biogeochemical cycle of silicon and related weathering processes. Geochimica et Cosmochimica Acta 61, 677-682.

Alvarado Induni, G.E., 1990. Características geólogicas de la Estación Biológica La Selva, Costa Rica. Tecnología en marcha 10, 11-22.

Auerbach, M., Shmida, A., 1987. Spatial scale and the determinants of plant species richness. Trends in Ecology & Evolution 2, 238-242.

Badgley, C., Smiley, T.M., Terry, R., Davis, E.B., DeSantis, L.R.G., Fox, D.L., Hopkins, S.S.B., Jezkova, T., Matocq, M.D., Matzke, N., McGuire, J.L., Mulch, A., Riddle, B.R., Roth, V.L., Samuels, J.X., Stromberg, C.A.E., Yanites, B.J., 2017. Biodiversity and Topographic Complexity: Modern and Geohistorical Perspectives. Trends Ecol Evol 32, 211-226.

Barboni, D., Bremond, L., Bonnefille, R., 2007. Comparative study of modern phytolith assemblages from inter-tropical Africa. Palaeogeography, Palaeoclimatology,

Palaeoecology 246, 454-470.

Bartoli, F., Wilding, L.P., 1980. Dissolution of biogenic opal as a function of its physical and chemical properties. Soil Science Society of America Journal 44, 873-878.

Bremond, L., Alexandre, A., Hély, C., Guiot, J., 2005. 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.

Burnham, R.J., Wing, S.L., Parker, G.G., 1992. The refelction of deciduous forest communities in leaf litter - implications for autochthnonous litter assemblages from the fossil record. Paleobiology 18, 30-49.

Cabanes, D., Weiner, S., Shahack-Gross, R., 2011. Stability of phytoliths in the archaeological record: a dissolution study of modern and fossil phytoliths. Journal of Archaeological Science 38, 2480-2490.

Charney, N., and Record, S., 2012. vegetarian: Jost Diversity Measures for Community Data. R package version 1.2. https://CRAN.Rproject.org/package=vegetarian

Chen, S.T., Smith, S.Y., 2013. Phytolith variability in Zingiberales: A tool for the reconstruction of past tropical vegetation. Palaeogeography, Palaeoclimatology, Palaeoecology 370, 1-12.

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

Cramer, M.J., Willig, M.R., 2005. Habitat heterogeneity, species diversity and null models. Oikos 108, 209-218.

Crifò, C., Strömberg, C.A., in pres. Small-scale spatial resolution of the soil phytolith record in a

Neotropical rainforest and a dry forest in Costa Rica: applications to the deep-time fossil phytolith record. Palaeogeography, Palaeoclimatology, Palaeoecology.

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

Finarelli, J.A., Badgley, C., 2010. Diversity dynamics of Miocene mammals in relation to the history of tectonism and climate. Proceedings of the Royal Society of London B: Biological Sciences 277, 2721-2726.

Fraser, R.H., 1998. Vertebrate species richness at the mesoscale: relative roles of energy and heterogeneity. Global Ecology & Biogeography Letters 7, 215-220.

Fraysse, F., Pokrovsky, O.S., Schott, J., Meunier, J.-D., 2006. Surface properties, solubility and dissolution kinetics of bamboo phytoliths. Geochimica et Cosmochimica Acta 70, 1939-1951.

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

Fredlund, G.G., Tieszen, L.L., 1997. Phytolith and carbon isotope evidence for late Quaternary vegetation and climate change in the southern Black Hills, South Dakota. Quaternary Research 47, 206-217.

Gauch, H.G., Gauch Jr, H.G., 1982. Multivariate analysis in community ecology. Cambridge University Press.

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

Harris, E.B., Strömberg, C.A., Sheldon, N.D., Smith, S.Y., Vilhena, D.A., 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.

Hijmans, R.J., 2017. geosphere: Spherical Trigonometry. R package version 1.5-7. https://CRAN.R-project.org/package=geosphere

Holdridge, L.R., 1967. Life zone ecology, Rev. ed. ed. San Jose, Costa Rica : Tropical Science Center, San Jose, Costa Rica.

Hodson, M.J., White, P.J., Mead, A., Broadley, M.R., 2005. Phylogenetic Variation in the Silicon Composition of Plants. Annals of Botany 96, 1027-1046.

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

Janis, C.M., Damuth, J., Theodor, J.M., 2004. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. Palaeogeography, Palaeoclimatology, Palaeoecology 207, 371-398.

Kalácska, M., Sanchez-Azofeifa, G.A., Calvo-Alvarado, J.C., Quesada, M., Rivard, B., Janzen, D.H., 2004. Species composition, similarity and diversity in three successional stages of a seasonally dry tropical forest. Forest Ecology and Management 200, 227-247.

Kalácska, M., Calvo-Alvarado, J.C., Sánchez-Azofeifa, G.A., 2005. Calibration and assessment of seasonal changes in leaf area index of a tropical dry forest in different stages of succession. Tree Physiology 25, 733-744.

Kerns, B.K., Moore, M.M., Hart, S.C., 2001. Estimating forest-grassland dynamics using soil phytolith assemblages and d13C of soil organic matter. Écoscience 8, 478-488.

Kleber, M., Schwendenmann, L., Veldkamp, E., Rößner, J., Jahn, R., 2007. Halloysite versus gibbsite: Silicon cycling as a pedogenetic process in two lowland neotropical rain forest

soils of La Selva, Costa Rica. Geoderma 138, 1-11.

Kohn, M.J., Fremd, T.J., 2008. Miocene tectonics and climate forcing of biodiversity, western United States. Geology 36, 783-786.

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

Legendre, P., Anderson, M.J., 1999. Distance-Based Redundancy Analysis: Testing Multispecies Responses in Multifactorial Ecological Experiments. Ecological Monographs 69, 1-24.

MacArthur, R.H., MacArthur, J.W., 1961. On Bird Species Diversity. Ecology 42, 594-598.

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

Minchin, P.R., 1987. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 69, 89-107.

Mitani, N., Ma, J.F., 2005. Uptake system of silicon in different plant species. Journal of Experimental Botany 56, 1255-1261.

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

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

Organization for Tropical Studies, 2019. https://tropicalstudies.org/

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2018. Package 'vegan'. Community ecology package, version 2.5-3. https://CRAN.R-project.org/package=vegan

Pearsall, D., 2000. Paleoethnobotany: A Handbook of Procedures Academic Press. San Diego, Calif.

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

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

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

Postek, M.T., 1981. The occurrence of silica in the leaves of Magnolia grandiflora L. Botanical Gazette 142, 124-134.

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

Prentice, I.C., 1985. Pollen Representation, Source Area, and Basin Size: Toward a Unified Theory of Pollen Analysis. Quaternary Research 23, 76-86.

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

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

Runge, F., 1999. The opal phytolith inventory of soils in central Africa —quantities, shapes, classification, and spectra. Review of Palaeobotany and Palynology 107, 23-53.

Stein, A., Beck, J., Meyer, C., Waldmann, E., Weigelt, P., Kreft, H., 2015. Differential effects of

environmental heterogeneity on global mammal species richness. Global Ecology and Biogeography 24, 1072-1083.

Strömberg, C.A.E., 2002. The origin and spread of grass-dominated ecosystems in the late Tertiary of North America: preliminary results concerning the evolution of hypsodonty. Palaeogeography Palaeoclimatology Palaeoecology 177, 59-75.

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

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 of the United States of America 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., Dunn, R.E., Madden, R.H., Kohn, M.J., Carlini, A.A., 2013. Decoupling the spread of grasslands from the evolution of grazer-type herbivores in South America. Nature communications 4, 1478-1478.

Strömberg, C.A.E., Di Stilio, V.S., Song, Z., 2016. Functions of phytoliths in vascular plants: an evolutionary perspective. Functional Ecology 30, 1286-1297.

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

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M.C., Schwager, M., Jeltsch, F.,

2004. Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. Journal of Biogeography 31, 79-92.

Tibshirani, R. and Leisch, F., 2017. bootstrap: Functions for the Book "An Introduction to the Bootstrap". R package version 2017.2.

Trembath-Reichert, E., Wilson, J.P., McGlynn, S.E., Fischer, W.W., 2015. Four hundred million years of silica biomineralization in land plants. Proceedings of the National Academy of Sciences 112, 5449-5454.

Venables, W.N. and Ripley, B.D., 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Weil, R., Brady, N.C., 2016. The nature and properties of soils: Pearson new international edition. Pearson Higher Ed.

Wilding, L.P., Drees, L.R., 1974. Contributions of forest opal and associated crystalline phases to fine silt and clay fractions of soils. Clays and Clay Minerals 22, 295-306.

Willig, M.R., Kaufman, D.M., Stevens, R.D., 2003. Latitudinal Gradients of Biodiversity: Pattern, Process, Scale, and Synthesis. Annual Review of Ecology, Evolution, and Systematics 34, 273-309.

Wing, S.L., Hickey, L.J., Swisher, C.C., 1993. Implications of an exceptional fossil flora for Late Cretaceous vegetation. Nature 363, 342.

Wing, S.L., Strömberg, C.A.E., Hickey, L.J., Tiver, F., Willis, B., Burnham, R.J., Behrensmeyer, A.K., 2012. Floral and environmental gradients on a Late Cretaceous landscape. Ecological Monographs 82, 23-47.

Zhang, Y.Q., Chen, J.M., Miller, J.R., 2005. Determining digital hemispherical photograph exposure for leaf area index estimation. Agricultural and Forest Meteorology 133, 166-

181.

Zhao, Z., Pearsall, D.M., 1998. Experiments for Improving Phytolith Extraction from Soils. Journal of Archaeological Science 25, 587-598.

Table 2.1. Phytolith yield and biomass contribution of tree species in Palo Verde, and their effect on the *FI-t* ratio.
Column 1 represent quadrats (15-25). Column 3 represent tree species phytolith yield. Phytolith yield is ranked according to four categories (0-4) where 0 = non-productive, 1= very low producer, 2= low producer, 3= moderate producer, 4 = high producer. Superscripts next to species names (column 3) indicate [1]deciduous species, [2]semideciduous species, and [3]evergreen species. Numbers in parenthesis next to species names (column 3) represent the species relative biomass with the quadrat. BA (%) values in column 4 indicate the percent biomass represented in each quadrat by species in each of the yield ranking categories. % deciduous biomass (column 5) indicate the percentage of the total tree biomass in each quadrat that corresponds to deciduous trees.
Arrows in column 6 and 7 illustrate the predicted effect of respectively the differential phytolith yield among species, and of the percent deciduous biomass on the *FI-t* ratio; arrows pointing up and down suggest respectively potential overestimation and underestimation of the *FI-t* ratio. Column 8 represent the *FI-t* ratio estimated by bootstrap analysis. Column 9 represent LAI values calculated from hemispheric photographs (column 7).

Quadrat	Yield Rank	Species	BA (%)	% deciduous biomass	Predicted yield and biomass effect on FI-t ratio	Predicted deciduousness effect on FI-t ratio	Estimated FI-t ratio	LAI
15	0	*Machaerium biovolatum*[1]	13.45	~77%	↓		19.16	3.05
	1	*Aphelandra scabra*[3] (2.12), *Astronium graveolens*[2] (31.25), *Erythroxylum havanense*[1] (2.48)	35.85					
	2	*Casearia corymbosa*[1] (0.07), *Tabebuia ochracea*[1] (16.27)	16.34					
	3	*Acacia centralis*[1] (33.62), *Lonchocarpus parviflorus*[1] (0.18)	33.80					
	4	*Alophyllus racemosus*[1]	0.57					
16	0	n.a.	0.00	~82%	↓		37.83	2.93
	1	*Erythroxylum havanense*[1] (9.38), *Manilkara chicle*[3] (1.17)	10.56					

	2	*Casearia corymbosa*[1] (0.2), *Luehea candida*[2] (0.95), *Tabebuia ochracea*[1] (18.08)	19.23				
	3	*Cordia callococca*[1] (0.37), *Guazuma ulmifolia*[1] (7.11), *Lonchocarpus parviflorus*[1] (26.25)	33.73				
	4	*Alophyllus racemosus*[1] (20.39), *Cordia dentata*[2] (0.07), Indetermined Malvaceae (0.95), *Simarouba glauca*[3] (15.07)	36.48				
17	0	*Coccoloba caracasana*[2] (2.5), *Thouinidium decandendrum*[2] (0.4)	2.90	~6%	↓↓↓	↓↓↓	41.46	3.5
	1	*Acacia collinsii*[1] (0.4), *Astronium graveolens*[2] (7), *Manilkara chicle*[3] (0.05), *Triplaris melanodendron*[3] (84.33)	91.78				
	2	*Psychotria carthagenesis*[3] (1.4)	1.39				
	3	n.a.	0.00				
	4	*Alophyllus racemosus*[1] (0.24), *Spondias mombin*[1] (3.7)	3.92				
19	0	n.a.	0.00	~84%	≈		84.45	3.5
	1	*Astronium graveolens*[2] (5.94)	5.94				
	2	*Calycophyllum candidissimum*[1] (15.9), *Capparidastrum frondosum*[3] (0.02)	15.88				
	3	*Acacia centralis*[1] (42.4) *Gauzuma ulmifolia*[1] (25.7), *Lonchocarpus parviflorus*[1] (0.4), *Rourea glabra*[3] (2.2)	70.64				
	4	*Garcia nutans*[3] (7.54)	7.54				

20	0	n.a.	0.00	~1%	≈	↓↓↓	59.13	3.98
	1	*Manilkara chicle*[3]	0.11					
	2	*Tabebouia ochracea*[1]	0.06					
	3	*Rourea glabra* (3.02)[3], *Tabebouia rosea*[1] (0.05)	3.07					
	4	*Garcia nutans*[3]	96.76					
21	0	n.a.	0.00	~98%	↓		62.15	3.78
	1	n.a.	0.00					
	2	*Cochlospermum candidissimum*[1] (73.45), *Tabebouia ochracea*[1] (0.16)	73.61					
	3	*Acacia centralis*[1] (5.76), *Guazuma ulmifolia*[1] 16.33) *Alophyllus racemosus*[1] (0.08), *Bonellia nervosa*[1] (0.09), *Croton niveus*[1] (1.72), *Lonchocarpus fasciculatus*[1] (0.45),	22.09					
	4	*Pithecellobium unguit-catii*[1] (0.06), *Simarouba glauca*[3] (2.44)	4.30					
22	0	n.a.	0.00	~84%	≈		62.71	2.28
	1	n.a.	0.00					
	2	*Caycophyllum candidissimum*[1] (42.47), *Casearia corymbosa*[1] (1.07), *Cochlospermum vitifolium*[1] (18.16)	61.71					
	3	*Acacia centralis*[1] (1.27), *Guazuma ulmifolia*[1] (1.82) *Bonellia nervosa*[1] (6.08),	3.10					
	4	*Croton niveus*[1] (5.57), *Pithecellobium unguis-catii*[1] (7.66), *Simarouba galuca*[3] (15.89)	35.20					

23	0	*Randia aculeata*[3] (0.03), *Randia armata*[2] (0.59)	0.62	~98%	≈↓		81.07	1.87
	1	*Acacia collinsii*[1] (0.59), *Erythroxyllum havanense*[1] (4.89), *Guettarda macrocarpa*[1] (36.26)	41.64					
	2	*Casearia corymbosa*[1] (35.52), *Guaiacum sanctum*[3] (1.03),	36.55					
	3	*Acacia centralis*[1] (0.13), *Trichilia hirta*[1] (2.19)	2.32					
	4	*Bonellia nervosa*[1] (0.59), *Croton niveus*[1] (3.38), *Lonchocarpus minimiflorus*[1] (14.85), *Pithecellobium unguis-catii*[1] (0.04)	18.86					
24	0	*Quadrella indica*[3]	0.39	~99%	≈		84.58	2.34
	1	*Acacia collinsii*[1] (28.84)	28.84					
	2	*Capparidastrum frondosum*[3] (0.04), *Guaiacum sanctum*[3] (0.24)	0.28					
	3	*Guazum ulmifolia*[1]	46.38					
	4	*Bonellia nervosa*[1] (0.04), *Pithecellobium unguis-catii*[1] (20.07	24.12					
25	0	*Quadrella indica*[3] (0.09), *Randia armata*[2] (0.34)	0.33	~15%	≈	↓↓	55.47	1.69
	1	n.a.	0.00					
	2	*Calycophyllum candidissimum*[1] (2.55), *Casearia corymbosa*[1] (0.92), *Cochlospermum vitifolium*[1] (7.67), *Guaicum sactum*[3] (0.3)	11.45					
	3	n.a.	0.00					
	4	*Bonellia nervosa*[2] (84.84), *Croton niveus*[1] (3.21),	88.23					

Lonchocarpus costaricensis[1] (0.18),	

Table 2.2. Phytolith counts from La Selva and Palo Verde quadrats.

Phytoliths are classified into eight functional groups (the seven defined groups, plus the group "Other", containing undetermined and/or non-diagnostic phytoliths) column 3-11). The *FI-t* ratio estimates in columns 11 are expressed as the percent of forest indicator (FI) phytoliths in a sum of FI and grass phytolith morphotypes (%FI/(FI+GSSCPs)). These estimates were obtained by bootstrap analysis of each sample, and are associated to a lower and an upper confidence interval, given in columns 12 and 13.

Site	Quadrat	Palms	Zingiberales	Palm/ Zingiberales	Palm/ Zingiberales/ Bromeliads	Other FI	Woody dicots	GSSCPs	Other	FI-t estimate	Lower 95% CI	Upper 95% CI
La Selva	Q2	135	4	0	0	111	50	1	37	99.67	98.98	100
	Q3	139	0	0	0	128	18	3	38	98.96	97.6	100
	Q1	99	0	0	0	116	13	13	19	94.61	91.56	97.15
	Q4	114	2	0	0	68	27	0	40	100	100	100
	Q6	39	0	0	0	103	46	3	144	98.43	96.41	100
	Q7	171	15	0	0	108	21	4	26	98.75	97.45	99.69
	Q8	288	0	0	0	103	29	7	79	98.36	96.99	99.52
	Q9	94	4	0	0	92	18	4	56	98.11	96.1	99.55
	Q10	139	0	0	0	67	25	1	28	99.57	98.68	100
	Q11	176	0	0	0	104	15	2	37	99.33	98.3	100
	Q12	96	1	0	0	125	16	6	50	97.54	95.42	99.2
	Q13	153	1	0	0	98	5	1	71	99.61	98.8	100
	Q14	141	7	0	0	91	18	5	34	98.09	96.24	99.62
	Q5	135	1	0	0	48	27	4	15	98.14	96.21	99.55
Palo Verde	Q15	1	0	0	0	59	4	270	15	19.16	14.88	23.21
	Q16	19	0	0	3	79	0	166	21	37.83	32.2	43.4
	Q17	6	0	0	0	77	2	120	46	41.46	34.72	48.31
	Q19	7	1	9	8	203	11	44	37	84.45	80.5	88.57
	Q20	2	0	5	2	138	2	103	35	59.13	52.46	65.32
	Q21	5	4	17	2	121	7	95	64	62.15	55.95	68.44
	Q22	28	1	7	27	83	3	88	34	62.71	56.96	68.7
	Q23	13	7	13	1	191	2	53	35	81.07	76.84	85.77
	Q24	3	2	0	0	157	8	31	26	84.58	79.5	89.34
	Q25	4	0	0	5	126	2	110	33	55.47	48.96	61.67

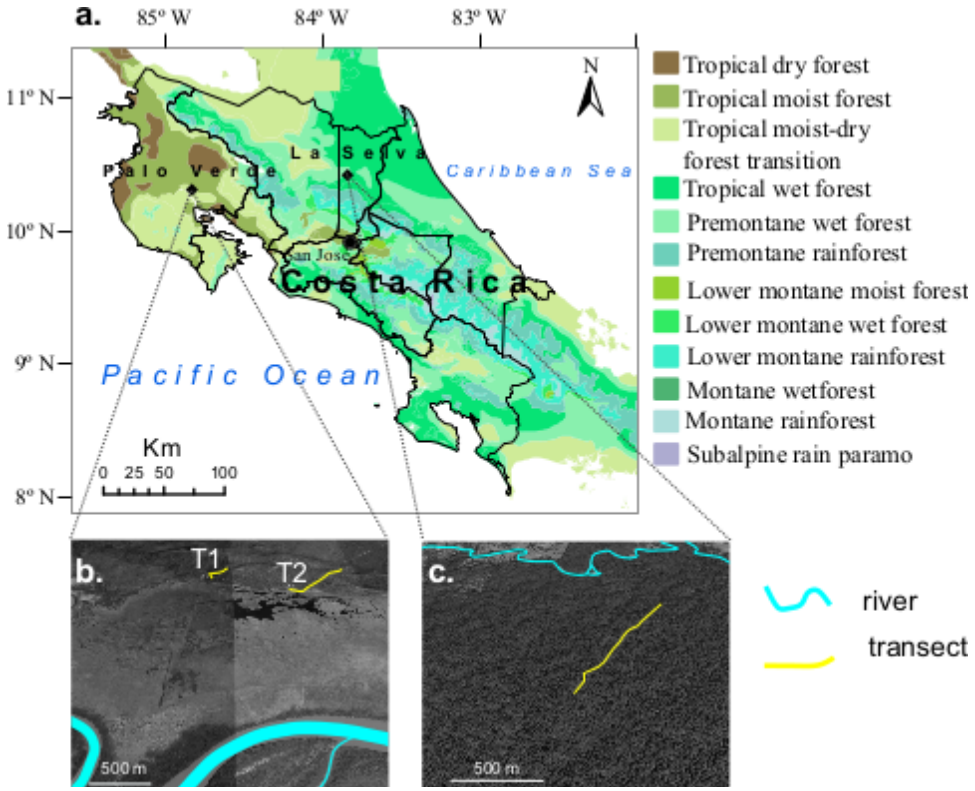


Figure 2.1. Map of the study sites. **a.** Map of Costa Rica showing the location of the two sites, and the location of transects within Palo Verde National Park; and **b**. La Selva Biological station; **c.** Transects are shown in yellow, and river bodies are shown in light blue.

