The legacy of 4,500 years of polyculture agroforestry in the eastern Amazon

S. Yoshi Maezumi, Daiana Alves, Mark Robinson, Jonas Gregorio de Souza, Carolina Levis, Robert L. Barnett, Edemar Almeida de Oliveira, Dunia Urrego, Denise Schaan and José Iriarte

The legacy of pre-Columbian land use in the Amazonian rainforest is one of the most controversial topics in the social1–10 and natural sciences11,12. Until now, the debate has been limited to discipline-specific studies, based purely on archaeological data1, modern vegetation2, modern ethnographic data3 or a limited integration of archaeological and palaeoecological data4. The lack of integrated studies to connect past land use with modern vegetation has left questions about the legacy of pre-Columbian land use on the modern vegetation composition in the Amazon, unanswered5. Here, we show that persistent anthropogenic landscapes for the past 4,500 years have had an enduring legacy on the hyperdominance of edible plants in modern forests in the eastern Amazon. We found an abrupt enrichment of edible plant species in fossil lake and terrestrial records associated with pre-Columbian occupation. Our results demonstrate that, through closed-canopy forest enrichment, limited clearing for crop cultivation and low-severity fire management, long-term food security was attained despite climate and social changes. Our results suggest that, in the eastern Amazon, the subsistence basis for the development of complex societies began ~4,500 years ago with the adoption of polyculture agroforestry, combining the cultivation of multiple annual crops with the progressive enrichment of edible forest species and the exploitation of aquatic resources. This subsistence strategy intensified with the later development of Amazonian dark earths, providing a food production system that sustained growing human populations in the eastern Amazon. Furthermore, these millennial-scale polyculture agroforestry systems have an enduring legacy on the hyperdominance of edible plants in modern forests in the eastern Amazon. Together, our data provide a long-term example of past anthropogenic land use that can inform management and conservation efforts in modern Amazonian ecosystems.

The extent to which pre-Columbian societies altered Amazonian landscapes is one of the most debated topics in botany6,13–15, archaeology6,1,2,3,7, palaeoecology7,12,16–19 and conservation20,21. New findings show that a disproportionate number of plants, accounting for half of all trees in the Amazon, are hyperdominant22 and domesticated species are five-times more likely to be hyperdominant than non-domesticated species13. This is particularly prevalent in archaeological sites, suggesting that the effect of pre-Columbian people on modern flora is more pronounced than previously thought1. The pre-Columbian anthropogenic soils, known as Amazonian dark earths (ADEs; traditionally called Terras Pretas de Índio), are one of the most distinct lines of evidence of human transformation of Amazonia because these modified soils are indicators of pre-Columbian sedentary occupation1,2,5. ADEs have been associated with sustained and intensive agriculture in the past and have been re-utilized by modern farmers because of their extremely high fertility22. Several studies have shown that (1) forests on ADEs have a distinct species composition, exhibiting greater richness and a higher abundance of domesticated and edible plants (used as food resources)19, (2) the more complex the ADE archaeological context (for example, multicomponent sites), the greater the floristic composition of cultivated useful plants in modern home gardens3, and (3) increased fertility associated with ADEs improves conditions for the establishment and growth of exotic species that are generally more nutrient demanding than native Amazonian species24.

However, the lack of detailed integrated archaeological or palaeoecological studies to connect past land use with modern vegetation have left fundamental questions about land-use practices and the effect of ADEs on modern Amazonian ecosystems unresolved. To address these issues, we integrate archaeological and archaeobotanical records that reflect local-scale vegetation histories with lake and terrestrial palaeoecology that reflects broader regional-scale vegetation histories, combined with palaeoclimate and modern botanical surveys to investigate the impact of the past 4,500 years of human land use in the eastern Amazon (Fig. 1).

The study area located within the protected rainforest of the FLONA Tapajós Reserve, provides an ideal setting because of the presence of extensive archaeological sites, high concentrations of ADE soils, a nearby lake with limited riverine influence and the existence of a nearby high-resolution palaeoclimate record26. We (1) collected a 210-cm sediment core dating to ~8,500 calendar years (cal) BP from Lake Caranã (~0.7 km in diameter, ~3 m in depth; 2° 50’ 08” S, 55° 02’ 33” W; 5 m above sea level) for palaeoecological analysis (see Methods section ‘Regional study area’ to ‘Pollen’); (2) carried out excavations and sampled three ADE soil profiles at the nearby Serra do Maguari-1 (SDM1) archaeological site (see Methods sections ‘Archaeological site selection’ and ‘Soil phytoliths’), ~5 km northeast of Lake Caranã on the crest of the upper slope of the Belterra Plateau (2° 47’ 87” S, 55° 03’ 53” W; 126 m above sea level) (Supplementary Fig. 4), compiled existing regional archaeological data (see Methods the ‘SPDs and site