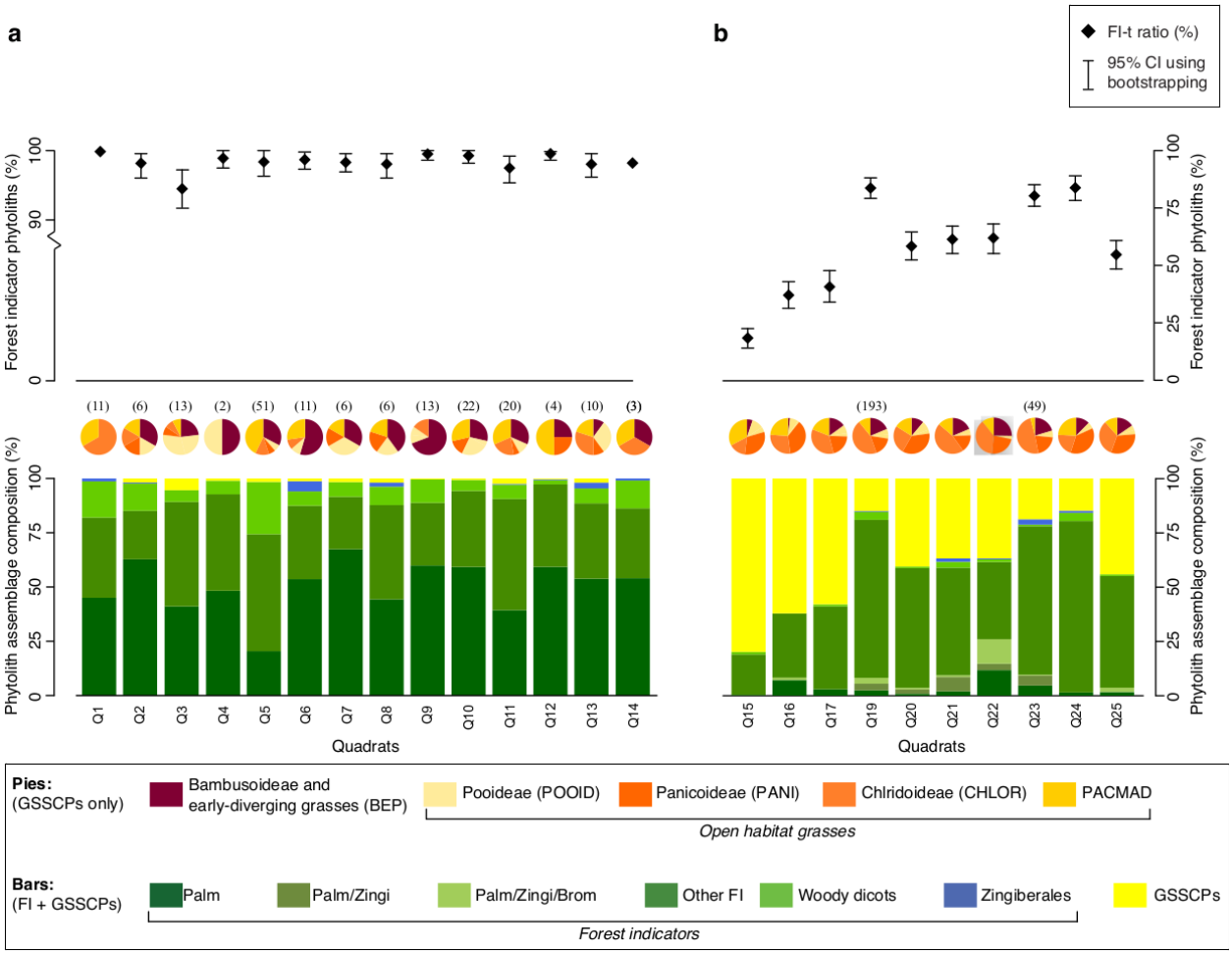


Figure 2.2. Phytolith assemblages from la Selva and Palo Verde.

a. La Selva; **b.** Palo Verde. The upper graphs show the *FI-t* ratio expressed as the percent forest indicator (FI) phytoliths in a sum of FI and grass phytolith morphotypes (%FI/(FI+GSSCPs)). Diamonds represents estimates of *FI-t* ratios from phytolith counts; vertical bars represent 95% confidence intervals based on bootstrap analysis of each sample. The bar graphs (lower part of the figure) represent phytolith assemblage composition of the samples indicating percent phytoliths in each of the 8 defined functional groups (in legend). Pie charts represent relative proportion of different GSSCP phytolith morphotypes in the phytolith assemblages (yellow portion of the bars in the bar graphs) in each of the 5 defined GSSCP functional groups (in

legend). Percent forest indicators in the upper graphs refer to the sum of all forest indicators (green and blue gradient) in the legend.

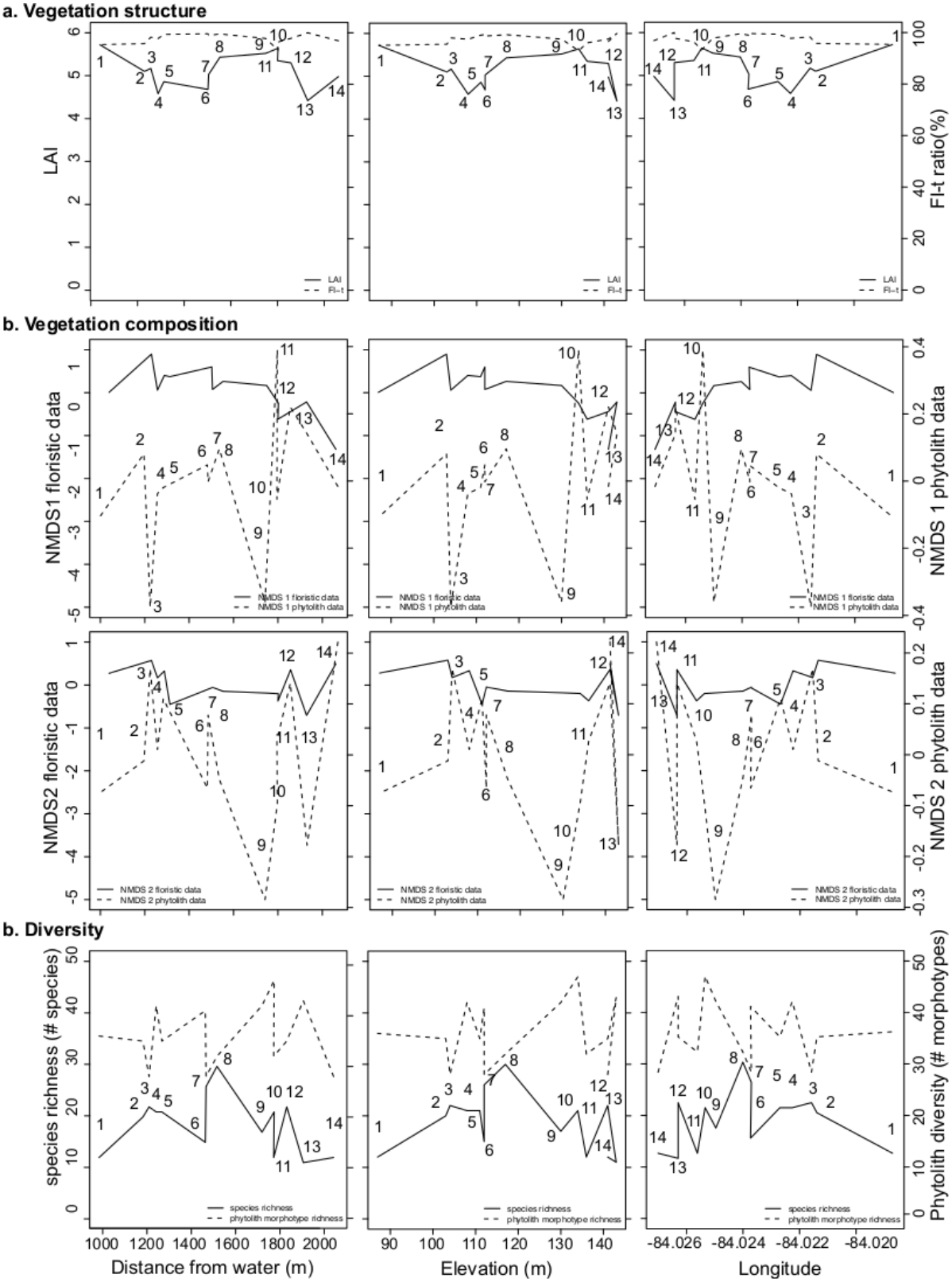


Figure 2.3. Gradients in vegetation structure, composition and diversity along La Selva transect.

a. Vegetation structure as measured by LAI and the *FI-t* ratio; **b.** vegetation composition as measured by quadrats loadings on axis 1 and 2 of the NMDS ordination according to species composition, and quadrats loadings on axis 1 and 2 of the NMDS ordination according to phytolith morphotype composition ; and **c.** diversity as measured by species α diversity, and number of phytolith morphotypes. Gradients are plotted along elevation, distance from the water, and longitude.

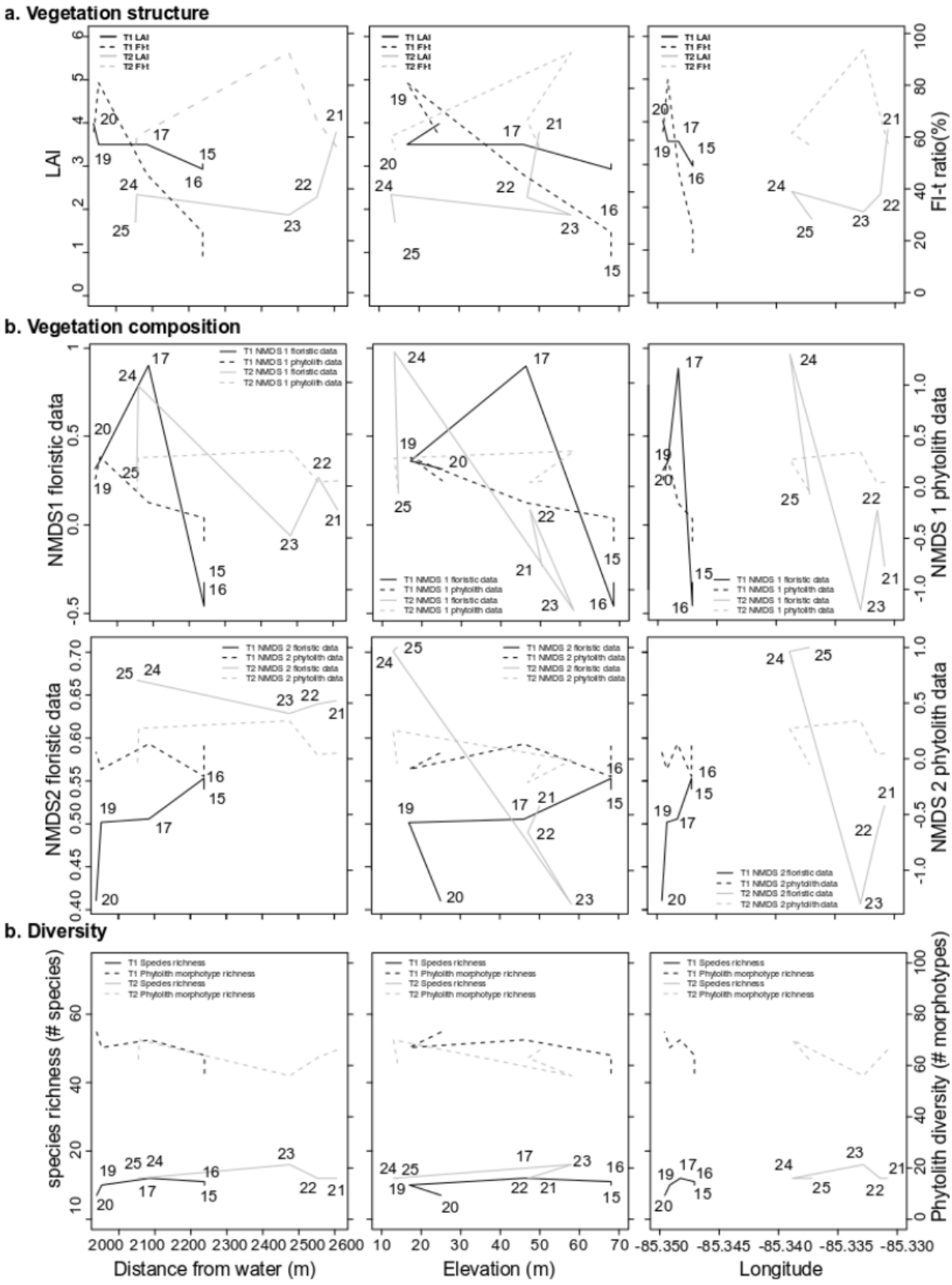


Figure 2.4. Gradients in vegetation structure, composition, and diversity along Palo Verde transects.

a. Vegetation structure as measured by LAI and the *FI-t* ratio; **b.** vegetation composition as measured by quadrats loadings on axis 1 and 2 of the NMDS ordination according to species composition, and quadrats loadings on axis 1 and 2 of the NMDS ordination according to phytolith morphotype composition ; and **c.** diversity as measured by species α diversity, and number of phytolith morphotypes. Gradients are plotted along elevation, distance from the water, and longitude.

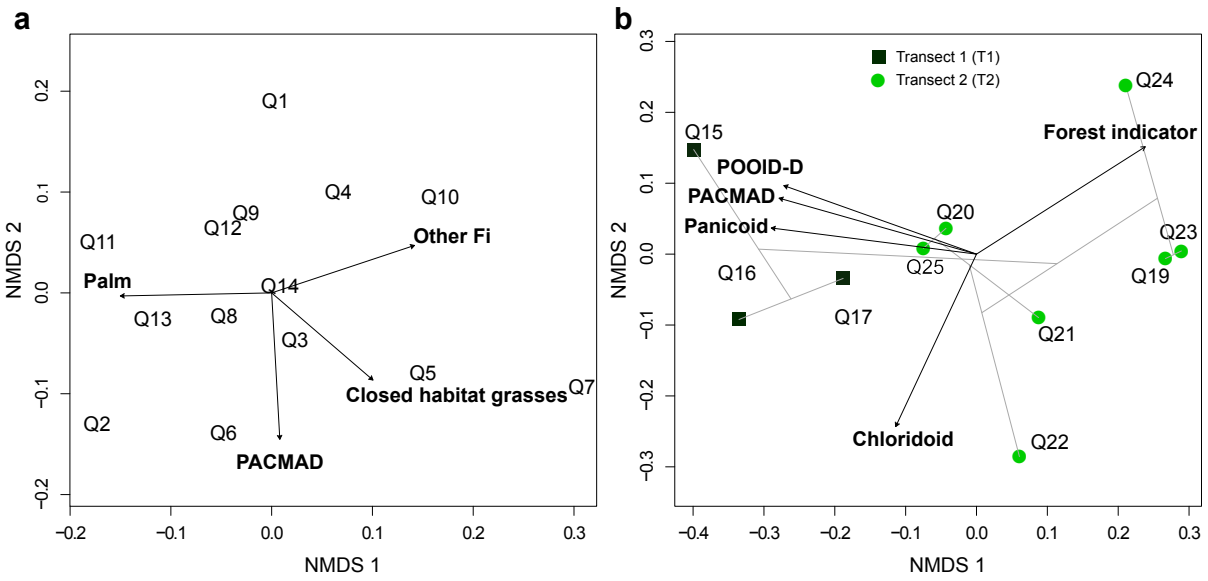


Figure 2.5. Ordination diagrams of phytolith assemblages within of La Selva and Palo Verde. **a.** NMDS ordinations of the soil samples across La Selva quadrats, and phytolith classes (black vectors). **b.** NMDS ordinations of the soil samples across Palo Verde quadrats, and phytolith classes (black vectors), with superimposed dendrogram (grey lines) obtained by cluster analysis. Only classes (i.e., phytolith compound variables) with significant ($p<0.01$) loadings on the ordination axes are shown here. The direction of the vectors indicates maximum correlation between the score of the samples in the quadrats (soil phytolith assemblages) and the phytolith classes (compound variables.

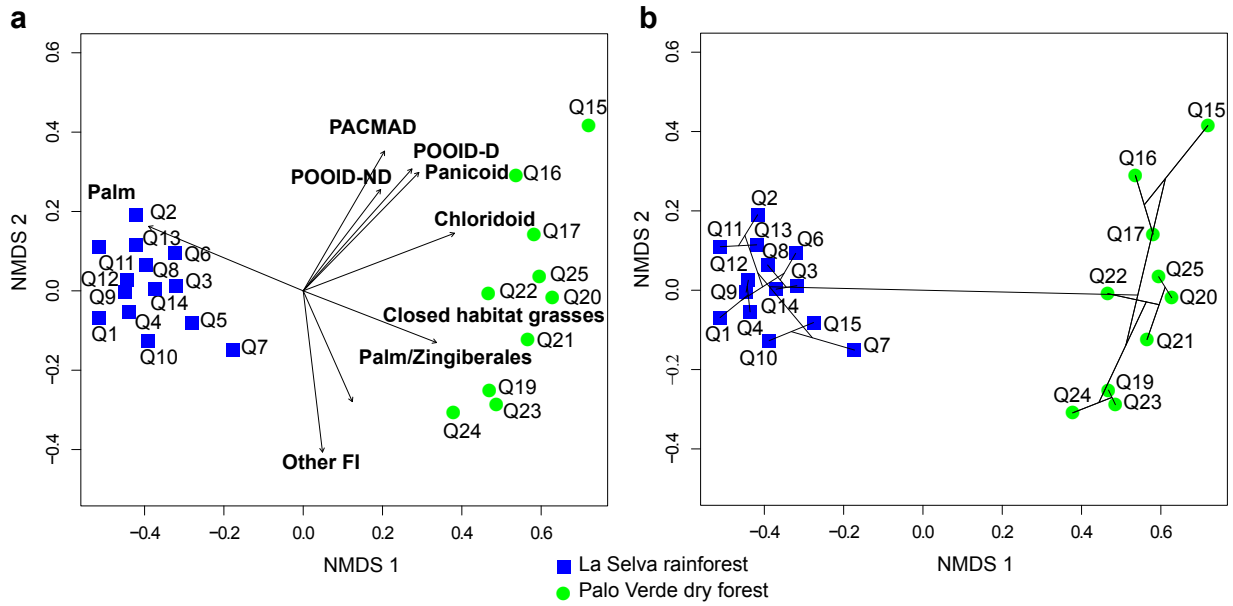


Figure 2.6. Ordination diagrams of phytolith assemblages across La Selva and Palo Verde quadrats.

a. NMDS ordination (stress value= 0.047) of the soil samples across La Selva (Q1-Q14) and Palo Verde (Q15-Q25) quadrats, and phytolith classes (black vectors). Only phytolith classes (i.e., compound variables) with significant ($p<0.01$) loadings on the ordination axes are shown here. The direction of the vectors indicates maximum correlation between the score of the samples in the quadrats (soil phytolith assemblages) and the phytolith classes (compound variables). **b.** NMDS ordination (as in **a**), overlaid by the dendogram of the quadrats at the two sites obtained from cluster analysis based on phytolith functional groups. Blue squares represent La Selva quadrats; green circles represent Palo Verde quadrats.

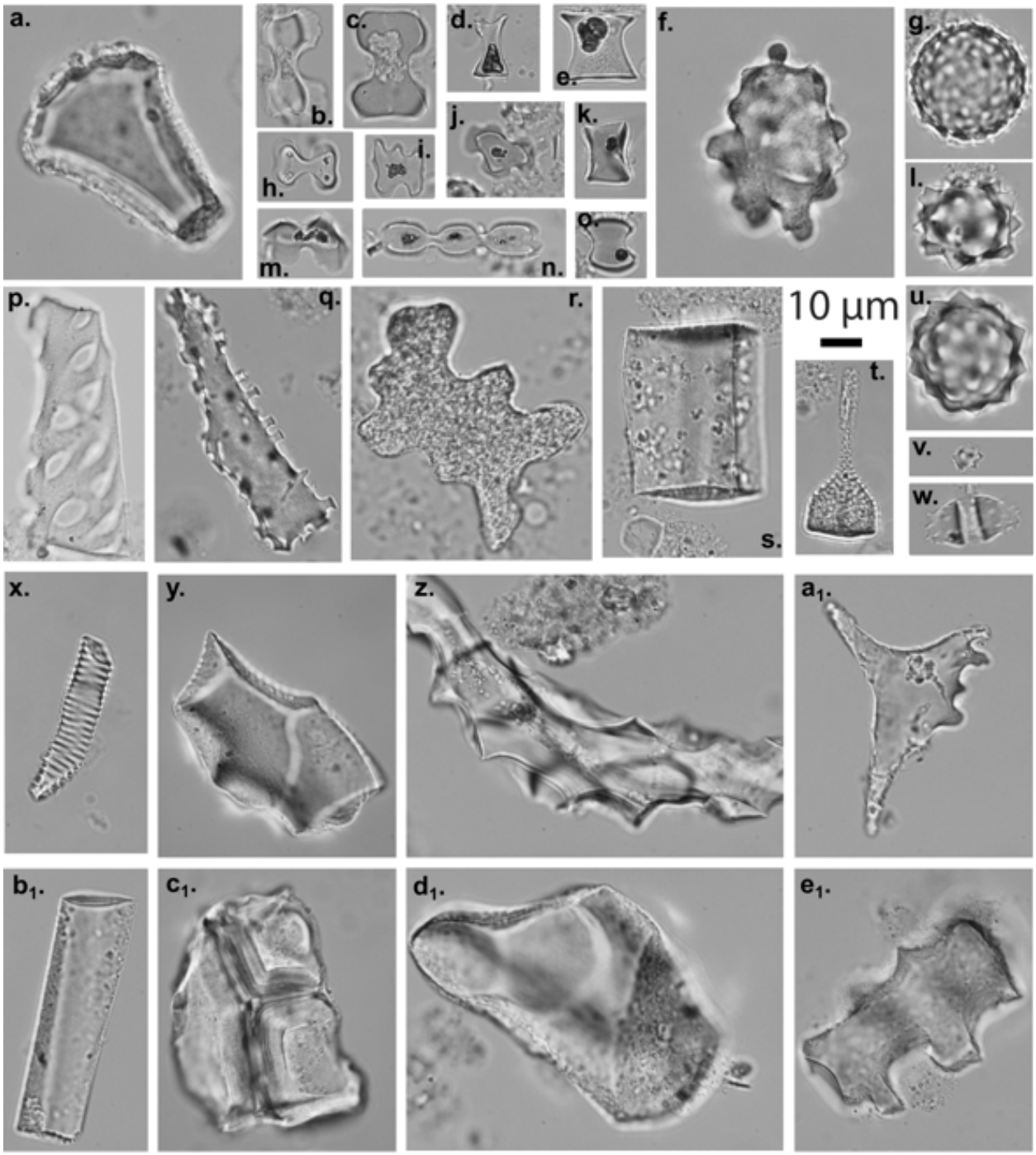


Figure 2.7. Examples of phytoliths from la Selva.

Phytolith morphotypes diagnostic of Poaceae (a-e; h-k; m-o), palms (g, l, v, w), Zingiberales (f, x), grass/monocots/conifers (s, b_1), other forest indicators (r, x, c_1), and woody dicots (p, y, z, a_1, d_1, e_1): **a)** Blo-10, bulliform B: keystone-shaped plate (GRASS); **b)** BI-6, symmetry C bilobate (inverted) GSSCP (PACMAD); **c)** BI-8, symmetry E bilobate (Panicoid type) GSSCP (Panicoideae); **d)** COF-1A, tall rondel with spiked top GSSCP (Bambusoideae); **e)** CO-1, generic

(truncated) conical rondel GSSCP (Pooideae); **f)** "subglobular papillate" from Marantaceae (Chen and Smith, 2013) (Zingiberales); **g)** Clm-2, echinate sphere (Palms); **h)** BI-5, symmetry D bilobate GSSCP (PACMAD); **i)** PO-4, polylobate with top and bottom of same size/shape GSSCP (symmetry E), (Panicoideae); **j)** CR4-4, four-lobed cross with rectangular or polyhedral top (Panicoideae); **k)** SA-3, collapsed saddle GSSCP (Bambusoideae); **l)** Clm-2, echinate sphere (Palms); **m)** BI-1, regular *Stipa*-type bilobate (symmetry A) GSSCP (Pooideae); **n)** PO-3, polylobate with larger top (symmetry C) GSSCP (PACMAD)
o) SA-1, true saddle GSSCP (Chloridoideae); **p)** unclassified vessel element (Woody dicots?); **q)** unclassified tracheary element (Woody dicots?); **r)** Epi-2, anticlinal epidermal (Other FI); **s)** Blo-1, rectangular (GRASS); **t)** unclassified papillate trichome (?) (Other FI?); **u-v)** Clm-2, echinate sphere (Palms); **w)** "T1 papillate trough" from Heliconia (Chen and Smith, 2013) (Zingiberales); **x)** Tra-1, helical tracheary element (Other FI); **y)** Scl-3, multifaceted S-body (Woody dicots); **z)** Undetermined multidimensionally faceted S-body (cf. Postek, 1981; Piperno, 1988; Runge, 1999) (Woody dicots); **a_1)** Scl-6, slim, tapering S-body; subtype 1 (Woody dicots?); **b_1)** Elo-3, faceted elongate (GRASS); **c_1)** VI-2, non-spherical VI (laminated) body (Other FI); **d_1-e_1)** unclassified sclereid bodies (Woody dicots?). Modified from Crifò and Strömberg, in press.

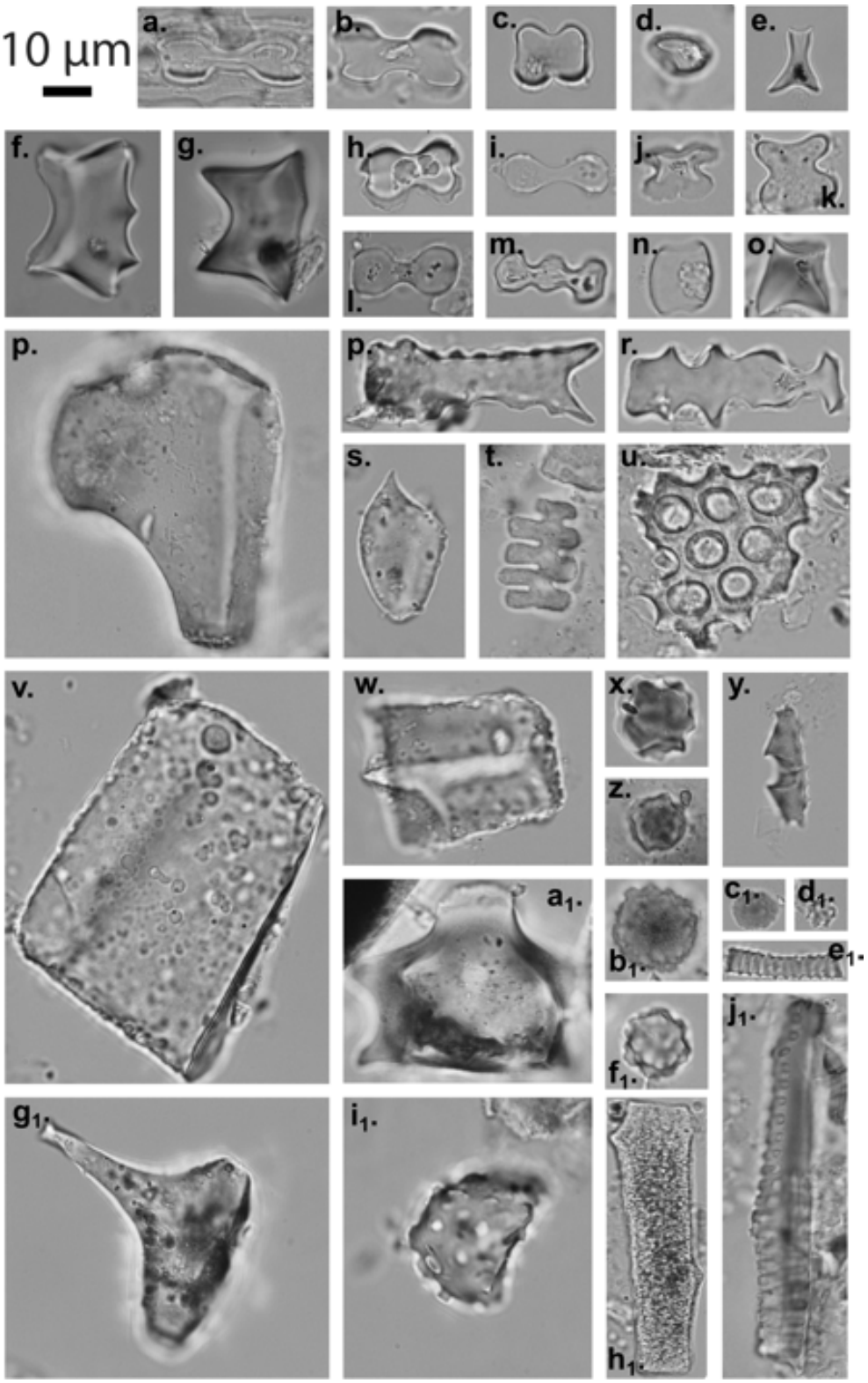


Figure 2.8. Examples of phytoliths from Palo Verde.

Phytolith morphotypes diagnostic of Poaceae (a-t), grass/monocots/conifers (u-w), Zingiberales (x-b_1), palms (f_1), palm/ Zingiberales (c_1, d_1), other forest indicators (e_1, g_1, I, j_1), as well as non-

diagnostic or produced by unknown plant taxa (h_1): **a)** BI-5, symmetry D bilobate GSSCP (PACMAD); **b)**, **l)** BI-8, symmetry E bilobate (Panicoid type) GSSCP (Panicoideae); **c)** CR4-2, four-lobed cross with near cross-shaped top GSSCP (Panicoideae); **d)** CO-1B, tall generic (truncated) conical rondel (top view) GSSCP (Pooideae); **e)** BI-13, shovel bilobate with rectangular top ("Chusqoid body") GSSCP (Bambusoideae); **f)** SA-3, collapsed saddle GSSCP (Bambusoideae/Oryzoideae/Early diverging grasses); **g)** SA-3, collapsed saddle GSSCP (Bambusoideae/Oryzoideae/Early diverging grasses); **h)** BI-7, symmetry D bilobate GSSCP (PACMAD clade); **i)** BI-5, symmetry B bilobate GSSCP (PACMAD clade); **j)** CR4–2, four lobed cross with near cross-shaped to cross-shaped top GSSCP (Panicoideae); **k)** CR3-1, "perfect" three-lobed cross GSSCP (Panicoideae); **m)** PO-3, symmetry C polylobate GSSCP (PACMAD clade); **n)** SA-1, true saddle GSSCP (Chloridoideae); **o)** SA-3, collapsed saddle GSSCP (Bambusoideae); **p)** Blo-10, bulliform B; keystone-shaped plate (GRASS); **q)** Epi-11, elongated epidermal with indented end (GRASS); **r)** Epi-12, papillate epidermal with large/elongate papilla (GRASS); **s)** Tri-8, spindle or teardrop-shaped trichome (GRASS); **t)** M-7, vertebral column mesophyll body (GRASS); **u)** M-6, crystal- like/well silicified straight file parenchyma (GRASS); **v)** Blo-2, thickened rectangular plate (Other); **w)** Blo-3 type A (grass type), faceted rectangular plate (GRASS); **x)**, **z)**, **b_1)** "D1- and D2-type druses" from Zingiberales (Chen and Smith, 2013) (Zingiberales); **y)** "T1 papillate trough" from Heliconiaceae (Chen and Smith, 2013) (Zingiberales); **a_1)** "Kn2 knobby" body from Maranta seed (Chen and Smith, 2013) (Zingiberales); **c_1)** Clm-1, *Nypa*-type body from palms or alternatively "rugose hat- shaped" from Marantaceae (Chen and Smith, 2013) (Palm/Zingiberales); **d_1)** Clm-2, echinate sphere from palms or alternatively, "globular microechinate" from Zingiberaceae (Chen and Smith, 2013) (Palm/Zingiberales); **e_1)**, **j_1)** Tra-1, helical tracheary element (Other FI); **f_1)** Clm-2 echinate sphere (Palm); **g_1)** Scl-4, spongy mesophyll body (Other FI); **i_1)** Scl-8, compact irregular S-body type A (Woody dicots); **h_1)** Elo-1, smooth elongate (Other). Modified from Crifò and Strömberg, in press.

Chapter 3. Vegetation shift during the Middle Miocene Climatic Optimum of Southern Patagonia (Argentina) recorded in phytolith assemblages

ABSTRACT

High global temperature and pCO_2 characterized the Middle Miocene Climatic Optimum (MMCO; ca. 17–14.5 Ma) and are thought to have promoted highly diverse ecosystems in warm and wet climates at high latitudes. However, only sparse observations inform interpretations of regional middle Miocene climate and biotas outside the Northern Hemisphere. The Santa Cruz Formation (SCF) of coastal Patagonia represents an exception, as the southernmost sedimentary sequence in the world recording the onset of the MMCO. SCF fossiliferous horizons yield one of the most species-rich and well-preserved vertebrate assemblages on Earth. While stratigraphic, paleontological, and geochemical studies have allowed refinement of SCF stratigraphy, and reconstruction of local climate, paleoenvironments, and ecology of the fauna, a detailed study of the SCF vegetation is lacking. Among the several lines of paleobotanical evidence that are present in the SCF, phytoliths hold the most promise for reconstructing Miocene plant community change; however, the SCF phytolith record have yet to be the object of quantitative, and detailed study through time.

Here we present results from analysis of phytolith assemblage composition (% plant functional types) documenting vegetation change during the onset of the MMCO, and interpret them in light of our current climate proxy data from the SCF. Our results show that right before the onset of the MMCO southeastern Patagonia was characterized by the presence of mixed habitats with abundant C_3 grasses, a woody component composed of conifers and dicots, as well as palms in varying abundance. We interpret this habitat as a mix of woodland or open

woodland/shrubland, including palm shrubland. In the upper part of the SCF sequence, during the initial stage of the MMCO inferred from isotopic data to be drier, our data reveal a decrease in grass abundance, and a concomitant increase in the woody component of the vegetation. In addition, grass communities were dominated C_3 pooid grassed, with low abundance of the tropical PACMAD clade (which includes both C_3 and C_4 grasses). We interpret these trends as reflecting the expansion of dry-adapted woody vegetation in response to MMCO climate change, and to the detriment of a C_3 grass community which was not adapted to dry conditions. Further, we suggest that PACMAD grasses at the SCF were likely primarily C_3, and that C_4 grass evolution, and the expansion of grass-dominated open-habitats did not take place in Patagonia until after the early middle Miocene.

3.1 INTRODUCTION

The mid-Miocene Climatic Optimum (MMCO, 17-14.5 Ma) is one of Earth's most recent prolonged greenhouse transition events (Zachos et al., 2001, 2008). The MMCO preceded rapid global cooling that began in the Middle Miocene (~16 Ma), and continued during the rest of the Neogene, culminating with the establishment of the Northern Hemisphere Ice Sheet in the late Pliocene, and the last Ice Age (Flower and Kennet, 1994; Zachos et al., 2001 Kürschner et al., 2008). The MMCO was characterized by increased surface and bottom ocean temperatures, and the retreat of the Antarctic Ice Sheet, as well as possible increase in atmospheric pCO_2 (Flower and Kennet, 1994; Kürschner et al., 2008; Shevenell et al., 2008; Tripati et al., 2009; Beerling and Royer, 2011; Foster et al, 2012; Holbourn et al, 2015). These changes had profound impacts on Earth's biota, promoting the expansion of tropical and subtropical terrestrial ecosystems into higher latitudes (e.g., Gayó et al., 2005; Palazzesi and Barreda, 2012; Pound et al., 2012). In

North America and elsewhere, browsing ungulate taxa reached abnormally high diversity levels relative to today during this period (e.g., Raza et al., 1984; Janis et al., 2000, 2004). Further, it has been hypothesized that climatic changes coupled with tectonic events in western North America promoted shifts in plant diversity, and favored mixed habitats characterized by high primary productivity resulting in increased mammal diversity (e.g. Kohn and Fremd, 2008; Finarelli and Badgley, 2010; Badgley et al., 2017).

Yet, despite a body of work concerning MMCO environmental and biotic transformation on a global scale, our knowledge of local ecosystem change during this time is limited (but see e.g., Harris et al., 2017; Smiley et al., 2018). In addition, the MMCO paleontological and climate proxy record is overall strongly biased toward the low latitudes of the northern hemisphere (You et al., 2009), whereas southern hemisphere sites are rare and even less commonly allow study of basin-level change (but see Warny et al., 2009). The Santa Cruz Formation (SCF, 47-52°S) in southern Patagonia (Argentina) which spans the Early-Middle Miocene, constitutes an exception.

The SCF, which has been studied by paleontologists for over 100 years (Vizcaíno et al., 2013), hosts one of the richest Cenozoic faunal localities of South America, representing the homonymous Santacrucian South American Land Mammal Age (SALMA) (Simpson, 1940, and references thereafter). Recent stratigraphic, paleontological, and geochemical studies have allowed refinement of SCF stratigraphy (e.g., Marshall et al., 1986; Tauber, 1994, 1997; Fleagle et al., 1995; Vizcaíno et al., 2012; Cuitiño et al., 2016, 2019; Trayler et al., 2019), and reconstruction of local climate (e.g., Vizcaíno et al., 2010, 2012; Kay et al., 2012a; Trayler et al., in review), paleoenvironments, and ecology of the fauna (e.g., Bown and Fleagle, 1993; Tauber, 1999; Tejedor et al., 2006; Townsend and Croft, 2008; Cassini and Vizcaíno, 2012; Kay et al., 2008, 2012b; Vizcaíno et al., 2006, 2010, 2012; Croft, 2013; Montalvo et al., 2019).

However, while the SCF fossil fauna has yielded abundant and well preserved remains, paleobotanical evidence is scarce. Palynofloras are virtually absent and plant macrofossil evidence (leaf impressions and wood) is scant or exhibits low levels of preservation, allowing only qualitative reconstructions of the SCF paleovegetation (Brea et al., 2012). The presence of well preserved phytoliths has been reported in several sections in the SCF (Brea et al., 2012; Raigemborn et al., 2018a, b) but a quantitative, detailed analysis of phytolith assemblages through time (and space) has yet to be performed (although see Raigemborn et al., 2018b). Thus, direct evidence for patterns of vegetation change in response to MMCO global warming is currently lacking.

In this study we present data from phytolith fossil assemblages recovered from strata of several coastal localities of the Santa Cruz Formation spanning about 1 million years at the onset of the MMCO (17.5 - 16.5 Ma; Trayler et al., 2019) to address the following questions:

1) How did vegetation structure change in response to the initial warming pulse of the MMCO? We assess changes in vegetation structure (canopy openness) by studying phytolith assemblages quantitatively through a well-dated stratigraphic sequence (Fig. 3.1), focusing on the abundance and types of forest indicator phytoliths).

2) How did grass community composition change in response to warmer and drier conditions (Trayler et al., 2019) at the onset of the MMCO? We assess changes in grass community composition (forest versus open habitat grasses) by calculating the proportion of grass phytoliths indicative of forest habitats versus open habitat grasses along a well-dated stratigraphic sequence. We then compare vegetation evidence from phytolith analysis with existing data on SCF climate, providing new insights on the relationship between climate change, and vegetation, during the MMCO of Patagonia.

3) Was the remarkably high diversity of the Santa Cruz fauna supported by habitats characterized by vegetation heterogeneity (i.e., a mix of forested and open vegetation areas) throughout the onset of the MMCO as would be predicted based on modern ecology (Ritchie & Olff, 1999; Badgley & Fox, 2000) and SCF faunal data (Tejedor et al., 2006; Bargo et al., 2012; Kay et al., 2012a, b)? We compare the composition of coeval phytolith assemblages, within two main stratigraphic levels, and across several localities of the SCF to assess whether the vegetation was homogenous across space. We further compare vegetation evidence from phytolith analysis with existing data on the dietary ecology of the SCF ungulate fauna (Trayler et al., in review).

3.1.1 *Cenozoic climate and vegetation history of Southern South America*

Existing paleobotanical and isotopic data from Southern South America paints a complex and sometimes contradictory picture of climate and plant community change during the Cenozoic. Mega- and palynofloras indicate that Paleocene tropical vegetation persisted until the Eocene-Oligocene, when cool temperate forests dominated by conifers and deciduous angiosperms are thought to have expanded in response to the Eocene-Oligocene global cooling (Barreda and Palazzesi, 2007; Iglesias et al., 2011); while a transitional zone (with both megathermal and sub-Antarctic taxa) formed towards the Atlantic coast (Iglesias et al., 2011). In contrast, phytolith and isotopic data suggest that tropical forest vegetation gave way to open, palm-dominated shrublands by the mid-late Eocene in stable, warm but drying climates, with no major change across the Eocene-Oligocene boundary (Kohn et al., 2004, 2015; Strömberg et al., 2013; Dunn et al., 2015; Selkin et al., 2015). Both palynological and phytolith records indicate that the Oligocene-early Miocene was characterized by an increase in xerophytic-adapted taxa,

with grasses continuing to be rare (Barreda and Palazzesi, 2007, 2014; Iglesias et al., 2011; Strömberg et al., 2013), followed by a re-greening of Patagonia during the MMCO (14.5-17.0 Ma), with warmer temperatures and marine ingressions causing southward migration of tropical taxa (Barreda and Palazzesi, 2007; Iglesias et al., 2011; Palazzesi and Barreda, 2012; Dunn et al., 2015). Grass-dominated habitats appear to have developed in southern South America only in the late Miocene based on recent interpretation of microfloral records (Palazzesi and Barreda, 2012; Palazzesi et al., 2014; Strömberg et al., 2013, 2014; but see Zucol et al., 2010, 2018).

3.1.2 *SCF climate – current knowledge*

The application of different climate proxies to the SCF has similarly resulted in several uncertain or conflicting interpretations. Wood physiognomy from the fossil assemblages of the lower SCF suggests mean annual temperatures (MAT) of 9.3 ± 1.7°C or 19.3 ± 1.7°C, mean annual precipitations (MAP) of 869 ± 940 mm/yr, and moderate seasonality (~7 months dry season). Calculations of primary and secondary productivity (based on estimates of body biomass, population size, metabolic rate, and on-crop biomass) indicate MAP < 1000 mm/yr (Vizcaíno et al., 2010). Finally, climate reconstructions using MAP and MAT constraints from both paleofloras and paleofaunas from the lower SCF indicate MAP > 1000 mm/yr and MAT > 14˚C (Kay et al., 2012a). Recent work on carbon and oxygen isotope data from tooth enamel from two combined stratigraphic sections (spanning ~ 17.4 to ~ 16.4 Ma) has sought to resolve these inconsistencies and provide a more robust and precise picture of climate change through the SCF. The reported $\partial^{13}C$ values point to a decrease in MAP from ~ 700 ± 130 mm/yr at 17.4 Ma to ~ 340 ± 90 mm/yr at ~16.9 Ma, followed by an increase to ~500 mm/yr around 16.5 Ma, thus indicating substantially drier climates for the MMCO of Patagonia overall. New $\partial^{18}O$ data

from tooth enamel also suggest higher temperatures than previously shown, with MAT ~20˚C before the onset of the MMCO, increasing to ~25˚C by 16.4 Ma (Trayler et al., in review).

3.1.3 *SCF paleoenvironments – current knowledge*

SCF paleofaunas suggest a complex mosaic of habitats, including dense forest and open vegetation (shrubland/grassland) (Kay et al., 2012a). This environment was inhabited by species adapted to forest and woodlands, such as anteaters, glyptodonts, terrestrial sloths (Bargo et al., 2012), and primates (Tejedor et al., 2006; Kay et al., 2012b), as well as dry adapted taxa such as armadillos (Vizcaíno et al., 2006). Many species of notoungulates are also present. Analysis of their tooth-enamel $\partial^{13}C$ and $\partial^{18}O$ values indicates a C_3 plant-dominated browsing diet requiring a woody component, as well as low seasonality in diet, and mixed habitats with moderate precipitation (Trayler et al., in review). In the lower portion of the SCF, these data are mostly in agreement with paleobotanical evidence. In fact, preliminary, qualitative analysis of plant macrofossils (wood and leaf compressions), and phytolith assemblages from the lower strata of the SCF had suggested the presence of mixed/mosaic vegetation (trees and grasses) where angiosperm taxa with megathermal affinity coexisted with semi-arid elements in a seasonal, dry temperate to temperate-warm climate (Brea et al., 2012). A more recent study of paleosols and associated ichnofossils, phytoliths, and vertebrates from four localities of the SCF suggests a shift from mixed vegetation to plant communities dominated by palms and grasses, and, towards the end of the sequence, to habitats dominated by microthermic (C_3) grasses in a context of a relatively stable, warm-humid climate (Raigemborn et al., 2018b). These results confirm the presence of mixed habitats at the base of the sequence but are in disagreement with the notion of a MMCO "regreening" of Patagonia suggested by some previously published pollen and

phytolith data (Iglesias et al., 2011; Palazzesi and Barreda, 2012; Dunn et al., 2015). They also conflict with the longstanding hypothesis of a shift from humid to dry conditions (Tauber, 1994), recently supported by the study of climate proxies (Trayler et al., in review).

Raigemborn and colleagues (2018b) did not propose an explanation for these discrepancies. However, it is possible that the observed vegetation trend is the result of variation in regional environmental conditions, and that these were overall more humid locally compared to other coeval sites elsewhere in Patagonia. Alternatively, the analytical approach used in that study might have prevented detailed assessment of the temporal trend of vegetation change. For instance, paleosol samples for phytolith analyses from several facies were lumped into three main groups representing three different chronological phases of the evolution of the SCF paleoenvironments (Raigemborn et al., 2018b). This mixing of phytolith assemblages of different ages could result in an obscured temporal vegetation signal. Hence, further, fine-scale analysis of phytolith assemblages along the SCF is needed to better constrain vegetation change during the onset of MMCO and link it to data from fauna assemblages and climate proxies.

3.2 GEOLOGIC SETTINGS – THE SANTA CRUZ FORMATION (SCF)

The SFC is part of the infill of the South American Austral Basin. This basin formed on the foreland side of the Southern Patagonian Andes starting with a rift stage in the Early Cretaceous and continuing with a foreland stage in the Late Cretaceous and Cenozoic, characterized by subsidence resulting from the development of a fold-and-thrust belt to the west (Cuitiño et al., 2016). During the Miocene, a major pulse in Andean uplift created the accommodation space that allowed the accumulation of marine and terrestrial deposits (Bown and Feagle, 1993; Blisniuk et al., 2005; Fosdick et al, 2013; Cuitiño et al., 2016). Marine

deposits, known as the “Patagonienense” transgression, resulted from a marine transgression that occurred between the late Oligocene and the early Miocene (Cuitiño and Scasso, 2010; Cuitiño et al., 2016). Terrestrial Miocene deposits overlaying the “Patagoniense” transgression include the Santa Cruz Formation in addition to the Monte León Formation, the Estancia 25 de Mayo Formation, and the El Chacay Formation. The SCF outcrops trend NW-SE, with a slight southward dip (Fleagle et al., 2012). They extend over a large area of the Santa Cruz Province of Argentina encompassing localities from the Andean foothills to the west, to the Atlantic coast of Patagonia to the east (Blisniuk et al., 2005; Cuitiño and Scasso, 2010; Fleagle et al., 2012; Marshall, 1976; Tauber, 1994, 1997). In this study we focus on eight of the coastal localities (marked by black stars in Fig 3.1).

Along the Atlantic coast, the stratigraphic features of the SCF have been documented by several authors from localities north of the Río Coyle (e.g., Bown and Fleagle, 1993) and south of the Río Coyle and North of the Río Gallegos (e.g., Tauber, 1997). North of the Río Coyle river, Bown and Fleagle (1993) described the SCF as a sequence of superimposed volcanoclastic mudstones with poorly developed paleosols that accumulated and formed, respectively, on a coastal alluvial plain. Relatively unaltered tuff beds were also described (Bown and Fleagle, 1993). Based on the study of the southernmost coastal localities, Tauber (1994) divided the unit into two members. The fossil-rich, lower member (Estancia La Costa Member; Tauber, 1997) is dominated by fine-grained beds with abundant pyroclasts, and it has later been interpreted as a low-energy floodplain with poorly developed paleosols and pyroclastic input (Matheos and Raigemborn, 2012; Raigemborn et al., 2015). The upper member (Estancia La Angelina Member; Tauber, 1997) is dominated by claystone, mudstone, and sandstone beds, and has yielded fewer fossils than the lower member.