---

1Department of Archaeology, College of Humanities, University of Exeter, Exeter, UK. 2Programa de Pós-Graduação em Ecologia, Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Brazil. 3Forest Ecology and Management Group, Wageningen University & Research, Wageningen, The Netherlands. 4Department of Geography, College of Life and Environmental Sciences, University of Exeter, Exeter, UK. 5Universidade do Estado de Mato Grosso, Campus of Nova Xavantina, Mato Grosso State, Brazil. 6Department of Anthropology, Federal University of Pará, Belém, Brazil. 7e-mail: s.y.maezumi@exeter.ac.uk
(frequencies’ section)); and (3) set up survey plots (three in ADE and three in non-ADE sites) to perform modern vegetation inventories (Supplementary Fig. 6; see Methods section ‘Modern botanical survey’). Owing to the extensive landscape-level modification of soils in this region, we define non-ADE plots as less-modified soils without ceramics, located at least 150 m from dark ADE soils. As the formation and utilization of ADEs are closely associated with food production, we focus our analysis on edible plant taxa in the pollen, phytolith and botanical assemblage. We classify edible plants as taxa that are ethnographically used as food resources in the Americas following Clement29, Levis et al.13 and Hanelt et al.30. These proxies are compared to the nearby speleothem record of the Paraíso Cave (Fig. 2c), which provides a high-resolution record of natural climate variability for the past ~45,000 years.

We used a combination of $^{210}$Pb and accelerator mass spectrometry (AMS) radiocarbon dating techniques to develop a robust chronology for the age-depth model for Lake Caranã. $^{210}$Pb dating was used to constrain the most recent palaeoenvironmental changes (<250 years), whereas AMS radiocarbon dating was used to date sediments that were >200 years (see Methods (the ‘Palaeoecological age-depth model’ section)). Based on the compiled charcoal, pollen and geochemical data, three main phases are identified at Lake Caranã that represent the past ~8,500 years (see Methods (from the ‘Palaeoecology site selection and core collection’ to the ‘Pollen’ sections); Figs. 2 and 3, Supplementary Figs. 1–3 and Supplementary Table 2).

**Phase 1.** Geochemical data (Methods (from the ‘Palaeoecological age-depth model’ to the ‘Magnetic susceptibility’ sections); Supplementary Fig. 2) characterize phase 1 (210–128 cm; before ~4,500 cal BP) as a high-energy environment with increased allochthonous inputs, probably associated with increased riverine influence (Supplementary Discussion 1). Pollen concentration is very low during this period (that is, <100 grains per cm$^2$) and is attributed to poor preservation associated with sandy soils and low organic material (Supplementary Figs. 1–3). Fire activity is low (Fig. 2d) and there is no palaeoecological evidence of human occupation in the near vicinity during this phase despite documented human activity in the region (Supplementary Discussion 2). Phase 1 is associated with the wettest period in the past ~45,000 years.

**Phase 2.** Phase 2 (128–79 cm; ~4,500–2,500 cal BP) begins with the formation of Lake Caranã following decreased riverine inputs, as indicated by the shift in geochemistry and an increase in sediment organic (Supplementary Fig. 2). Pollen is dominated by >50% rainforest taxa (Fig. 2c). Approximately 30% of the total pollen taxa and the frequency of herbs and grasses remain low throughout the length of the record (<10%). Our record documents the earliest arrival of maize (Zea mays) to the eastern Amazon, which is then present consistently after ~4,300 cal BP, combined with sweet potato (Ipomoea batatas), recorded at ~3,200 cal BP, the assemblage indicates that the lake inhabitants practiced polyculture (mixed cropping), including cereal and tubers crops (Fig. 2c and Supplementary Fig. 3). The occurrence of maize pollen is consistent with a temporal gradient of maize dispersal that begins outside Amazonia and reaches the eastern Amazon ~4,300 cal BP (Fig. 1a and Supplementary Table 1), ~3,800 years before the development of ADE soils at SDM1. The formation of Lake Caranã is followed by increased charcoal accumulation, indicating low-severity fire activity around the lake (Fig. 2d). Although regional climate data document a gradual shift towards drier conditions from ~4,500 cal BP, which continues into phase 3 (Fig. 2e), the synchronous onset of fire activity combined with the presence of maize pollen suggests intentional human-caused ignitions associated with local forest clearance for crop cultivation around the lake. The sum of the calibrated probability distributions (SPDs) from dated archaeological contexts in the lower Tapajós indicate an increase in regional-scale pre-Columbian activity after ~4,300 cal BP (see Methods (the ‘SPDs and site frequencies’ section); see bottom panel in Fig. 2b and Supplementary Table 3). During phase 2, the residents of Lake Caranã were probably hunting and fishing and utilizing the seasonally flooded, nutrient-rich soils surrounding the lake shore, practicing agroforestry and exploiting wild plants, combined with low-level fire activity to clear land for polyculture.