Correlations among coastal localities and between coastal and inland SCF localities have recently been refined by Cuitiño et al. (2016) and Trayler et al. (2019), (Fig. 3.2). Trayler et al., (2019) also developed an age model for the SCF coastal localities using new, high-precision, isotope dilution U-Pb ages (ranging from 16.78 ± 0.03 Ma to 17.62 ± 0.03 Ma) from zircons of seven tuff layers across five localities (Cerro Observatorio, CO; Cañadón de la Vacas, CV; Rincón del Buque 3, RB3; Killik Aike Norte, KAN; and Cabo Buen Tiempo, CBT), as well as Ar^{40}/Ar^{49} ages previously reported by Perkins et al., (2012) from one locality (Cañadón de las Vacas).

3.3 METHODOLOGY

3.3.1 *Sampling*

The localities included in this study are from north to south: Cañadón de las Vacas (CV), Cerro Observatorio (CO), Rincón de Buque 2 (RB2), Rincón de Buque 3 (RB3), Cañadón del Indio (CI), Puesto Estancia La Costa (PLC), Monte Tigre (MT), Cabo Buen Tiempo (CBT), and Killik Aike Norte (KAN) (synonomous with Killik-Aike Norte in Ragiemborn et al., 2018b; Fig. 3.1). Paleosols, tuffs, and other fine-grained layers were sampled for phytolith analysis approximately every two meters throughout a composite section obtained from CV (the most complete section), RB3, CI, PLC, MT, and KAN. Two tuff layers, namely the tuff CO, and the tuff Toba Blanca (highlighted in pink and green respectively in Fig. 3.2), were also sampled laterally within and across several localities. This sampling resulted in a total of 292 samples, of which 165 are from logged sections and 127 from lateral transects. Correlations between localities where based on Trayler et al., (2019); hence, we follow the same terminology and

abbreviations for sites and tuff levels. Along the six stratigraphic sections logged, we described lithology at ~ 2 m intervals. All collected samples were processed for phytolith extraction, and about 1/3 yielded biosilica (phytoliths, diatoms, sponges etc.) in sufficient abundance and degree of preservation for quantitative analysis.

3.3.2 *Phytolith extraction*

To homogenize sediment samples, we crushed them with a hammer. A subsample of one gram per sample was further crushed using a mortar and pestle. We used standard methods for phytolith extraction (e.g., Strömberg, 2005), which include carbonate removal using hydrochloric acid; removal of large particles using a 250 μm sieve; removal of organic material using Schulze's solution; sediment deflocculation with a 53 μm sieve; and biosilica separation through centrifugation with zinc-bromide-based heavy liquid. Of the 127 lateral samples, 58 were productive. Of the 165 samples from logged sections, 87 came from the main section (CV) but only 16 yielded dense and well preserved enough phytoliths for analysis. Of the remaining 78 samples from other smaller sections, 25 yield phytoliths. A total number of 29 phytolith assemblage samples were selected for phytolith analysis with the objective of covering as much of the available stratigraphic interval as possible. 16 samples came from the main section (CV), 6 from other sections. Three additional samples (AR15-276, AR15-290, and AR15-292; KAN1, KAN2, and KARG-15-12 respectively in Trayler et al., 2019) corresponding to identified guide tuff layers within the section were collected and used for both phytolith analyses and U-Pb dating. In addition, to investigate regional variation in vegetation we chose several samples from two dated tuff layers from the bottom (CO) and top (Toba Blanca) of the section (Trayler et al.,

2019), which we were able to trace across multiple localities (i.e., CV, RB2, CI, and PLC for the CO tuff and RB3, MT, and CBT for the Toba Blanca tuff).

3.3.3 *Phytolith identification and counting*

For each sample we prepared two sets of slides; the first one using a fixed mounting medium (Cargille MeltMount®) and the second one using immersion oil, which allows rotation of phytoliths under the scope for optimal visualization of their 3D morphology. The first set was used to count all phytolith morphotypes and to compare relative abundances of open habitat indicators in the form of grass silica short cell phytoliths (GSSCP) and forest indicator phytolith morphotypes (FI). The second set was used for more detailed GSSCP identification. We counted GSSCP for only 16 of the 29 samples, corresponding to assemblages with moderate to high GSSCP abundance. GSSCP counts from the oil slides were then scaled within the GSSCP totals from the fixed slides to maintain proportional GSSCP abundances. Phytoliths were identified and counted at 1000 X magnification using a Nikon i80 compound microscope with mounted Nikon DS-Fi1 camera. At least 200 diagnostic phytoliths were counted per slide (both fixed and oil immersion) for statistical robustness (Pearsall, 2000; Strömberg, 2009). We scanned the remaining part of each slide for presence of rare (<5%) morphotypes which might be ecologically significant but might had been missed in the count; these morphotypes were not included in quantitative analysis but were taken into consideration in vegetation inference (see discussion in Crifò and Strömberg, 2019). Samples and slides are reposited at the Burke Museum of Natural History and Culture, University of Washington, Seattle (UWBM). Sediment replicas were deposited at Centro de Investigación Cientifica y de Transferencia Teconológica a la

Producción – Consejo Nacional de Investigaciones Cientificas y Técnicas (CICyTTP-CONICET) in Diamante, Argentina.

3.3.4 *Phytolith classification*

We followed the classification scheme outlined in Strömberg et al., (2013) which builds on Strömberg (2003) and later publications and is based on modern reference collections housed at UWBM. To make inferences about the composition and structure of phytolith assemblages we used the following phytolith categories, referred to as compound variables and corresponding roughly to plant functional group affinities (see Strömberg, 2003; Strömberg et al., 2013): 1) Forest indicator (FI) morphotypes found in palms (PALM), conifers (CONI), Zingiberales (ZINGI), and in woody or herbaceous dicots and ferns (Other FI); 2) grass silica short cell phytoliths (GSSCP), which are found exclusively in grasses (family Poaceae), and include taxa with affinity to closed habitats (CH, e.g. bambusoid and early-diverging grasses), open habitat grasses (OH) in the Pooideae tribe (POOID) and PACMAD clade (C_3 and C_4 grasses in the subfamilies Panicoideae (PANI), Aristidoideae, Chloridoideae (CHLOR), Micrairoideae, Arundinoideae, and Danthonioideae); 3) Phytoliths produced by wetland taxa (AQ) including sedges and *Equisetum*; and 4) non-diagnostic phytoliths (OTH), which includes non-diagnostic and unknown non- grass phytoliths (NDO), unknown GSSCP (OTHG) and non-diagnostic (potentially) grass phytolith (NDG). Non-diagnostic, and undeterminable non-grass (NDO) and certain grass and potential grass phytolith classes (NDG, and OTHG) were excluded for vegetation analyses because they are not considered diagnostic (Strömberg and McInerney, 2011). In addition, we created an additional class (cf. CONI) composed of the morphotypes combining the functional group CONI (see Strömberg, 2003) and phytoliths corresponding to the morphotype Blo-1, which tends to be very abundant in some samples. We assigned conifer

affinity to this morphotype based on its morphology, which has been described by several authors in many conifer taxa (e.g., An, 2016; Solomonova et al., 2017; Silantyeva et al., 2018; Lisztes-Szabó et al., 2019). Although this affinity is tentative, the prominence of this morphotype relative to others in our assemblages strongly suggest that it derives from a non-grass plant and likely a shrub or tree. The sum of forest indicator phytolith morphotypes corresponds to *FI tot.* In addition, we counted diatoms and sponge spicule relative abundance (compared to phytoliths) to assess proximity to local water sources (e.g., Piperno, 1988; Clarke, 2003). We did not use these data for quantitative analysis as sponges and diatoms were rare in most assemblages. Chrysophyte cysts were extremely rare in the SCF samples, and therefore they were not counted. To reconstruct vegetation structure and changes in habitat type we used the relative abundance of FI phytoliths (*FI-tot*) within the sum of GSSCPs and *FI tot* (categories 1 and 2 above), referred to as the *FI-t* ratio (*FI-t*), (Strömberg et al., 2007), and calculated as follows:

$$FI-t=\frac{FI_{tot}}{(FI_{tot}+GSSCPT)}\times 100 \quad (4.1)$$

However, we also took into consideration cross-sample variation in dominant types of forest indicators. To characterize grass communities and reconstruct changes in composition through time, we focused on major changes across samples in terms of relative abundances of phytoliths diagnostic of different grass groups, namely: 1) closed-habitat (CH) grasses (forest grasses in the Bambusoideae and early-diverging Poaceae lineages, all C_3 grasses; 2) POOID (open-habitat grasses in the Pooideae, all C_3); 2) CHLOR (C_4 open-habitat grasses in the Chloridoideae); 3) PANI (open-habitat grasses in the Panicoideae, mostly C_4); and 4) PACMAD (other open-habitat grasses in the PACMAD clade, including both C_3 and C_4 grasses). We used the proportion of closed-habitat grasses in the sum of GSSCP (herein called *CH grass*), calculated as:

$$CHgrass(\%) = \frac{CH}{\left(CH + POOID + CHLOR + PANI + PACMAD\right)} \quad (4.2)$$

In addition, within grass phytolith assemblages, we focus on the relative abundance of grass subfamilies that differ in the photosynthetic pathway (C_3, or C_4) in order to evaluate the presence of these grass types and their ecological significance in the SCF vegetation. In particular, we focused on the relative abundance of phytoliths of the Chloridoideae and the Panicoideae subclades where respectively all, and most species are C_4; of phytoliths from other subfamilies of the PACMAD clade, in which C_3 and C_4 grasses morphotypes are ndistinguishable; and phytoliths of the Pooideae subfamilies, which are C_3. However, we did not quantitatively estimate relative abundance of C_3 versus C_4 grasses in the SCF based on the fact that $\partial^{13}C$ values from herbivore tooth enamel (Trayler et al., in press), indicate that the SCF landscape was strongly dominated by C_3 vegetation.

3.3.5 *Statistical analyses*

To evaluate robustness of the calculated *FI-t* ratios we conducted bootstrap analysis (10,000 resampling with replacement) on the diagnostic phytolith counts yielding 95% confidence intervals (CIs) (Strömberg, 2009) using the statistical software R, version 3.6.1 (R Core Team, 2019). We use the same methods to evaluate robustness of the calculated proportion of open (OH) versus closed habitat grasses (CH) from the GSSCP counts.

Relative abundance of phytolith morphotypes in sediment assemblages do not directly represent plant abundances on the landscape because of phytolith *multiplicity* (a single species produces many different morphotypes) and *redundancy* (multiple species produce the same morphotype) (Rovner, 1971), as well as differential phytolith production among plant taxa (e.g.,

Hodson et al., 2005; Strömberg et al., 2016). Therefore, to make inferences about vegetation change through time we focus on changes in relative abundance rather than absolute abundance of phytolith morphotypes.

First, to visualize compositional difference in phytolith assemblages between samples we used Principal Component Analysis. The advantage of this technique is that it allows us to observe differences between samples (assemblages) in term of combinations of plant functional types rather than single functional types. PCA was conducted on variance-covariance matrices using the R function *prcomp* (R package *vegan*). We then tested for assemblage compositional trends through time by performing linear regressions of the sample scores on different PC axes against sample ages. The age for each sample was obtained using Trayler's et al.'s (2019) age model. We ran three PCA (PCA1-3) using different phytolith functional type classes. 1) In PCA1, we included phytoliths produced by the following plant functional types: other forest indicator (FI, excluding palms, and conifers), palms (PALM), GSSCP, conifers and potential conifers (cf. CONI); 2) In PCA2, , we included phytoliths produced by the following plant functional types: other forest indicator (Other FI, FI excluding palms and conifers), palms (PALM), conifers and potential conifers (cf. CONI), closed habit grasses (CH), Pooid grasses (POOID), PACMAD grasses (PACMAD), Panicoid grasses (PANI), and Chloridoid grasses (CHLOR); 3) In PCA3 , we included phytoliths produced by the following plant functional types: other forest indicator (Other FI, FI excluding palms and conifers), palms (PALM), phytoliths produced by conifers and potential conifers (cf. CONI), open habitat grasses (OH, i.e., POOID, PACMAD, PANI, CHLOR), and closed habitat grasses (CH). In addition, to test for changes over time in the relative abundances of individual phytolith groups (classes) we used generalized linear models with a logit link function (see Harris et al., 2017). The logit link

function converts morphotype proportions (Y_i, bounded between 0 and 1) into a logarithm of the ratio $Y_i /(1- Y_i)$, which can be then used in the regression. The analysis was performed using the R function *glm*. We calculated a 95% confidence interval band around the regression model, assuming that the error was normally distributed and multiplied 1.96 (0.975 quantile of the normal distribution) by the standard error of the regression fit. Further, to compensate for differences in sample size (*n*, number of phytoliths counted) between assemblages, we weighted the samples according to their size *n*, so that assemblages with bigger *n* (for which the calculated morphotype proportions can be considered more reliable) had more weight in the regression fit than assemblages with lower *n*.

3.4 RESULTS

3.4.1 *FI-t ratio and functional group composition of phytolith assemblages*

The *FI-t* ratio of phytolith assemblages (29 samples) ranges from 32% (95% CI = 26.3–37.8%) to 99.6% (95% CI = 98.8–100%), (table 4.1). Phytoliths indicative of wetland (AQ, e.g. sedges), and phytolith produced by plants in the order Zingiberales (ZINGI), commonly interpreted as indicative of moist habitats (e.g., Chen and Smith, 2013) are virtually absent in most assemblages or very rare (<5%). Phytoliths produced by palms (PALM), also typically interpreted as indicative of warm and moist habitats (see Strömberg et al., 2013; Dunn et al., 2015) are more abundant through the SCF, ranging from 0% to 50.6% (95% CI = 44.6–56.5%), but with relative abundance >20% in only five assemblages. Phytoliths assigned to the functional group cf. CONI are relatively abundant in all assemblages, ranging from 1.7% (95% CI = 0.4–3.4%) to 36.5% (95% CI = 30.3–43.1%) but with relative abundance >20% in only four assemblages. Phytolith produced by other forest indicators (FI except Palm, and cf. CONI) vary

widely in abundance, ranging from 12.9% (95% FI = 9.1-17%) to 92.1% (95% CI = 88.8-95.2%), with relative abundance >20% in all but two assemblages.

Regression analysis of all analyzed phytolith assemblages (29 samples) reveals a significant increase of the *FI-t* ratio over time ($p<<0.01$) (Fig. 3.3a). In addition, when the different classes within the FI functional group, as well as the class cf. CONI are studied separately, regression analysis reveals other, more detailed trends in phytolith assemblage composition through time. Specifically: 1) an increase in Other FI phytoliths (forest indicators excluding PALM, and cf. CONI classes), ($p<<0.01$) (Fig. 3.3b); 2) a slight decrease in PALM phytoliths ($p<<0.01$), (Fig. 3.3c); and 3) a decrease of GSSCP ($p<0.01$), (Fig. 3.3e). Abundance of cf. CONI phytolith morphotypes does not change significantly through time ($p>0.05$), (Fig. 3.3d).

In the seven samples taken along the oldest tuff layer, CO, the FI-t ratio shows considerable variability, from 32% (95% CI = 26.3-37.7%) to 94.2% (95% CI = 90.8-97.2), (Table 4.1). Phytolith produced by palms (PALM) range from 4.2% (95% CI = 1.9-6.9%) to 50.6% (95% CI = 44.6-56.4%) but with relative abundance >20% in only one sample, while phytoliths assigned to the functional group cf. CONI have relatively low abundance in all assemblages, ranging from 5.6% (95%% CI = 3-8.5%) to 12.6% (95% CI = 8.5-17.1%). Other forest indicator phytoliths (FI except PALM, and cf. CONI) range in relative abundance from 7.9% (95% CI = 4.8-11.2%) to 77% (95% CI = 71.4-82.4%). Along this stratigraphic layer no discernable gradients are observed across localities. However, within the locality CV, the two southernmost samples present lower *FI-t* than the northmost sample.

Along the younger tuff, Toba Blanca, the *FI-t* ratio of the three samples analyzed herein shows a smaller range of variation, from 65.7% (95% CI = 56.6-.74.7%) to 99.19% (95% CI =

98-100), (Table 4.2). Phytolith produced by palms (PALM) range from 0% to 23.1% (95% CI = 15.7-30.6%), whereas phytoliths assigned to the functional group cf. CONI have similar abundance in all assemblages, ranging from 11.1% (95% CI = 5.9-17.2%) to 12.5% (95% CI = 8.4-17.7%). Phytolith produced by Other FI phytoliths (FI except PALM, and cf. CONI) range from 30.4% (95% CI = 24.7-35.5%) to 87.1% (95% CI = 83.1-90.8%) but represent <40% in all but one assemblage. Along this stratigraphic layer no discernable gradients are observed across localities

3.4.2 *Grass phytolith (GSSCP) assemblage composition*

Within the counted GSSCP assemblages (15 samples), the proportion of closed-habitat grass phytolith verges between 1.3% (95% CI = 0–4.2%) and 24.4% (95% CI = 19.2–29.8%), (Table 4.2). Among open habitat grasses, Pooid (POOID) phytoliths are dominant with relative abundance >50% in all but two assemblages and ranging from 8.3% (95% CI = 4.1–13.2%) to 93.1% (95% CI = 90–95.9%). Within Pooid grass phytoliths, diagnostic morphotypes (POOID-D) are dominant, making up >50% is most assemblages, but with a range from 7.6% (95% CI = 3.5-12.2%) to 87.6% (95% CI = 83.7-91.3%); and non-diagnostic morphotypes (POOID-ND) are relatively rare, ranging from 0.7% (95% CI = 0-2.1%) to 12.9% (95% CI = 8.4-18%). Panicoid grass phytolith are rare (<6%) in all assemblages except one, in which they represent 36.1% of the GSSCP morphotypes (95% CI = 28.5-44%). Chloridoid (CHLOR) grass phytoliths are rare in most assemblages (<5%) and have low abundance in five assemblages, ranging from 5.7% (95% CI = 3–8.6%) to 12.3% (95% CI = 8.1–16.6%). Similarly, PACMAD grass phytoliths are rare (<5%) in many assemblages but are moderately abundant in others, ranging from 0% to 29.9% (95% CI = 22.5-37.5%).

Regression analyses show significant trends through time in GSSCP phytolith relative abundance within the entire phytolith assemblages (including FI), with both closed-habitat (CH) and open-habitat (OH) grass phytolith abundance decreasing through time ($p<<0.01$), (Fig. 3.3f-g). However, within GSSCP assemblages only closed-habitat grass (CH) abundance decreases through time ($p<<0.001$) (Fig. 3.3m). Conversely, open-habitat grass phytoliths of the Pooideae (POOID-D), and Panicoideae (PANI) subfamilies increases through time relative to other GSSCP classes (respectively $p<<0.001$, and $p<0.01$) (Fig. 3.3i-j). Chloridoideae (CHLOR), and PACMAD phytolith abundance does not show any significant trend through time ($p>0.05$) (Fig. 3.3k).

Along the oldest tuff layer (CO), GSSCP counts were obtained for five samples. These GSSCP assemblages are dominated by phytoliths of the subfamily Pooideae (POOID) ranging from 46.1% (95% CI = 29.9-65.4%) to 80.3% (95% CI = 74.7-85.8%) (Table 4.2). Panicoid phytoliths (PANI) are rare with <5% abundance in all but one assemblage, where they represent only 5.6% (95% CI = 3.1-8.4%). Similarly, Chloridoid phytoliths (CHLOR) represent <5% in most assemblages but one, where they are only moderately abundant (6.7%; CI=3.9-9.7%). PACMAD phytoliths show more variable abundance, ranging from 0 to 26.9% (95% CI = 11.5-46.1%). In the two samples collected along the younger tuff layer (Toba Blanca) for which GSSCP were counted, Pooideae phytoliths (POOID) show high abundances, namely 93.1% (95% CI = 90.1-95.9%), and 84.7 (95% CI = 79.7-89.2%) (Table 4.2); Panicoideae phytoliths (PANI) show very low abundance, respectively, 2.4% (95% CI = 0.7-4.2%), and 2.7% (95% CI = 0.9-4.9%), Chloridoideae phytoliths (CHLOR), respectively, 0% and 4.9% (95% CI = 2.5-8.1%), and PACMAD phytoliths, respectively 0 and 4.9% (95% CI = 2.5-8.1%).

3.4.3 *Principal component analysis of phytolith assemblages*

In PCA1 (Fig. 3.4a), the first PCA axis (PC1) accounts for 56.3% of the variation, with other forest indicators (Other FI), and potential conifer phytoliths (cf. CONI) loading positively, and GSSCP phytoliths (GSSCP) loading negatively along this axis. PC2 accounts for 27.2% of the variation, with palm phytoliths (PALM) and potential conifer phytoliths (cf. CONI) loading positively along this axis. In PCA2 (Fig. 3.4b), PC1 accounts for 37.4% of the variation, with other forest indicators (Other FI) loading positively, and diagnostic and non-diagnostic Pooideae phytoliths (POOID-D, and POOID-ND) loading negatively along this axis. Other phytolith classes (CH, PANI, and CHLOR) do not load heavenly on any of the PC axes. PC2 accounts for 20.9% of the variation, with palm phytoliths (PALM) loading positively and PACMAD phytoliths (PACMAD) loading negatively along this axis. In PCA3 (Fig. 3.4c), PC1 accounts for 51.5% of the variation, with other forest indicator (Other FI) and potential conifer phytoliths (cf. CONI) loading positively, and both open-habitat (OH) and closed-habitat (CH) grass phytoliths loading negatively on this axis. PC2 accounts for 26.1% of the variation, with other forest indicator phytoliths (Other FI) loading positively and palm phytoliths (PALM) loading negatively along this axis. For all three PCAs, PC1 correlates negatively with time ($p<0.01$, $R^2=0.21$ for PCA1; $p<0.001$, $R^2=0.57$ for PCA2; and $p<0.001$, $R^2=0.58$ for PCA3). PC2 does not correlate with time for any of the PCAs. Further, for all three PCSs, across the ordination space phytolith assemblages from the lower part of the section (in pink and purple, toward the negative end of PC1 in Fig. 3.4a-c) separate from phytolith assemblages from the upper part of the section (in light and dark green, toward the positive end of PC1) along PC1. This separation along axis 1 is clearer in PCA2 and PCA3 than in PCA1.

3.5 DISCUSSION

3.5.1 *Vegetation change through time*

Phytolith analysis of 29 samples from eight localities across >50 km of the coastal Santa Cruz Formation points to a complex signal of vegetation change. Although the interpretation of the temporal patterns of vegetation change is complicated by spatial variability in the phytolith signal within and between localities, some trends emerge that allow us to address plant community response at the onset of the MMCO. Overall, the samples analyzed herein show a trend of increasing proportion of forest indicators (*FI-t* ratio) and decreasing proportions of grass and palm phytoliths through time. Within grass communities, the proportion of open habitat grasses (OH) increases whereas the proportion of closed habitat grasses (CH) decreases through time.

The overall increase in the proportion of forest indicators appears to be in agreement with the re-greening pattern previously suggested for the MMCO of Patagonia based on phytolith and pollen analysis (Iglesias et al., 2011; Palazzesi and Barreda, 2012; Dunn et al., 2015). However, previous work in the Eocene-Miocene of Patagonia suggests that phytolith assemblages dominated by forest indicators are not necessarily indicative of closed habitats (Dunn et al., 2015). Specifically, in some phytolith assemblages Dunn et al. (2015) observed that high *FI-t* ratios (in particular driven by palm and woody dicot phytoliths) were associated with relatively low rLAI (reconstructed Leaf Area Index, a measure of canopy openness inferred from the morphology of leaf epidermis phytoliths of non-grass taxa). Because of the uncertainty in the interpretation of FI phytoliths in the Cenozoic of Patagonia, we chose to interpret our data in the light of recent climate reconstructions based on isotopic data of the SCF (Trayler et al., in review), rather than on traditional interpretation of phytolith functional types composition of the

assemblages (e.g. Strömberg et al., 2013). Recent climate reconstruction of the SCF point to a shift toward drier and warmer conditions at the onset of the MMCO, which would not be consistent with a regreening scenario. Specifically, at the base of the section (~17.4 Ma), reconstructed mean annual temperatures (~20˚C) and precipitations (~700±130 mm/yr) are similar to levels recorded in modern-day shrublands and woodlands in subtropical or warm temperate climates. Un South America, a similar climate is experienced today in the central and western part of the Chaco ecoregion in Argentina (e.g., Adamoli et al., 1990). In the upper part of the section (~17-16.5 ma), corresponding to the onset of the MMCO, reconstructed temperature (MAT ~25˚C) and precipitation (MAP = ~340±90 to 500 mm/yr), are reminiscent with conditions experienced in modern-days subtropical, semi-arid climates, characterized by scrubby, dry-adapted vegetation. Similar conditions characterize the Caatinga of northeastern Brazil today. This ecoregion consists of a mosaic of vegetation types including dry thorn forest shrubland/woodland, and open shrubby vegetation with mostly C_4 grasses (da Costa et al., 2007).

Several patterns in phytolith assemblage composition observed at the SCF support this drying scenario as opposed to a regreening scenario. First, while our analysis shows an overall increase in the forest component of the vegetation (FI functional group), it also reveals that within this functional group, abundance of palms, usually indicative of moist conditions and closed forest habitats (but see Dunn et al., 2015; Kohn et al., 2015) decreased through time. Second, within grass communities the relative proportion of closed-habitat grasses– which are plesiomorphically adapted to shady (hence wetter) conditions or more humid microhabitats, such as forest understorey (Gallaher et al., 2019)–decreases through time. Conversely, the relative proportion of open habitat grasses of the subfamilies Pooideae (C_3, temperate grasses) and Panicoideae (mostly C_4, mesic to dry-adapted grasses) increases. These trends are in agreement

with a warmer and more arid climate recorded by isotope analysis (Trayler et al., in press). On the other hand, within the grass community, relative abundance of potential C_4 grasses (Chloridoideae, Panicoideae, and other PACMADs), which have an advantage in hot, dry environments with warm-season precipitation, does not increase through time (with the exception of Panicoideae).

To explain the discrepancy between increased *FI-t* ratios and reconstructed (warmer and drier) climate, we suggest that this trend reflects the expansion of a dry-adapted, woody component of the vegetation perhaps corresponding to a non-analogue woodland or shrubland-like vegetation, including palm shrublands. To explain the lack of an increase in potential C_4 grasses given the reconstructed climate, we suggest that the majority of PACMAD grasses in the SCF were likely C_3, and therefore not adapted to dry conditions (Ibrahim et al., 2008). This is consistent with previous work suggesting that C_4 grasses, and grasslands expansion in Patagonia did not take place during the onset of the MMCO –and conceivably not until the late Miocene (Barreda and Palazzesi, 2007, 2014; Iglesias et al., 2011; Strömberg et al., 2013; Dunn et al., 2015).

How can we reconcile our vegetation reconstruction with other phytolith data from Patagonia (i.e., Strömberg et al., 2013, and Dunn et al, 2015), and in particular with the observed regreening pattern? First, it is possible that the differences in vegetation type between the SCF and other Patagonian localities are driven by regional differences in precipitation patterns or other climatic factors. In fact, the SCF represent a coastal floodplain environment whereas other Patagonians sites are located further inland (Dunn et al., 2015) or northward, and closer to the Andes foothills (Strömberg et al., 2013; Dunn et al., 2015). Second, differences between the vegetation recorded at the SCF and other localities in Patagonia (Strömberg et al., 2013; Dunn et

al., 2015) can at least in part be explained by the fact that these localities cover different time intervals. The phytolith record from Gran Barranca (Strömberg et al, 2013), spans a long interval of the Cenozoic, staring (~43 Ma - ~18 Ma), preceding the SCF phytolith record by 0.5 Ma. Similarly, the vegetation record described by Dunn et al., (2015) spans an interval from ~49 to ~12 Ma but lacks samples between ~18.8 and 15.7 Ma, meaning that there is no overlap with the time interval corresponding to the SCF phytolith record. Therefore, we argue that previous vegetation reconstructions should be viewed as complimentary pieces of information that can help us better understand vegetation change at the onset of the MMCO in the Santa Cruz Formation, rather than contradictory vegetation reconstructions.

Phytolith assemblage data and rLAI reconstruction primarily from Gran Barranca point to the existence of nonanalog palm shrubland with a discontinuous canopy during the middle Eocene-Oligocene, shifting to open forests without a continuous grass understory in the early Miocene (Strömberg et al., 2013; Dunn et al., 2015). Grass abundance is much higher at the base of the SCF (~17.5 Ma) compared to the early Miocene (~18 Ma) of Gran Barranca, perhaps continuing the trend towards drier habitats observed at Gran Barranca. Similarly, at the base of the SCF palm abundance is lower compared to Gran Barranca, consistent with a drier climate. It is possible that this trend toward drier habitats was enhanced in the early middle Miocene by the Andes uplift and the subsequent orographic rain shadow driving long term aridification of Patagonia (Blisniuk et al., 2005; Palazzesi et al., 2014)– although the magnitude of the uplift between 16.8 Ma and 17.6 Ma in unclear (Trayler et al., in review). Changes in phytolith assemblage composition and increased *FI-t* ratio in the upper part of the SCF marks a further shift towards more dry-adapted vegetation dominated by a woody component at the onset of the MMCO. This record predates by about 1 Ma a brief regreening phase observed in the Patagonian

lowlands according to rLAI data (~ 15.7 to 14.6 Ma) that coincides with an increase in atmospheric pCO_2 and a reduction in the extent of the Antarctic ice sheet. Although these records could simply reflect different chapters of the vegetation history of southern Patagonia, there is evidence to suggest that they record regional differences in plant communities. Indeed, phytolith assemblages from the SCF are remarkably distinct from those from inland Patagonia (Dunn et al., 2015). The latter completely lacks the phytolith morphotypes potentially indicative of conifers (cf. CONI) that characterize many SCF samples, pointing to major floristic differences between the coastal and inland Patagonian vegetation. Therefore, it is likely that the phytolith records from these two regions reflect regional differences in vegetation responses to the MMCO, whereby higher temperatures and pCO_2 as well as wetter conditions (Kohn et al., 2015) favored a regreening of the inland region, whereas a stronger aridification effect favored more open, dry adapted vegetation towards the coast.

However, we cannot completely rule out the possibility that a later regreening phase, not encompassed by the available SCF record, took place after initial aridification and habitat opening. In this context, it is important to note, that Trayler et al., (in press) hypothesized that regional aridification in the SCF might have taken place at the onset of the MMCO as a result of a negative feedback induced by increasing global temperatures and reduced Antarctic ice volume; these authors also highlight that the drop in precipitations was shortly followed by more stable and higher precipitation (coupled with higher temperatures). Perhaps, under such conditions the hypothesis of a brief regreening phase in the SCF vegetation could find support. Unfortunately, the lack of more recent data from the SCF does not allow us to test whether this seemingly reversed trend in precipitation (following aridification) continued over a long time interval, and how vegetation responded to it. However, it has been well documented that the

uplift of the Patagonian Andes has caused a rain shadow effect on Patagonia since ~14-12 Ma with profound effects on the landscape. It is therefore reasonable to expect that climate and vegetation aridification starting at the onset of the MMCO has continued up to the present in coastal Patagonia.

3.5.2 *Heterogeneity across the landscape*

The variability of phytolith assemblage composition within the tuff layer CO suggest that plant communities were heterogeneous both in terms of grass cover, and relative abundances and types of forest indicator phytoliths, including palms and potential conifers. In addition, although GSSCP abundance is high in most assemblages, grasses are associated with different FI elements in different samples, adding to complexity of pattern. For example, at RB2, phytolith assemblages (sample 21) are dominated by palms and grasses, with low abundance of potential conifer phytoliths and other forest indicators. Such high palm abundance in combination with grasses is unique to this locality. Scanning of several additional samples (all collected along this layer at the same locality) shows that this pattern is consistent at RB2 (over at least ~400 meters) and suggests that unique plant communities, perhaps corresponding to a palm shrubland habitat, dominated locally. In addition, tuff samples within Cañodón de Las Vacas show similar variation between the northern and southern side of the outcrop (separated by several hundred meters), with higher grass abundance toward the south (samples 22 and 24) and lower grass abundance toward the north (sample 23). This pattern also points to heterogeneous habitats characterized by more open patches with higher grass abundance and patches dominated by woody vegetation. Phytolith assemblage composition is less variable along the Toba Blanca tuff, pointing to more homogeneous habitats. However, this trend is to be interpreted with caution, because only three

assemblages were analyzed (and GSSCP were counted in only two assemblages). Therefore, additional samples are needed to characterize habitat heterogeneity along this stratigraphic level.

3.5.3 *Caveats*

Taphonomic biases can sometimes throw off interpretations of vegetation change. For example, the original phytolith assemblage composition can be altered by size sorting or preferential preservation of certain forms over others (e.g., round vs. elongate), resulting in inaccurate vegetation reconstructions. We exclude the influence of such biases on our interpretations because all the samples that were quantitatively analyzed in this study yielded abundant phytoliths with good preservation and no sign of size sorting or preferential preservation. Likewise, the influence of taphonomic biases linked to certain depositional environments (such as ash deposits) on the observed trend in phytolith assemblages can be excluded for several reasons. First, although many of our samples were collected from tuff layers (which might be indicative of a disturbance event), the composition of the resulting phytolith assemblages does not seem to reflect vegetation primary vegetation succession that follows disturbance. All phytolith functional groups are represented in the assemblages, including taxa that are not typically indicative of disturbed habitats (e.g., palms show relative abundance up to ~50%) (e.g., Baker and Couvreur, 2013). Second, if phytolith assemblages in tuffs represented just the immediate colonization of ash-covered ground following disturbance, we would predict a strong dominance of a single component of the vegetation, such as grasses. In contrast, sampled phytolith assemblages within a single layer (e.g., tuff CO, tuff Toba Blanca) are highly heterogeneous, pointing to marked spatial variation in vegetation. We interpret this pattern to indicate that soil formation in these tuffaceous layers occurred over hundreds of years or more;

thus, phytolith assemblages contained within them should capture primarily "climax" vegetation. Third, the ash layers of the SCF do not appear to represent local volcanic events that would prompt vegetation succession in response to disturbance such as deforestation. Rather, these reworked tuff layers likely result from fluvial deposition of pyroclastic material originating from the west/southwest in the Andean ranges (Cuitiño and Scasso, 2013).

Although our data indicate vegetation heterogeneity across the coastal localities of the Santa Cruz formation, our vegetation reconstruction is based on only a few samples across a >50 km long region. In addition, at each locality our samples do not cover the complete temporal sequence, and lateral sampling was only possible at a few localities. It is therefore likely that we cannot entirely detect differences in local vegetation and climate across the SCF. In fact, a single sample at a single stratigraphic level might only be representative of a particular microhabitat, without capturing the heterogeneity and complexity of a mosaic vegetation. For this reason, more extensive sampling is needed across localities in order to reconstruct regional and local vegetation heterogeneity.

Moreover, to improve our interpretation of past vegetation change in response to climate change, much work is needed on the taxonomic and ecological significance of phytoliths from different functional groups such as palms (e.g., Armos et al., 2019), as well as on the evolutionary history and ecology of important taxa such as open habitat grasses (e.g., Gallaher et al., in prep).

3.6 CONCLUSION

We analyzed quantitatively assemblage composition of phytoliths extracted from paleosols and tuff layers along a well dated stratigraphic sequence of the Santa Cruz Formation,

Patagonia, to investigate vegetation response to climate change at the onset of MMCO. In addition to uncovering temporal trends in vegetation, we also investigated local vegetation heterogeneity by studying phytoliths samples collected across a single stratigraphic level. Our findings indicate the presence of mixed and heterogeneous vegetation right before the initiation of the MMCO, in a subtropical warm and dry climate, as inferred from isotopic data. Habitats were characterized by patches of woodland/open woodland or shrubland, including palm shrubland, where C_3 grasses were abundant. Grass communities were dominated by open-habitat grasses (mostly Pooideae), although some closed habitat grasses (of the subfamily Bambusoideae) were present. The onset of the MMCO appeared to have triggered drier and warmer conditions leading to an overall shift in vegetation. While dry-adapted tree and shrub vegetation expanded, grass communities contracted. This contraction suggests that SCF grass communities were not dry-adapted, and that grasses in the PACMAD clade were likely primarily C_3. Thus, we hypothesize that these C_3 grasses were outcompeted under abruptly drying climatic conditions, giving way to a largely non-analog woody dry-adapted vegetation. Our results support previous hypotheses based on the study of the SCF flora, and faunal ecology and diet, that the vegetation of the SCF was characterized by mixed habitats, with open-woodland or mixed woodland scrubland environments, and that closed-canopy vegetation was lacking. In addition, our result support previous hypothesis that the expansion of grasslands in Patagonia did not take place until the late Miocene. Our data also point to differences in vegetation response within Patagonia, whereby sites further to the west point to a brief regreening followed by aridification, our data from the SCF coastal localities indicate that aridification favored xerophytic vegetation adapted to semi-arid conditions since the onset of the MMCO and onward.

In light of the vegetation patterns described above, it is important to note that the SCF vegetation, although characterized by the presence of modern lineages, might represent a habitat type that has no modern analogue in term of ecological characteristics of single plant taxa and communities. In particular, it is possible that the ecology and climatic ranges of plant lineages like PACMAD grasses and palms differed from those of their modern relatives. Linking climate and vegetation proxies (such as phytolith analysis) allow us to refine our understanding of ancient plant ecology and its interaction with the environment, perhaps also broadening our perspective on projected ecosystem response to climate change.

ACKNOWLEDGMENTS

We are grateful to Lucas Cataldi, Juan Carlos Fernicola, Santiago Hernández Del Pino, Nahuel Muñoz, Néstor Toledo for their help in the field, and Luciano Zapata for providing sample AR15-292. We also thank Madeline Mamer, Katherine Anderson, Ron Eng, and the Burke Museum for helping with sample cataloging. Additional thanks to the undergraduate researchers in the Strömberg lab, Thy Huyhn, Erin Sofinoski, Sarah Larson, Kevin Jackson, Matt Butrim, Una O'Connell, Matt Bloch, Gabrielle Alampay, Kirsten Olson, Kristen Hamel, and Alexander Arrendale, who helped with phytolith sample processing. Funding for this project was provided by Benjamin Hall Washington Research Foundation, University of Washington, Philosophical Society of America, Geological Society of America, Paleontological Society, and Quaternary Research Center awards to C.C.; National Science foundation grant number EAR-1349530 to C.A.E.S; National Science foundation grant number EAR-1349741 to R.F.K., National Science foundation grant number EAR-1349749 to M.J.K; and grants PITC 2013-0389, and UNLP 11/N750 (SVF) to M.S.B and S.F.V.

REFERENCES

Adamoli, J., Sennhauser, E., Acero, J.M., Rescia, A., 1990. Stress and disturbance: vegetation dynamics in the dry Chaco region of Argentina. Journal of biogeography, 491-500.

An, X.-H., 2016. Morphological characteristics of phytoliths from representative conifers in China. Palaeoworld 25, 116-127.

Armos, B., Lavin, S., Akbar, S., Brightly, W., Crifò C., Gallaher, T., Lowe, A., Novello, A., Wilson, P., Strömberg, C.A.E., 2019. The utility of palm phytoliths for inferring the evolution and paleoecology of Arecaceae. Botany 2019, Tucson, AZ, USA .

Baker, W.J., Couvreur, T.L., 2013. Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. I. Historical biogeography. Journal of Biogeography 40, 274-285.

Badgley, C., Fox, D.L., 2000. Ecological biogeography of North American mammals: species density and ecological structure in relation to environmental gradients. Journal of Biogeography 27, 1437-1467.

Badgley, C., Smiley, T.M., Terry, R., Davis, E.B., DeSantis, L.R.G., Fox, D.L., Hopkins, S.S.B., Jezkova, T., Matocq, M.D., Matzke, N., McGuire, J.L., Mulch, A., Riddle, B.R., Roth, V.L., Samuels, J.X., Strömberg, C.A.E., Yanites, B.J., 2017. Biodiversity and Topographic Complexity: Modern and Geohistorical Perspectives. Trends Ecol Evol 32, 211-226.

Bargo, M., Toledo, N., Vizcaíno, S., 2012. Paleobiology of the Santacrucian sloths and anteaters (Xenarthra, Pilosa). Early Miocene Paleobiology in Patagonia: High-latitude Paleocommunities of the Santa Cruz Formation. Cambridge University Press, Cambridge, 216-242.

Barreda, V., Palazzesi, L., 2007. Patagonian vegetation turnovers during the Paleogene-early Neogene: origin of arid-adapted floras. The botanical review 73, 31-50.

Barreda, V.D., Palazzesi, L., 2014. Response of plant diversity to Miocene forcing events: the case of Patagonia. Ann. Missouri Bot. Gard.

Beerling, D.J., Royer, D.L., 2011. Convergent Cenozoic CO_2 history. Nature Geoscience 4, 418-420.

Blisniuk, P.M., Stern, L.A., Chamberlain, C.P., Idleman, B., Zeitler, P.K., 2005. Climatic and ecologic changes during Miocene surface uplift in the Southern Patagonian Andes. Earth and Planetary Science Letters 230, 125-142.

Bown, T.M., Fleagle, J.G., 1993. Systematics, biostratigraphy, and dental evolution of the Palaeothentidae, later Oligocene to early-middle Miocene (Deseadan-Santacrucian) caenolestoid marsupials of South America. Memoir (The Paleontological Society), 1-76.

Brea, M., Zucol, A.F., Iglesias, A., 2012. Early Miocene Paleobiology in Patagonia: Fossil plant studies from late Early Miocene of the Santa Cruz Formation: paleoecology and paleoclimatology at the passive margin of Patagonia, Argentina.

Cassini, G.H., Vizcaíno, S.F., 2012. An Approach to the Biomechanics of the Masticatory Apparatus of Early Miocene (Santacrucian Age) South American Ungulates (Astrapotheria, Litopterna, and Notoungulata): Moment Arm Estimation Based on 3D Landmarks. Journal of Mammalian Evolution 19, 9-19.

Clarke, J., 2003. The occurrence and significance of biogenic opal in the regolith. Earth-Science Reviews 60, 175-194.

Cuitiño, J.I., Scasso, R.A., 2010. Sedimentología y paleoambientes del Patagoniano y su transición a la Formación Santa Cruz al sur del Lago Argentino, Patagonia Austral. Revista de la Asociación Geológica Argentina 66, 406-417.

Cuitiño, J.I., Scasso, R.A., 2013. Reworked pyroclastic beds in the early Miocene of Patagonia: Reaction in response to high sediment supply during explosive volcanic events. Sedimentary Geology 289, 194-209.

Cuitiño, J.I., Fernicola, J.C., Kohn, M.J., Trayler, R., Naipauer, M., Bargo, M.S., Kay, R.F., Vizcaíno, S.F., 2016. U-Pb geochronology of the Santa Cruz Formation (early Miocene) at the Río Bote and Río Santa Cruz (southernmost Patagonia, Argentina): Implications for the correlation of fossil vertebrate localities. Journal of South American Earth Sciences 70, 198-210.

Cuitiño, J.I., Vizcaíno, S.F., Bargo, M.S., Aramendía, I., 2019. Sedimentology and fossil vertebrates of the Santa Cruz Formation (early Miocene) in Lago Posadas, southwestern Patagonia, Argentina. Andean Geology 46.

Crifò, C., Strömberg, C.A., 2019. Small-scale spatial resolution of the soil phytolith record in a Neotropical rainforest and a dry forest in Costa Rica: applications to the deep-time fossil phytolith record. Palaeogeography, Palaeoclimatology, Palaeoecology.

Croft, D.A., 2013. What constitutes a fossil mammal community in the early Miocene Santa Cruz Formation? Journal of Vertebrate Paleontology 33, 401-409.

da Costa, R.C., de Araújo, F.S., Lima-Verde, L.W., 2007. Flora and life-form spectrum in an area of deciduous thorn woodland (caatinga) in northeastern, Brazil. Journal of Arid Environments 68, 237-247.

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

Fleagle, J.G., Bown, T., Swisher, C., Buckley, G., 1995. Age of the Pinturas and Santa Cruz formations, Congreso Argentino de Paleontología y Bioestratigrafía pp. 129-135.

Flower, B.P., Kennett, J.P., 1994. The middle Miocene climatic transition: East Antarctic ice sheet development, deep ocean circulation and global carbon cycling. Palaeogeography, palaeoclimatology, palaeoecology 108, 537-555.

Finarelli, J.A., Badgley, C., 2010. Diversity dynamics of Miocene mammals in relation to the history of tectonism and climate. Proceedings of the Royal Society of London B: Biological Sciences 277, 2721-2726.

Fosdick, J.C., Grove, M., Hourigan, J.K., Calderon, M., 2013. Retroarc deformation and exhumation near the end of the Andes, southern Patagonia. Earth and Planetary Science Letters 361, 504-517.

Foster, G.L., Lear, C.H., Rae, J.W.B., 2012. The evolution of pCO_2, ice volume and climate during the middle Miocene. Earth and Planetary Science Letters 341-344, 243-254.

Gallaher, T.J., Adams, D.C., Attigala, L., Burke, S.V., Craine, J.M., Duvall, M.R., Klahs, P.C., Sherratt, E., Wysocki, W.P., Clark, L.G., 2019. Leaf shape and size track habitat transitions across forest-grassland boundaries in the grass family (Poaceae). Evolution.

Gayó, E., Hinojosa, L.F., Villagrán, C., 2005. On the persistence of Tropical Paleofloras in central Chile during the Early Eocene. Review of Palaeobotany and Palynology 137, 41-50.

Harris, E.B., Strömberg, C.A., Sheldon, N.D., Smith, S.Y., Vilhena, D.A., 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.

Hodson, M.J., White, P.J., Mead, A., Broadley, M.R., 2005. Phylogenetic Variation in the Silicon Composition of Plants. Annals of Botany 96, 1027-1046.

Holbourn, A., Kuhnt, W., Kochhann, K.G.D., Andersen, N., Sebastian Meier, K.J., 2015. Global perturbation of the carbon cycle at the onset of the Miocene Climatic Optimum. Geology 43, 123-126.

Ibrahim, D.G., Gilbert, M.E., Ripley, B.S., Osborne, C.P., 2008. Seasonal differences in photosynthesis between the C3 and C4 subspecies of *Alloteropsis semialata* are offset by frost and drought. Plant, Cell & Environment 31, 1038-1050.

Iglesias, A., Artabe, A.E., Morel, E.M., 2011. The evolution of Patagonian climate and vegetation from the Mesozoic to the present. Biological Journal of the Linnean Society 103, 409-422.

Janis, C.M., Damuth, J., Theodor, J.M., 2000. Miocene ungulates and terrestrial primary productivity: Where have all the browsers gone? Proceedings of the National Academy of Sciences 97, 7899-7904.

Janis, C.M., Damuth, J., Theodor, J.M., 2004. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. Palaeogeography, Palaeoclimatology, Palaeoecology 207, 371-398.

Kay, R.F., Vizcaíno, S.F., Bargo, M.S., Perry, J.M.G., Prevosti, F.J., Fernicola, J.C., 2008. Two new fossil vertebrate localities in the Santa Cruz Formation (late early – early middle Miocene, Argentina), ~51° South latitude. Journal of South American Earth Sciences 25, 187-195.

Kay, R.F., Vizcaino, S.F., Bargo, M.S., 2012a. A review of the paleoenvironment and paleoecology of the Miocene Santa Cruz Formation, in: Vizcaíno, S.F., Kay, R.F., Bargo,

M.S. (Eds.), Early Miocene Paleobiology in Patagonia: high-latitude paleocommunities of the Santa Cruz Formation. Cambridge: Cambridge University Press, pp. 331-365.

Kay, R.F., Perry, J.M., Malinzak, M., Allen, K.L., Kirk, E.C., Plavcan, J.M., Fleagle, J., 2012b. The paleobiology of Santacrucian primates. Early Miocene Paleobiology in Patagonia: High-Latitude Paleocommunities of the Santa Cruz Formation, 306-330.

Kohn, M.J., Josef, J.A., Madden, R., Kay, R., Vucetich, G., Carlini, A.A., 2004. Climate stability across the Eocene-Oligocene transition, southern Argentina. Geology 32, 621.

Kohn, M.J., Fremd, T.J., 2008. Miocene tectonics and climate forcing of biodiversity, western United States. Geology 36, 783-786.

Kohn, M.J., Strömberg, C.A.E., Madden, R.H., Dunn, R.E., Evans, S., Palacios, A., Carlini, A.A., 2015. Quasi-static Eocene–Oligocene climate in Patagonia promotes slow faunal evolution and mid-Cenozoic global cooling. Palaeogeography, Palaeoclimatology, Palaeoecology 435, 24-37.

Kürschner, W.M., Kvaček, Z., Dilcher, D.L., 2008. The impact of Miocene atmospheric carbon dioxide fluctuations on climate and the evolution of terrestrial ecosystems. Proceedings of the National Academy of Sciences 105, 449-453.

Lisztes-Szabó, Z., Braun, M., Csík, A., Pető, Á., 2019. Phytoliths of six woody species important in the Carpathians: characteristic phytoliths in Norway spruce needles. Vegetation History and Archaeobotany, 1-14.

Marshall, L.G., 1976. Fossil localities for Santacrucian (early Miocene) mammals, Santa Cruz Province, southern Patagonia, Argentina. Journal of Paleontology, 1129-1142.

Marshall, L.G., Drake, R.E., Curtis, G.H., Butler, R.F., Flanagan, K.M., Naeser, C.W., 1986. Geochronology of type Santacrucian (middle Tertiary) land mammal age, Patagonia, Argentina. The Journal of Geology 94, 449-457.

Matheos, S.D., Raigemborn, M.S., 2012. Sedimentology and paleoenvironment of the Santa Cruz Formation, in: Vizcaíno, S.F., Kay, R.F., Bargo, M.S. (Eds.), Early Miocene Paleobiology in Patagonia: High Latitude Paleocommunities of the Santa Cruz Formation. Cambridge University Press, Cambridge, UK, pp. 59-82.

Montalvo, C.I., Raigemborn, M.S., Tomassini, R.L., Zapata, L., Bargo, M.S., Uncal, M.C.M., Vizcaíno, S.F., 2019. Floodplain Taphonomic Mode of Early Miocene Vertebrates of Southern Patagonia, Argentina. Palaios 34, 105-120.

Palazzesi, L., Barreda, V., 2012. Fossil pollen records reveal a late rise of open-habitat ecosystems in Patagonia. Nature communications 3, 1294.

Palazzesi, L., Barreda, V.D., Cuitiño, J.I., Guler, M.V., Telleria, M.C., Ventura Santos, R., 2014. Fossil pollen records indicate that Patagonian desertification was not solely a consequence of Andean uplift. Nat Commun 5, 3558.

Pearsall, D., 2000. Paleoethnobotany: A Handbook of Procedures Academic Press. San Diego, California.

Perkins, M. E., Fleagle, J. G., Heizler, M. T., Nash, B., Bown, T., Tauber, A., & Dozo, M., 2012. Tephrochronology of the Miocene Santa Cruz and Pinturas Formations, Argentina. In S. F. Vizcaíno, R. F. Kay, & M. S. Bargo (Eds.), *Early Miocene Paleobiology in Patagonia: High-Latitude Paleocommunities of the Santa Cruz Formation* (pp. 23-40).

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

Pound, M.J., Haywood, A.M., Salzmann, U., Riding, J.B., 2012. Global vegetation dynamics and latitudinal temperature gradients during the Mid to Late Miocene (15.97–5.33 Ma). Earth-Science Reviews 112, 1-22.

R Core Team, 2019. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012. URL https://www. R-project. org.

Raza, S.M., Barry, J.C., Meyer, G.E., Martin, L., 1984. Preliminary report on the geology and vertebrate fauna of the Miocene Manchar Formation, Sind, Pakistan. Journal of Vertebrate Paleontology 4, 584-599.

Raigemborn, M.S., Matheos, S.D., Krapovickas, V., Vizcaíno, S.F., Bargo, M.S., Kay, R.F., Fernicola, J.C., Zapata, L., 2015. Paleoenvironmental reconstruction of the coastal Monte Léon and Santa Cruz formations (Early Miocene) at Rincón del Buque, Southern Patagonia: A revisited locality. Journal of South American Earth Sciences 60, 31-55.

Raigemborn, M.S., Krapovickas, V., Beilinson, E., Peral, L.E.G., Zucol, A.F., Zapata, L., Kay, M.R.F., Bargo, M.S., Vizcaíno, S.F., Sial, A.N., 2018a. Multiproxy studies of Early Miocene pedogenic calcretes in the Santa Cruz Formation of southern Patagonia, Argentina indicate the existence of a temperate warm vegetation adapted to a fluctuating water table. Palaeogeography, palaeoclimatology, palaeoecology 500, 1-23.

Raigemborn, M.S., Krapovickas, V., Zucol, A.F., Zapata, L., Beilinson, E., Toledo, N., Perry, J., Lizzoli, S., Martegani, L., Tineo, D., 2018b. Paleosols and related soil-biota of the early Miocene Santa Cruz Formation (Austral-Magallanes Basin, Argentina): a multidisciplinary approach to reconstructing ancient terrestrial landscapes. Latin American Journal of Sedimentology and Basin Analysis 25.

Ritchie, M.E., Olff, H., 1999. Spatial scaling laws yield a synthetic theory of biodiversity. Nature 400, 557.

Shevenell, A.E., Kennett, J.P., Lea, D.W., 2008. Middle Miocene ice sheet dynamics, deep-sea temperatures, and carbon cycling: A Southern Ocean perspective. Geochemistry, Geophysics, Geosystems 9.

Selkin, P.A., Strömberg, C.A.E., Dunn, R., Kohn, M.J., Carlini, A.A., Davies-Vollum, K.S., Madden, R.H., 2015. Climate, dust, and fire across the Eocene-Oligocene transition, Patagonia. Geology 43, 567-570.

Silantyeva, M., Solomonova, M., Speranskaja, N., Blinnikov, M.S., 2018. Phytoliths of temperate forest-steppe: A case study from the Altay, Russia. Review of Palaeobotany and Palynology 250, 1-15.

Simpson, G.G., 1940. Review of the mammal-bearing Tertiary of South America. Proceedings of the American Philosophical Society, 649-709.

Smiley, T.M., Hyland, E.G., Cotton, J.M., Reynolds, R.E., 2018. Evidence of early C4 grasses, habitat heterogeneity, and faunal response during the Miocene Climatic Optimum in the Mojave Region. Palaeogeography, Palaeoclimatology, Palaeoecology 490, 415-430.

Solomonova, M.Y., Silantyeva, M., Speranskaya, N.Y., 2017. Phytolith research in the South of Western Siberia. Ukrainian Journal of Ecology 7.

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

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 of the United States of America 102, 11980-1198.

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., Friis, E.M., Liang, M.-M., Werdelin, L., Zhang, Y.-L., 2007. 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., McInerney, F.A., 2011. The Neogene transition from C-3 to C-4 grasslands in North America: assemblage analysis of fossil phytoliths. Paleobiology 37, 50-71.

Strömberg, C.A., Dunn, R.E., Madden, R.H., Kohn, M.J., Carlini, A.A., 2013. Decoupling the spread of grasslands from the evolution of grazer-type herbivores in South America. Nature communications 4, 1478.

Strömberg, C.A., Dunn, R.E., Madden, R.H., Kohn, M.J., Carlini, A.A., 2014. Evolution of Grazer Morphologies in the Absence of Grasslands in Southern South America. The Paleontological Society Special Publications 13, 113-113.

Strömberg, C.A.E., Di Stilio, V.S., Song, Z., 2016. Functions of phytoliths in vascular plants: an evolutionary perspective. Functional Ecology 30, 1286-1297.

Tauber, A.A., 1994. Estratigrafía y vertebrados fósiles de la Formación Santa Cruz (Mioceno inferior) en la costa atlántica entre las rías del Coyle y Río Gallegos, Provincia de Santa Cruz, República Argentina, República Argentina [Ph. D. dissertation]: Facultad de Ciencias Exactas, Físicas y Naturales, Córdoba, Universidad Nacional de Córdoba.

Tauber, A.A., 1997. Bioestratigrafía de la Formación Santa Cruz (Mioceno Inferior) en el extremo sudeste de la Patagonia. Ameghiniana 34, 413-426

Tauber, A.A., 1999. Los vertebrados de la Formación Santa Cruz (Mioceno inferior medio) en el extremo sureste de la Patagonia y su significado paleoecológico. Revista Española de Paleontología 14, 173-182.

Tejedor, M.F., Tauber, A.A., Rosenberger, A.L., Swisher, C.C., 3rd, Palacios, M.E., 2006. New primate genus from the Miocene of Argentina. Proc Natl Acad Sci U S A 103, 5437-5441.

Townsend, K.B., Croft, D.A., 2008. Diets of notoungulates from the Santa Cruz Formation, Argentina: new evidence from enamel microwear. Journal of Vertebrate Paleontology 28, 217-230.

Trayler, R.B., Schmitz, M.D., Cuitiño, J.I., Kohn, M.J., Bargo, M.S., Kay, R.F., Strömberg, C.A.E., Vizcaíno, S.F., 2019. An improved approach to age-modeling in deep time: Implications for the Santa Cruz Formation, Argentina. GSA Bulletin. https://doi.org/10.1130/B35203.1

Trayler, R.B., Schmitz, M.D., Cuitiño, J.I., Kohn, M.J., In rewiev. An improved approach to age-depth modeling in deep time: implications for the santa cruz formation, argentina.

Tripati, A.K., Roberts, C.D., Eagle, R.A., 2009. Coupling of CO_2 and ice sheet stability over major climate transitions of the last 20 million years. science 326, 1394-1397.

Vizcaíno, S.F., Bargo, M.S., Kay, R.F., Milne, N., 2006. The Armadillos (Mammalia, Xenarthra, Dasypodidae) Of The Santa Cruz Formation (Early–Middle Miocene): An Approach To Their Paleobiology. Palaeogeography, Palaeoclimatology, Palaeoecology 237, 255-269.

Vizcaíno, S.F., Bargo, M.S., Kay, R.F., Fariña, R.A., Di Giacomo, M., Perry, J.M., Prevosti, F.J., Toledo, N., Cassini, G.H., Fernicola, J.C., 2010. A Baseline Paleoecological Study For The

Santa Cruz Formation (Late–Early Miocene) At The Atlantic Coast Of Patagonia, Argentina. Palaeogeography, Palaeoclimatology, Palaeoecology 292, 507-519.

Vizcaíno, S.F., Kay, R.F., Bargo, M.S., 2012. Early Miocene paleobiology in Patagonia: high-latitude paleocommunities of the Santa Cruz Formation. Cambridge University Press.

Vizcaíno, S., Bargo, M., Fernicola, J., 2013. Expediciones paleontológicas durante los Siglos XIX y XX A la Formación Santa Cruz (Mioceno Inferior, Patagonia) y destino de los fósiles. Salta), Actas, 231-246.

Warny, S., Askin, R.A., Hannah, M.J., Mohr, B.A., Raine, J.I., Harwood, D.M., Florindo, F., Team, S.S., 2009. Palynomorphs from a sediment core reveal a sudden remarkably warm Antarctica during the middle Miocene. Geology 37, 955-958.

You, Y., Huber, M., Müller, R., Poulsen, C., Ribbe, J., 2009. Simulation of the middle Miocene climate optimum. Geophysical Research Letters 36.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686-693.

Zachos, J.C., Dickens, G.R., Zeebe, R.E., 2008. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. Nature 451, 279-283.

Zucol, A., Brea, M., Bellosi, E., Madden, R., Carlini, A., Vucetich, M., Kay, R., 2010. Phytolith studies in Gran Barranca (central Patagonia, Argentina): the middle-late Eocene. The paleontology of Gran Barranca: evolution and environmental change through the Middle Cenozoic of Patagonia, 317-340.

Zucol, A.F., Krause, J.M., Brea, M., Raigemborn, M.S., Matheos, S.D., 2018. Emergence of Grassy Habitats during the Greenhouse–Icehouse Systems Transition in the Middle Eocene of Southern South America. Ameghiniana 55, 451-483.

Table 3.1. Table of phytolith counts of the Sant Cruz Formation assemblages.
Counts include all diagnostic and non-diagnostic phytolith classes as well as classes not used in the analysis (e.g. AQ), and FI-t values estimated from bootstrap analysis with associated 95% confidence intervals.

Sample #	UWBM Locality #	Abbreviation from Fig. 3.2 & 3.4	AQ	ZINGI	FI	cf. CONI	PALM	GSSCP	NDG	NDO	Tree	Grass	*FI-t* estimate from bootstrap (%)	Lower 95% CI (%)	Upper 95% CI (%)
AR15-087		1	0	1	345	33	79	17	36	136	440	17	96.28	94.42	97.86
AR15-080		2	0	0	139	32	10	2	29	12	157	2	98.74	96.84	100.00
AR15-084		3	0	0	119	80	13	7	52	40	167	7	95.98	92.86	98.38
AR15-079		4	2	0	416	68	46	31	82	105	477	31	93.90	91.77	95.91
AR15-198		5	0	0	126	76	22	13	68	61	177	13	93.16	89.25	96.45
AR15-197		6	0	0	42	6	0	35	5	10	45	35	56.25	45.45	67.11
AR15-290		7	1	0	132	30	14	64	31	115	159	64	71.30	65.26	76.99
AR15-076		8	0	0	243	35	12	1	36	34	258	1	99.61	98.81	100.00
AR15-075		9	0	0	160	23	16	15	26	73	183	15	92.42	88.48	95.88
AR15-167		10	0	0	243	34	0	2	44	10	244	2	99.19	97.95	100.00
AR15-171		11	0	0	89	4	12	136	19	28	102	136	42.86	36.67	49.15
AR15-105		12	0	0	40	13	27	37	29	44	71	37	65.74	56.64	74.75
AR15-273		13	0	0	83	10	63	59	17	96	146	59	71.22	64.88	77.30
AR15-272		14	0	0	200	64	25	4	55	31	236	4	98.33	96.60	99.59
AR15-068		15	0	0	145	8	51	30	9	49	197	30	86.78	82.24	90.99
AR15-292		16	0	3	74	23	63	80	38	28	153	80	65.67	59.40	71.85
AR15-268		17	2	2	69	39	82	29	44	24	153	29	84.07	78.65	89.25
AR15-261		18	1	0	64	11	21	137	19	34	87	137	38.84	32.41	45.25
AR15-158		19	0	1	60	10	4	14	15	99	65	14	82.28	73.33	90.36
AR15-041		20	2	0	127	4	17	12	10	23	144	12	92.31	87.97	96.20
AR15-010		21	0	0	21	15	135	96	26	48	160	96	62.50	56.60	68.30
AR15-030		22	0	0	17	6	7	26	7	28	24	26	48.00	34.04	62.00
AR15-152		23	0	0	177	29	12	12	39	105	196	12	94.23	90.78	97.17
AR15-205		24	0	0	103	15	17	177	48	82	122	177	40.80	35.12	46.28
AR15-213		25	0	0	69	10	11	172	45	61	81	172	32.02	26.29	37.75
AR15-243		26	0	0	127	20	21	115	40	139	153	115	57.09	51.13	62.96
AR15-289		27	2	0	106	23	35	90	52	56	144	90	61.54	55.32	67.83
AR15-026		28	0	0	23	5	26	86	14	72	50	86	36.76	28.78	44.88
AR15-255		29	0	0	117	72	43	24	43	41	193	24	88.94	84.62	92.89

Table 3.2. Table of GSSCP counts the Sant Cruz Formation assemblages.

Counts include all GSSCP diagnostic and non-diagnostic classes, and CHO grass estimates from bootstrap analysis with associated 95% confidence intervals.

Sample #	UWBM Locality #	Abbreviation from Fig. 3.2 & 3.4	Closed-Habitat (CH)	POOID-D	POOID-ND	PANI	CHLOR	PACMAD GEN	OTHG	CH estimate by bootstrap (%)	lower 95% CI (%)	upper 95% CI (95)	Total # GSSCP counted
AR15-084	3		8	82	11	6	11	1	4	6.72	2.52	11.67	123
AR15-197	6		1	56	3	4	1	9	3	1.35	0.00	4.17	77
AR15-290	7		4	49	1	0	0	11	2	6.15	1.49	12.50	67
AR15-171	11		13	255	16	7	0	0	12	4.47	2.35	6.92	303
AR15-273	13		44	113	17	4	7	41	10	19.47	14.47	24.78	236
AR15-068	15		23	11	1	52	14	43	2	15.97	10.34	22.22	146
AR15-292	16		6	170	18	6	11	11	0	2.70	0.90	4.95	222
AR15-268	17		19	200	15	4	15	12	10	7.17	4.15	10.45	275
AR15-261	18		26	194	7	6	9	15	19	10.12	6.64	14.00	276
AR15-010	21		33	130	25	4	1	0	10	17.10	11.92	22.40	203
AR15-030	22		6	11	1	0	1	7	0	23.08	7.69	42.31	26
AR15-205	24		30	184	26	16	19	10	13	10.53	7.07	14.19	298
AR15-213	25		51	157	21	12	4	35	15	18.21	13.81	22.86	295
AR15-243	26		62	118	11	2	3	58	11	24.41	19.20	29.80	265
AR15-026	28		24	141	30	4	29	8	6	10.17	6.44	14.29	242

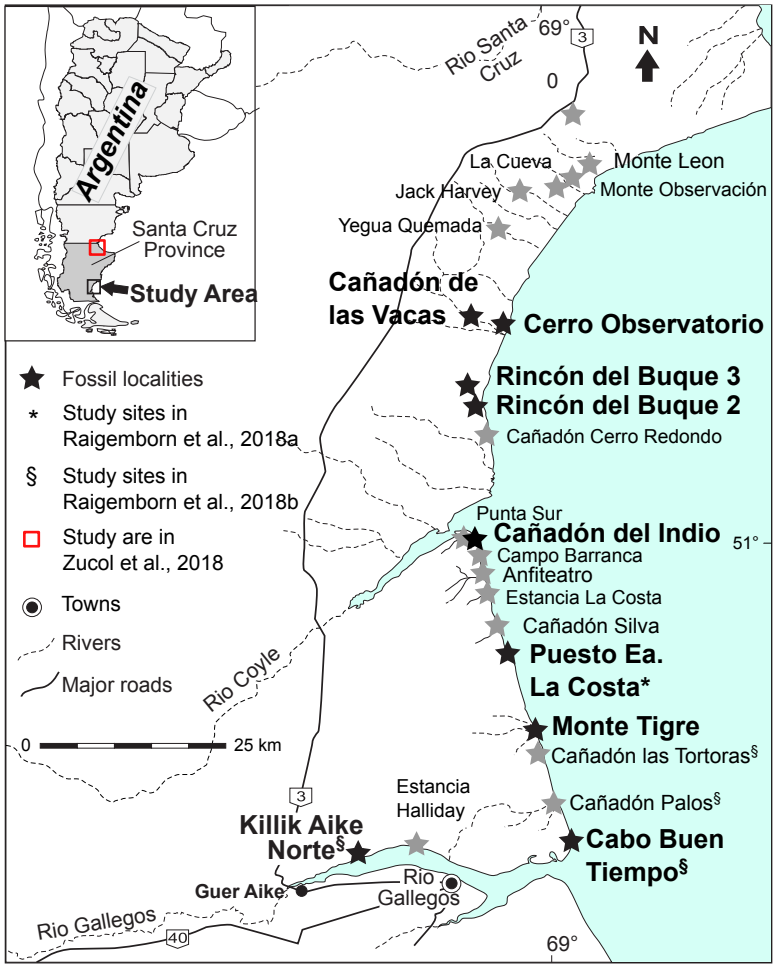


Figure 3.1. Map of the Santa Cruz Formation fossil localities

Gray stars represent known fossil localities. Bold stars represent localities included in this study, other symbols represent localities and study areas in previous phytolith studies (modified from Vizcaíno et al., 2012).

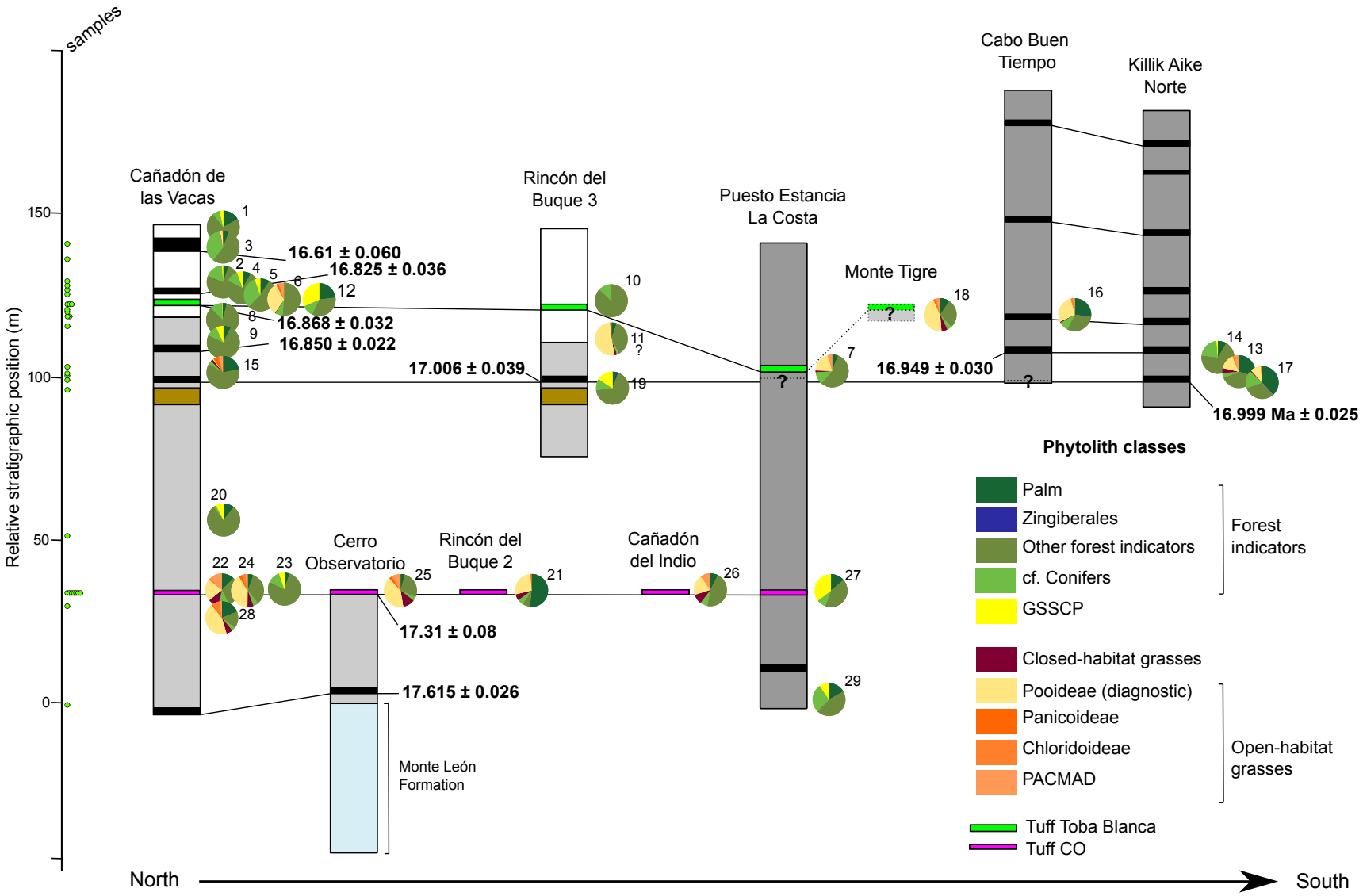


Figure 3.2. Composite stratigraphic section of SCF and phytolith assemblages.

Correlations between localities (black lines), and ages of dated tuffs (in bold) are based on Trayler et al., (2019). The geographic location of each locality (arranged from north to south in this diagram) is shown in Figure 3.1. Green dots along the stratigraphic column represent the stratigraphic position of samples. Pie charts represent relative proportion of different phytolith morphotypes in the assemblages. The two tuff layers from which lateral samples were collected are colored in pink (tuff CO), and green (tuff Toba Blanca).

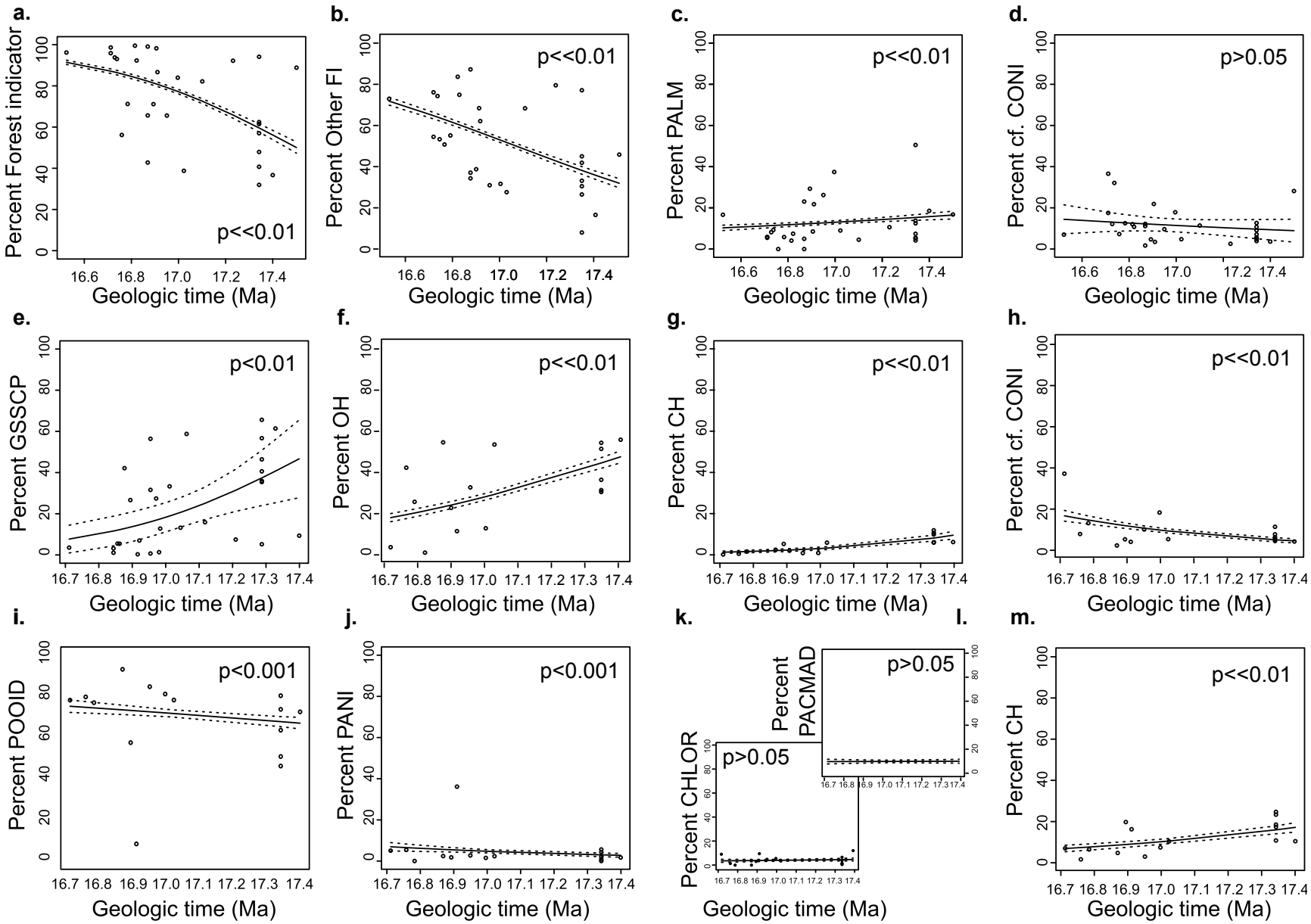

a.
Percent Forest indicator
p<<0.01
Geologic time (Ma)
b.
Percent Other FI
p<<0.01
Geologic time (Ma)
c.
Percent PALM
p<<0.01
Geologic time (Ma)
d.
Percent cf. CONI
p>0.05
Geologic time (Ma)
e.
Percent GSSCP
p<0.01
Geologic time (Ma)
f.
Percent OH
p<<0.01
Geologic time (Ma)
g.
Percent CH
p<<0.01
Geologic time (Ma)
h.
Percent cf. CONI
p<<0.01
Geologic time (Ma)
i.
Percent POOID
p<0.001
Geologic time (Ma)
j.
Percent PANI
p<0.001
Geologic time (Ma)
k.
Percent CHLOR
p>0.05
l.
Percent PACMAD
p>0.05
Geologic time (Ma)
m.
Percent CH
p<<0.01
Geologic time (Ma)

Figure 3.3. Relative abundance of phytoliths corresponding to various functional groups through time

a. relative abundance of FI calculated over the sum of forest indicator (FI tot) and open habitat indicator (GSSCP) functional groups. **b.-e.** relative abundances calculated over the sum of Other FI (forest indicators excluding palms, and conifers), PALM, cf. CONI, and GSSCP functional groups. **f.-h.** relative abundances calculated over the sum of Other FI (forest indicators excluding palms, and conifers), PALM, cf. CONI, CH, and OH functional groups. **i.-m.** relative abundances calculated over the sum of all GSSCP classes (CH, POOID, PANI, CHLOR, and PACMAD). Dotted lines represent 95% confidence intervals calculated around the fit of the generalized linear models.

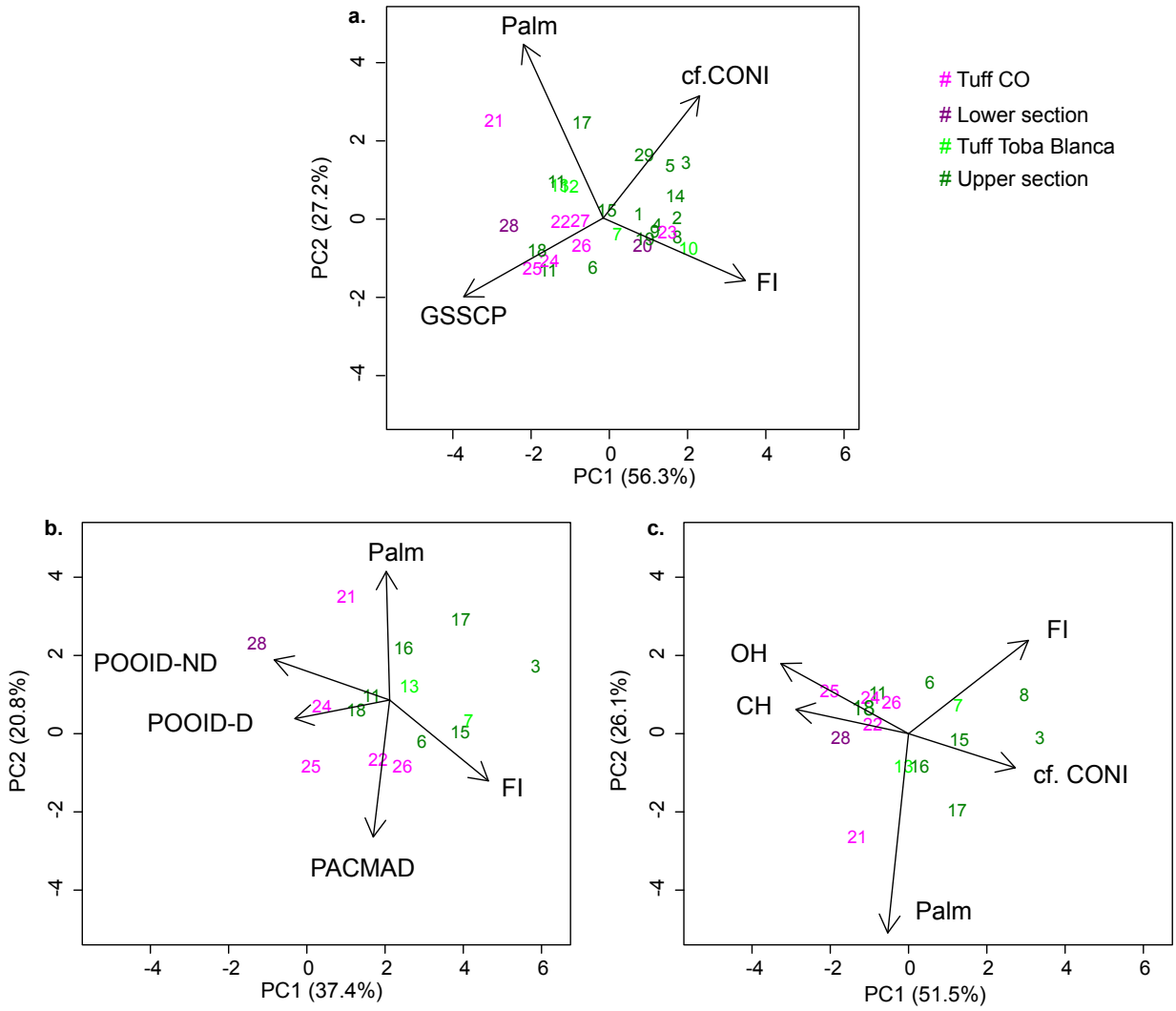


Figure 3.4. PCA ordination of the phytolith assemblages.

a. PCA1 with the variables FI (forest indicator excluding palms, conifers, and cf. conifers), PALM, cf. CONI (phytoliths produced by conifers, as well as Blo-1), and GSSCP. **b.** PCA2 with the variables FI (forest indicator excluding palms, conifers, and cf. conifers), PALM, cf. CONI (phytoliths produced by conifers, as well as Blo-1), CH, POOID-D, POOID-ND, PANI, CHLOR, and PACAMD. **c.** PCA3 with the variables FI (forest indicator excluding palms, conifers, and cf. conifers), PALM, cf. CONI (phytoliths produced by conifers, as well as Blo-1), CH, and OH.

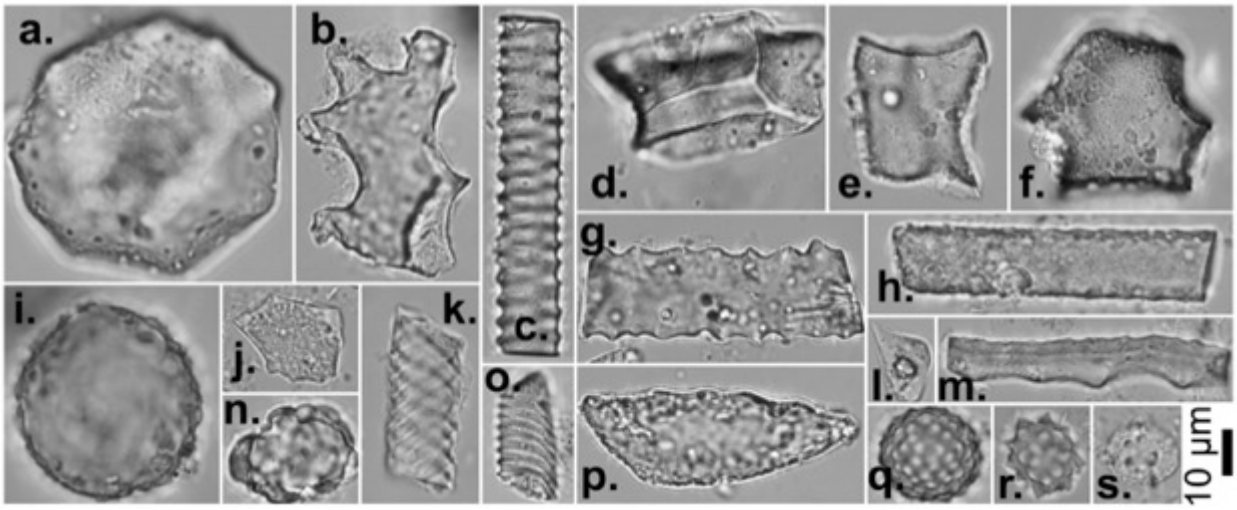


Figure 3.5. Selected non-GSSC phytolith form the Santa Cruz Formation

a. Blo-11, large, faceted/scalloped sphere produced by dicots (Other FI); **b.** Epi-2, anticlinal epidermis produced by dicots (Other FI); **c.** Tra-1, hollow and infilled helix (helical tracheary element), (Other FI); **d-f.** Blo-3 type B, faceted rectangular plate (CONI), in oblique (**d.**), top (**e.**) and end (**f.**) view); **g.** Epi-8, wavy elongate (NDG); **h**. Elo-1, smooth elongate (OTH); **i.** Large, verrucate sphere, produced by dicots (Other FI); **j.** Epi-1, polyhedral epidermis, produced by dicots (Other FI); **k.** Tra-1, hollow and infilled helix (Other FI); **l.** Tri-11, trichome with spindle-shaped or teardrop-shaped infilling (NDG); **m.** Elof-2, smooth elongate with faceted edge (OTH); **n.** Cl-8, large nodular body, produced by woody dicots (Other FI); **o.** Tra-1, hollow and infilled helix (helical tracheary element), (Other FI); **p.** Spindle-shaped body (trichome filling), (NDG); **q.** Clm-2, echinate sphere (PALM); **r.** Clm-2, echinate sphere (PALM); **s.** Clm-1, *Nypa*-type body; bowler hat with tubercles (PALM).

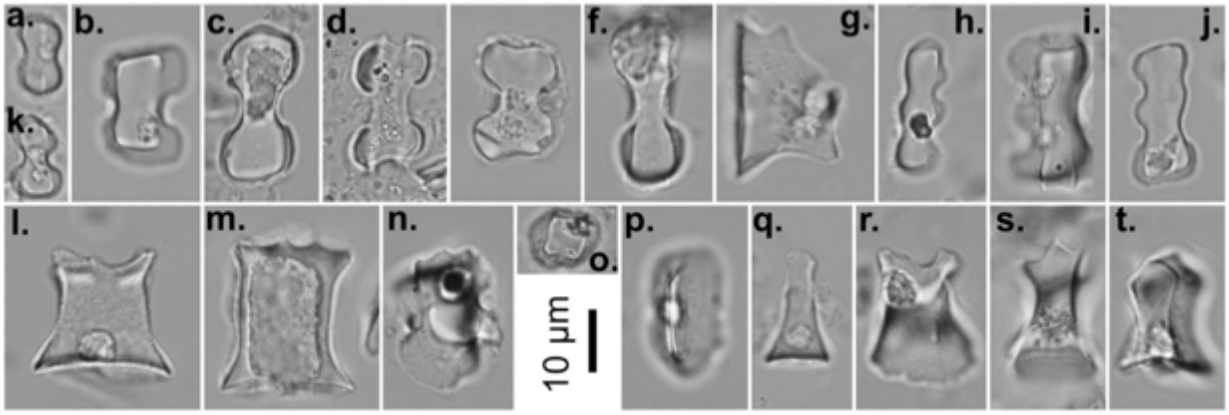


Figure 3.6. Selected GSSCP form the Santa Cruz formation

a., k. BI-1, regular *Stipa*-type bilobate (symmetry A) (POOID-D); **b.** BI-3, bilobate with rectangular top (POOID-D); **c.** BI-5, symmetry B bilobate (PACMAD); **d.** BI-7, symmetry D bilobate (PACMAD);

e. BI-8, symmetry E bilobate (PANI); **f.** BI-10, *Lygeum*-type wreath bilobate (POOID-D); **g.** CE-6, keel-crenate (OTHG); **h.** CE-1, crenate with symmetry A (POOID-D); **i.** CE-2, crenate with symmetry B (POOID-D); **j.** PO-1, polylobate with symmetry A (POOID-D); **l.** BIF-10, tall BI-1 type bilobate with ridges on top parallel to long axis (BOP); **m.** BIF-11type A, tall BI-1 type bilobate with ridges on top perpendicular to long axis (BOP); **n.** CEF-8, tall crenate with ornamented top (BOP); **o.** CO-1, generic (truncated conical rondel) (POOID-D); **p-q.** KR-1, regular keeled rondel (POOID-D); **r.** COF-6 A, tall conical rondel with concave plate (BOP); **s.** COF-1 B, tall rondel with spiked top (BOP); **t.** COF-4 A, hourglass shaped rondel (BOP).

Appendices

APPENDIX 1. SUPPLEMENTARY MATERIAL FOR CHAPTER 1

A.1.1. Data table with phytolith counts (functional groups) for all samples from La Selva and Palo Verde. The collection number is the number deposited at the Burke Museum

									Phytolith counts							
Collection number	Field Number	Site	Latitude	Longitude	Quadrat	Horizon	Sample type		Palms	Zingiberales	Palm/Zingiberales	Other forest indicators	Woody dicots	GSSCPs	Grass/monocots	Other
PR 9709	CRO16-001	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_1	Single		206	0	n.a.	48	21	12	18	43
PR 9710	CRO16-002	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_1	Single		171	0	n.a.	57	21	17	16	37
PR 9711	CRO16-003	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_2	Single		338	3	n.a.	53	15	3	17	58
PR 9712	CRO16-004	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_2	Single		338	3	n.a.	57	20	1	18	53
PR 9713	CRO16-005	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_3	Single		135	3	n.a.	51	9	7	25	33
PR 9714	CRO16-006	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_3	Single		179	0	n.a.	46	25	5	9	32
PR 9715	CRO16-007	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_4	Single		112	0	n.a.	52	33	8	7	35
PR 9716	CRO16-008	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_4	Single		166	0	n.a.	54	18	13	11	27
PR 9717	CRO16-009	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_5	Single		166	0	n.a.	65	14	8	27	25
PR 9718	CRO16-010	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_5	Single		124	1	n.a.	91	38	6	44	76

PR 9719	CRO16-011	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_6	Single	108	3	n.a.	82	17	2	18	30
PR 9720	CRO16-012	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_6	Single	99	2	n.a.	150	15	9	14	49
PR 9721	CRO16-013	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_7	Single	173	1	n.a.	75	14	12	10	13
PR 9722	CRO16-014	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_7	Single	172	0	n.a.	116	13	13	13	19
PR 9723	CRO16-015	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_8	Single	245	2	n.a.	60	15	9	17	23
PR 9724	CRO16-016	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_8	Single	216	0	n.a.	105	16	11	14	17
PR 9725	CRO16-017	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_9	Single	148	0	n.a.	59	5	3	12	17
PR 9726	CRO16-018	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_9	Single	148	1	n.a.	47	10	4	10	17
PR 9727	CRO16-019	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_{10}	Single	86	1	n.a.	47	4	2	11	17
PR 9728	CRO16-020	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_{10}	Single	170	0	n.a.	47	4	2	8	17
PR 9729	CRO16-021	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	U_{11}	Composite	161	1	n.a.	48	13	9	14	18
PR 9730	CRO16-022	La Selva	N 10° 24' 56.9"	W 084° 01' 32.4"	1	L_{11}	Composite	209	0	n.a.	62	19	11	19	27
PR 9731	CRO16-075	Palo Verde	N 10° 20' 55.4"	W 085° 20' 49.5"	2	U_{Q2}	Single	1	0	0	31	n.a.	181	12	7
PR 9732	CRO16-076	Palo Verde	N 10° 20' 55.4"	W 085° 20' 49.5"	2	L_{Q2}	Single	1	0	0	63	n.a.	270	11	15
PR 9733	CRO16-077	Palo Verde	N 10° 20' 55.4"	W 085° 20' 49.5"	2	U_{mixQ2}	Composite	1	0	0	44	n.a.	200	9	8
PR 9734	CRO16-078	Palo Verde	N 10° 20' 55.4"	W 085° 20' 49.5"	2	L_{mixQ2}	Composite	1	0	0	41	n.a.	225	14	8
PR 9735	CRO16-079	Palo Verde	N 10° 20' 52.7"	W 085° 20' 53.8"	3	U_{Q3}	Single	22	0	22	62	n.a.	180	22	27
PR 9736	CRO16-080	Palo Verde	N 10° 20' 52.7"	W 085° 20' 53.8"	3	L_{Q3}	Single	22	0	3	79	n.a.	166	32	21
PR 9737	CRO16-081	Palo Verde	N 10° 20' 52.7"	W 085° 20' 53.8"	3	U_{mixQ3}	Composite	14	0	15	82	n.a.	148	33	20
PR 9738	CRO16-082	Palo Verde	N 10° 20' 52.7"	W 085° 20' 53.8"	3	L_{mixQ3}	Composite	8	0	5	35	n.a.	133	27	13
PR 9739	CRO16-083	Palo Verde	N 10° 20' 49.4"	W 085° 20' 57.4"	4	U_{Q4}	Single	5	0	0	94	n.a.	126	40	37
PR 9740	CRO16-084	Palo Verde	N 10° 20' 49.4"	W 085° 20' 57.4"	4	L_{Q4}	Single	6	0	3	79	n.a.	120	34	46
PR 9741	CRO16-085	Palo Verde	N 10° 20' 49.4"	W 085° 20' 57.4"	4	U_{mixQ4}	Composite	4	0	1	89	n.a.	124	27	36
PR 9742	CRO16-086	Palo Verde	N 10° 20' 49.4"	W 085° 20' 57.4"	4	L_{mixQ4}	Composite	3	0	2	92	n.a.	108	46	59
PR 9743	CRO16-109	Palo Verde	N 10° 20' 48.4"	W 085° 20' 57.0"	5	U_{Q5}	Single	16	0	14	139	n.a.	85	15	15
PR 9744	CRO16-110	Palo Verde	N 10° 20' 48.4"	W 085° 20' 57.0"	5	L_{Q5}	Single	7	1	17	214	n.a.	44	38	37
PR 9745	CRO16-111	Palo Verde	N 10° 20' 48.4"	W 085° 20' 57.0"	5	U_{mixQ5}	Composite	8	4	23	295	n.a.	45	27	26

PR 9746	CRO16-112	Palo Verde	N 10° 20' 48.4"	W 085° 20' 57.0"	5	L_{mixQ6}	Composite	3	4	17	174	n.a.	42	40	25
PR 9747	CRO16-113	Palo Verde	N 10° 20' 51.8"	W 085° 20' 58.5"	6	U_{Q6}	Single	1	1	4	146	n.a.	97	35	43
PR 9748	CRO16-114	Palo Verde	N 10° 20' 51.8"	W 085° 20' 58.5"	6	L_{Q6}	Single	2	0	6	140	n.a.	103	20	35
PR 9749	CRO16-115	Palo Verde	N 10° 20' 51.8"	W 085° 20' 58.5"	6	U_{mixQ6}	Composite	1	2	21	146	n.a.	112	40	38
PR 9750	CRO16-116	Palo Verde	N 10° 20' 51.8"	W 085° 20' 58.5"	6	L_{mixQ6}	Composite	5	7	12	139	n.a.	101	40	45
PR 9751	CRO16-117	Palo Verde	N 10° 20' 55.1"	W 085° 19' 51.0"	7	U_{Q7}	Single	8	1	4	106	n.a.	94	33	46
PR 9752	CRO16-118	Palo Verde	N 10° 20' 55.1"	W 085° 19' 51.0"	7	L_{Q7}	Single	5	4	19	128	n.a.	95	55	64
PR 9753	CRO16-119	Palo Verde	N 10° 20' 55.1"	W 085° 19' 51.0"	7	U_{mixQ7}	Composite	6	0	14	150	n.a.	86	63	55
PR 9754	CRO16-120	Palo Verde	N 10° 20' 55.1"	W 085° 19' 51.0"	7	L_{mixQ7}	Composite	2	2	12	135	n.a.	107	31	33
PR 9755	CRO16-121	Palo Verde	N 10° 20' 53.2"	W 085° 19' 53.2"	8	U_{Q8}	Single	23	4	9	97	n.a.	101	58	38
PR 9756	CRO16-122	Palo Verde	N 10° 20' 53.2"	W 085° 19' 53.2"	8	L_{Q8}	Single	28	1	34	85	n.a.	88	46	34
PR 9757	CRO16-123	Palo Verde	N 10° 20' 53.2"	W 085° 19' 53.2"	8	U_{mixQ8}	Composite	16	0	18	128	n.a.	88	91	71
PR 9758	CRO16-124	Palo Verde	N 10° 20' 53.2"	W 085° 19' 53.2"	8	L_{mixQ8}	Composite	11	0	20	129	n.a.	76	92	47
PR 9759	CRO16-125	Palo Verde	N 10° 20' 51.4"	W 085° 19' 58.3"	9	U_{Q9}	Single	2	7	7	179	n.a.	43	44	36
PR 9760	CRO16-126	Palo Verde	N 10° 20' 51.4"	W 085° 19' 58.3"	9	L_{Q9}	Single	13	7	14	193	n.a.	53	32	35
PR 9761	CRO16-127	Palo Verde	N 10° 20' 51.4"	W 085° 19' 58.3"	9	U_{mixQ9}	Composite	7	9	23	192	n.a.	10	34	32
PR 9762	CRO16-128	Palo Verde	N 10° 20' 51.4"	W 085° 19' 58.3"	9	L_{miQx9}	Composite	17	8	11	189	n.a.	14	21	39
PR 9763	CRO16-129	Palo Verde	N 10° 20' 38.9"	W 085° 20' 19.8"	10	U_{Q10}	Single	5	0	4	127	n.a.	79	34	14
PR 9764	CRO16-130	Palo Verde	N 10° 20' 38.9"	W 085° 20' 19.8"	10	L_{Q10}	Single	3	2	0	165	n.a.	31	18	26
PR 9765	CRO16-131	Palo Verde	N 10° 20' 38.9"	W 085° 20' 19.8"	10	U_{mixQ10}	Composite	2	11	6	148	n.a.	34	20	35
PR 9766	CRO16-132	Palo Verde	N 10° 20' 38.9"	W 085° 20' 19.8"	10	L_{mixQ10}	Composite	3	15	6	112	n.a.	84	18	12
PR 9767	CRO16-133	Palo Verde	N 10° 20' 38.8"	W 085° 20' 13.8"	11	U_{Q11}	Single	2	0	11	154	n.a.	107	46	39
PR 9768	CRO16-134	Palo Verde	N 10° 20' 38.8"	W 085° 20' 13.8"	11	L_{Q11}	Single	4	0	5	128	n.a.	110	34	33
PR 9769	CRO16-135	Palo Verde	N 10° 20' 38.8"	W 085° 20' 13.8"	11	U_{mixQ11}	Composite	19	4	11	139	n.a.	139	22	25
PR 9770	CRO16-136	Palo Verde	N 10° 20' 38.8"	W 085° 20' 13.8"	11	L_{mixQ11}	Composite	4	0	17	126	n.a.	113	37	27

A.1.2. Data table with *FI-t* values and associated 95% confidence intervals of phytolith assemblages from La Selva primary quadrat 1).

Profile (Quadrat 1)	Sample	FI-t (%)	Lower CI (95%)	Upper CI (95%)	Differences in % FI-t (percentage points) U-L	U-U_{mix}	L-L_{mix}
1	U_1	95.83	93.36	97.96	2.22	0.29	2.74
	L_1	93.61	90.64	96.23			
2	U_2	99.27	98.3	100	0.49	3.15	3.42
	L_2	99.76	99.08	100			
3	U_3	96.59	93.81	99	1.45	0.24	1.69
	L_3	98.04	96.21	99.6			
4	U_4	96.1	93.36	98.54	1.28	0.02	1.52
	L_4	94.82	91.89	97.57			
5	U_5	96.84	94.49	98.81	0.85	0.72	1.35
	L_5	97.69	95.69	99.61			
6	U_6	99.06	97.62	100	2.33	2.94	0.38
	L_6	96.73	94.58	98.6			
7	U_7	95.64	93.04	97.83	0.22	0.48	0.49
	L_7	95.86	93.57	97.83			
8	U_8	97.28	95.14	98.81	0.44	1.16	0.49
	L_8	96.84	94.87	98.55			
9	U_9	98.6	96.8	100	0.51	2.48	1.75
	L_9	98.1	96.05	99.53			
10	U_{10}	98.57	96.35	100	0.53	2.45	2.76
	L_{10}	99.1	97.73	100			
MIX	U_{mix}	96.12	93.67	98.31	0.22	n.a.	n.a.
	L_{mix}	96.35	93.96	98.06			
	AVERAGE	97.13±1.57	95.01±2.12	98.86±1.03	0.96±0.76	1.39±1.23	1.66±1.05

A.1.3. Data table with *FI-t* values and associated 95% confidence intervals of phytolith assemblages from Palo Verde quadrats.

					Differences in % FI-t (percentage points)			
Quadrat	**Sample**	**FI-t (%)**	**Lower CI (95%)**	**Upper CI (95%)**	U-L	$U-U_{mix}$	$L-L_{mix}$	$U_{mix}-L_{mix}$
15	U_{Q2}	15.02	9.91	19.63	4.14	3.34	3.43	2.64
	L_{Q2}	19.16	14.95	23.57				
	U_{mixQ2}	18.37	13.77	23.27				
	L_{mixQ2}	15.73	11.41	20.08				
16	U_{Q3}	37.06	31.14	42.60	1.46	5.79	**12.00**	***16.34***
	L_{Q3}	38.52	32.57	44.16				
	U_{mixQ3}	42.86	36.82	48.67				
	L_{mixQ3}	26.52	20.00	33.34				
17	U_{Q4}	44.00	37.22	50.47	1.69	0.88	5.01	4.20
	L_{Q4}	42.31	35.88	48.78				
	U_{mixQ4}	43.12	36.62	49.54				
	L_{mixQ4}	47.32	40.50	54.32				
19	U_{Q5}	66.54	60.94	71.89	***17.92***	**21.46**	1.95	5.50
	L_{Q5}	84.45	80.43	88.70				
	U_{mixQ5}	88.00	84.51	91.15				
	L_{mixQ5}	82.50	77.64	87.08				
20	U_{Q6}	61.04	55.06	67.22	2.08	0.76	2.78	1.46
	L_{Q6}	58.96	52.59	65.08				
	U_{mixQ6}	60.28	54.71	65.82				
	L_{mixQ6}	61.74	55.90	67.88				
21	U_{Q7}	55.87	49.53	62.04	6.28	10.54	3.62	7.88
	L_{Q7}	62.15	56.25	67.89				
	U_{mixQ7}	66.41	60.31	71.75				
	L_{mixQ7}	58.53	52.27	64.82				
22	U_{Q8}	56.84	50.42	63.14	5.87	7.96	5.08	3.00
	L_{Q8}	62.71	56.59	68.78				
	U_{mixQ8}	64.80	58.94	70.59				
	L_{mixQ8}	67.80	61.95	73.67				
23	U_{Q9}	81.93	77.21	86.67	0.86	**13.92**	**13.07**	1.71
	L_{Q9}	81.07	76.70	85.61				
	U_{mixQ9}	95.85	93.22	97.98				
	L_{miQx9}	94.14	91.13	96.75				

24	U_{Q10}	63.26	56.87	69.49	***21.32***	**19.83**	**22.76**	***21.27***
	L_{Q10}	84.58	79.50	89.34				
	U_{mixQ10}	83.08	77.78	88.24				
	L_{mixQ10}	61.82	55.91	68.06				
25	U_{Q11}	60.95	54.81	66.55	5.48	5.50	1.07	1.09
	L_{Q11}	55.47	49.39	61.57				
	U_{mixQ11}	55.45	49.84	60.51				
	L_{mixQ11}	56.54	50.75	62.50				
	AVERAGE	58.07±21.08	52.54±21.36	63.50±20.71	6.71±7.12	9±7.37	7.08±6.83	6.51±6.89

APPENDIX 2: SUPPLEMENTARY MATERIAL FOR CHAPTER 2

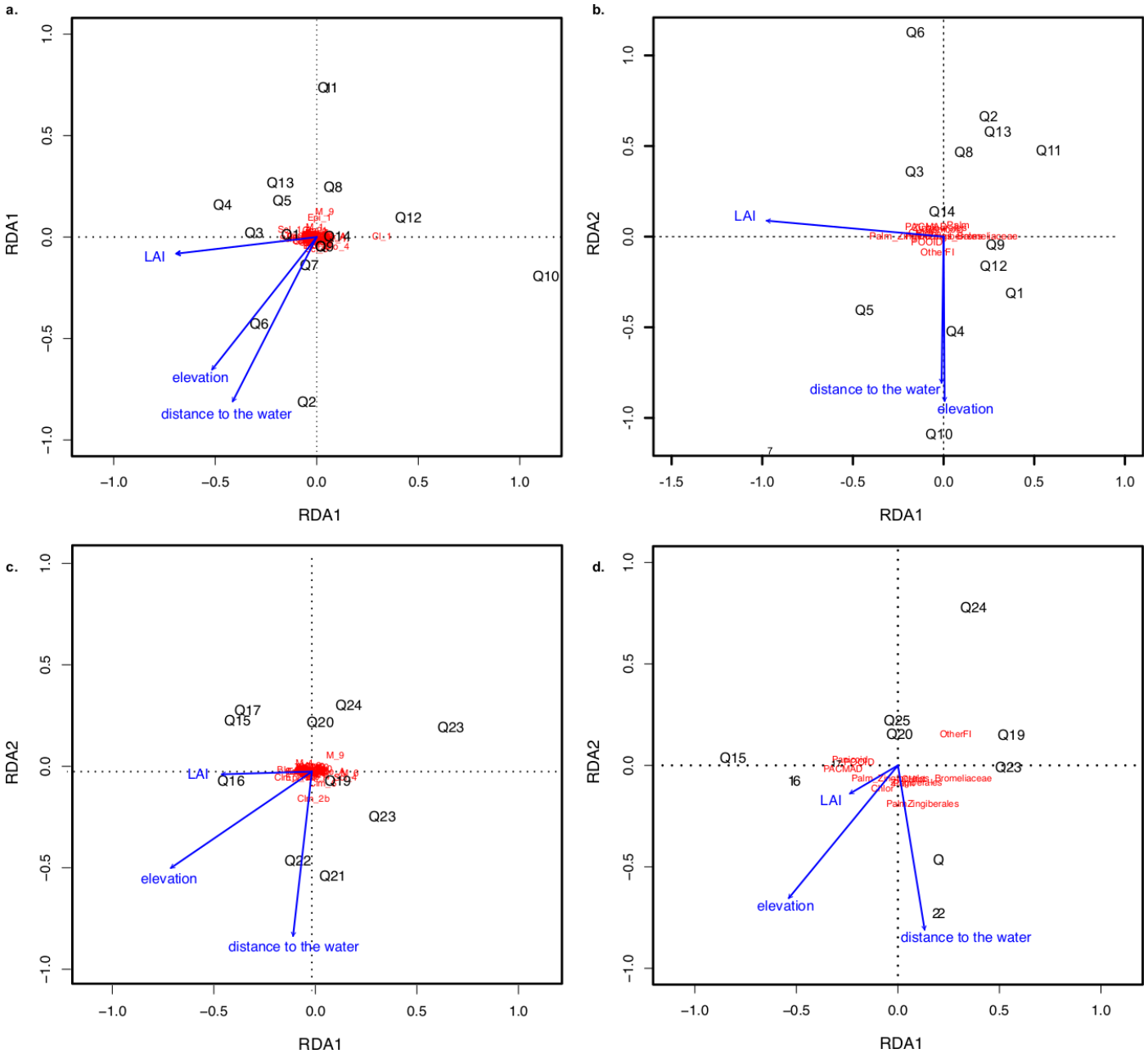
A.2.1. List of plant species from Palo Verde with estimated silica yield (rY_{Si}), number of phytoliths counted in 50 fields of view, number of diagnostic phytoliths in 50 fields of view (n_{phyto}), relative phytolith yield ($rYield_{phyto}$) and standard diagnostic phytolith yield (Y_{phyto}) for each species.

UWBM herbarium accession #	Family	Genus	Species	Species acronym	rY_{Si}	# of phytoliths in 50 fields of view*	n_{phyto}	rY_{phyto}	Y_{phyto}
415121	BIGNONIACEAE	*Tabebuia*	*ochracea*	TABOCH	1	88	88	1.76	1.76
421396	FABACEAE	*Machaerium*	*biovulatum*	MACBIO	0	0	0	0	0
421398	SALICACEAE	*Casearia*	*corymbosa*	CASCOR	1	25	24	0.48	0.48
421399	ACANTHACEAE	*Aphelandra*	*scabra*	APHSCA	1	5	5	0.1	0.1
421400	ERYTHROXYLACEAE	*Erythroxylum*	*havanense*	ERYHAV	1	24	23	0.46	0.46
421401	SAPINDACEAE	*Alophyllus*	*racemosus*	ALORAC	3	138	135	2.7	8.1
421402	ANACARDIACEAE	*Astronium*	*graveolens*	ASTGRA	1	91	91	1.82	1.82
421404	ZYGOPHYLLACEAE	*Guaiacum*	*sanctum*	GUASAN	1	32	31	0.62	0.62
421405	FABACEAE	*Acacia*	*centralis*	ACACEN	2	99	99	1.98	3.96
421406	FABACEAE	*Lonchocarpus*	*fasciculatus*	LONCFAS	4	107	103	2.06	8.24
NA	MALVACEAE	*Indet.*		INDET2	4	108	108	2.16	8.64
421407	MALVACEAE	*Luehea*	*candida*	LUECAN	1	113	112	2.24	2.24
421408	SIMAROUBACEAE	*Simarouba*	*glauca*	SIMGLA	2	250	214	4.28	8.56
421410	SAPOTACEAE	*Manilkara*	*chicle*	MANCHI	1	15	15	0.3	0.3
421411	BORAGINACEAE	*Cordia*	*dentata*	CORDEN	5	53	50	1	5
421412	BORAGINACEAE	*Cordia*	*callococca*	CORCAL	4	49	37	0.74	2.96
NA	MALVACEAE	*Guazuma*	*ulmifolia*	GUAULM	3	75	74	1.48	4.44
421413	ANACARDIACEAE	*Spondias*	*mombin*	SPOMOM	4	98	97	1.94	7.76
421414	FABACEAE	*Acacia*	*collinsii*	ACACOL	1	5	5	0.1	0.1
421415	CAPPARIDACEAE	*Capparidastrum*	*frondosum*	CAPFRON	2	75	74	1.48	2.96
421416	RUBIACEAE	*Psychotria*	*carthagenensis*	PSYCAR	2	35	35	0.7	1.4
421417	SAPINDACEAE	*Thouinidium†*	*decandrum*	THODEC	1	25	24	0.48	0.48
421420	POLYGONACEAE	*Coccoloba*	*caracasana*	COCCAR	0	0	0	0	0
421419	POLYGONACEAE	*Triplaris*	*melaenodendron*	TRIMEL	1	4	3	0.06	0.06
421421	CONNARACEAE	*Rourea*	*glabra*	ROUGLA	3	47	47	0.94	2.82
421422	PRIMULACEAE	*Ardisia*	*revoluta*	ARDREV	1	71	53	1.06	1.06
421423	RUBIACEAE	*Calycophyllum*	*candidissimum*	CALCAN	3	42	42	0.84	2.52
421424	BIGNONIACEAE	*Tabebuia*	*rosea*	TABROS	3	45	43	0.86	2.58

421425	PHYTOLACACEAE	*Petiveria*	*alliacea†*	PETALL	0	0	0	0	0
421426	MORACEAE	*Brosimum*	*alicastrum*	BROALI	5	40	39	0.78	3.9
421427	EUPHORBIACEAE	*Garcia*	*nutans*	GARNUT	3	113	113	2.26	6.78
421428	FABACEAE	*Lonchocarpus*	*parviflorus*	LONPAR	3	71	71	1.42	4.26
421432	BIXACEAE	*Cochlospermum*	*vitifolium*	COCVIT	2	63	63	1.26	2.52
421433	APOCYNACEAE	*Stemmadenia*	*pubescens*	STEPUB	0	0	0	0	0
421434	FABACEAE	*Pithecellobium*	*unguis-cati*	PITUNG	5	89	88	1.76	8.8
421435	EUPHORBIACEAE	*Croton*	*niveus*	CRONIV	3	139	138	2.76	8.28
421436	PRIMULACEAE	*Bonellia*	*nervosa*	BONNER	2	580	580	11.6	23.2
421438	RUBIACEAE	*Randia*	*armata*	RANARM	0	0	0	0	0
421439	MELIACEAE	*Trichilia*	*hirta*	TRIHIR	3	97	96	1.92	5.76
421440	RUBIACEAE	*Randia*	*aculeata*	RANACU	0	0	0	0	0
421441	FABACEAE	*Lonchocarpus*	*minimiflorus*	LONMIN	3	143	131	2.62	7.86
421443	RUBIACEAE	*Guettarda*	*macrocarpa*	GUEMAC	1	14	14	0.28	0.28
421444	MALVACEAE	*Pavonia*	*dasypetala*	PAVDAS	3	53	52	1.04	3.12
421445	ACANTHACEAE	*Carlowrightia*	*arizonica*	CARARI	1	15	15	0.3	0.3
421446	CAPPARACEAE	*Quadrella*	*indica*	QUAIND	0	0	0	0	0
421447	ARECACEAE	*Bactris*	*guineensis*	BACGUI	5	227	111	2.22	11.1
421449	MALVACEAE	*Melochia*	*lupulina*	MELUP	1	50	50	1	1
421450	FABACEAE	*Lonchocarpus*	*costaricensis*	LONCOS	5	140	135	2.7	13.5
421451	ACANTHACEAE	*Barleria*	*oenotheroides*	BAROEN	1	102	99	1.98	1.98

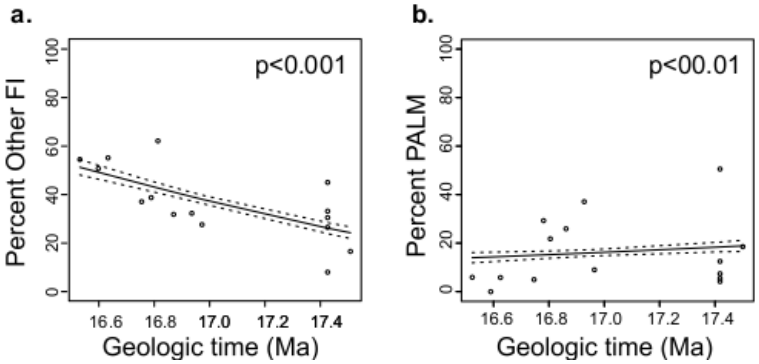
*including diagnostic and non-diagnostic morphotypes

A.2.2. Redundancy Analysis (RDA) ordination diagrams showing the soil profiles of **a-b.** La Selva rainforest, **c-d.** Palo Verde dry forest. Q1-Q24 correspond to quadrat numbers. Acronym in red correspond to phytolith morphotypes (in **a.** and **c.**), and phytolith functional groups (in **b.** and **d.).** Blue vectors represent loading of environmental variables (LAI, distance to the water, and elevation). The direction of the vectors indicates maximum correlation between the scores of the samples (quadrats, and phytolith morphotypes/functional groups), and the environmental variables; however, p>0.05 for all RDA ordinations.



APPENDIX 3: SUPPLEMENTARY MATERIAL FOR CHAPTER 3

A.3.1. Phytolith abundance through time of: **a.** Other FI, calculated over the sum of FI (forest indicators excluding palms, and conifers), PALM, cf. CONI, CH, POOID, PANI, CHLOR, and PACMAD; and **b.** Palm phytolith abundance through time through time, calculated over the sum of FI (forest indicators excluding palms, and conifers), PALM, cf. CONI, CH, POOID, PANI, CHLOR, and PACMAD.



VITA

Camilla Crifò was born in Rome, Italy. She became interested in Paleoecology and Paleobotany during her undergraduate studies in Organismal Biology at the University of Montpellier, France. For the final project of her Masters in Paleobiology, at University of Montpellier, she spent four month conducting research at the Smithsonian Tropical Research Institute (STRI), under the supervision for. Dr. Carlos Jaramillo. After an additional year spent at STRI she did a second Master thesis under the supervision of Dr. Ellen Currano, at Miami University, Ohio. Her research at Panama and at Miami University resulted in a peer review article on angiosperm venation density and its use a as proxy for forest structure in the fossil record, in the journal *Geology*. Camilla began her Ph.D. studies in 2013 under the supervision of Dr. Caroline Strömberg at the University of Washington. During her time at UW, she held various teaching assistant positions in the Biology department. In addition, she spent a significant amount of time training and mentoring undergraduates in the lab. During Dr. Strömberg's sabbatical (2015/2016), Camilla worked as lab and project manager in the Strömberg lab, guiding a group of undergraduates through a research project, and ensuring the basic functioning of the laboratory. In addition, Camilla acquired field experience by participating to various fieldwork campaign in South America, including Argentina, Bolivia, Costa Rica, Peru, and Colombia. She participated to various outreach activities at the Burke Museum and collaborated to an outreach program sponsored by the Nasa space grant consortium at UW. In addition to her dissertation work, Camilla coauthored a peer-reviewed article on leaf epidermal morphology and its physiological implications, published on the journal *New Phytologist*, as well as a book chapter on the applications of phytolith analysis in paleoecology.