**Phase 3.** Phase 3 (79–0 cm; 2,500 cal BP to modern) is characterized by an increase in the number of edible plants (from ~45% to >70% of terrestrial pollen taxa), a decrease in the number of non-edible plants (from ~50% to 30%), followed by the arrival of manioc (Manihot esculenta) ~2,250 cal BP (Fig. 2c). The increase in the number of edible plants is not associated with significant change
in the regional climate data, suggesting that pre-Columbian forest enrichment is the driver of this abrupt change in forest composition. Lake Caranã exhibited an increase in fire activity ~1,250–500 cal BP, associated with the increase in pre-Columbian activity (see bottom panel in Fig. 2b), coupled with the arrival of squash (Cucurbita sp.) ~600 cal BP in the lake pollen.

Phase 3 is contemporaneous with the archaeobotanical data from SDM1, a ~15-ha mounded village with a central plaza surrounded by a mosaic of ADE sites (Supplementary Fig. 4). Mound construction and ADE formation at SDM1 occurred between ~530 and 450 cal BP (Supplementary Table 3), and ceramic material at the site is characteristic of the late pre-Columbian Tapajós Period (Supplementary Discussion 2). Herb phytoliths from the three profiles account for ~4–18% (seetoppanelinFig. 2band Supplementary Fig. 5). Phytolith data indicate a gradual increase in edible plants (see Methods (the ‘Soil phytoliths’ section); Supplementary Tables 9 and 10) following the formation of ADE soils that reach the highest levels (>66%) in modern surface soils. Squash is present in the phytolith assemblage at SDM1 from before the formation of the ADE, consistent with the presence of squash in the lake; however, maize does not appear at SDM1 until the ADE has formed. Soil charcoal is present in all three profiles at SDM1 and increases with the formation of ADE soils ~530 cal BP (see top panel in Fig. 2b). The increase in soil charcoal in the ADE layers suggests that in-field burning was probably implemented to ameliorate nutrient-poor ferralsols on Belterra Plateau to cultivate nutrient-demanding crops,
such as maize. Similar practices have been recorded in indigenous Amazonian groups\textsuperscript{31}. However, sediment charcoal from the lake indicates an overall decrease in fire activity in the watershed at this time, synchronous with the driest regional climate conditions in the past 5,500 years (Fig. 2). Modern fire activity in the eastern Amazon is associated with increased droughts\textsuperscript{32}; thus, this decrease in fire activity suggests that pre-Columbian fire management probably suppressed large wildfires during the apex of pre-Columbian activity in the region. This interpretation is further supported by the continued presence of rainforest vegetation at Lake Caranã despite climate and socioeconomic changes. The patterns observed in the Lake Caranã record are concomitant with key periods of cultural transformation in the lower Tapajós. The abrupt increase in fire activity after ~1,250 cal bp followed by the arrival of new cultivars (\textit{Cucurbita} sp.) ~600 cal bp coincide with the development of the Santarém culture, centred at the modern city of Santarém, where the largest site comprising ~16 ha of ADE is located\textsuperscript{5}. The Santarém polity purportedly comprised an area of 23,000 km\textsuperscript{2}, with sites extending for hundreds of miles along river bluffs and interior plateaus\textsuperscript{40}. During this period, the lower Tapajós concentrated one of the highest population densities in Amazonia\textsuperscript{41–43}. Initially interpreted as a warlike, tribute-based chiefdom that persisted until colonial times\textsuperscript{47}, the degree of centralization of the Santarém polity has been recently questioned given the absence of differential access to prestige goods or other clear evidences of hierarchy\textsuperscript{46}. Moreover, peripheral sites were shown to be independent of Santarém influence during the early stages of that cultural expansion, although they were ultimately abandoned after ~1,000 cal bp\textsuperscript{40}. Irrespective of the social organization implied, the changes observed in the Lake Caranã record after ~1,250 cal bp could reflect changing land

*Fig. 3 | A conceptual landscape drawing of the changing vegetation and disturbance regimes. The conceptual drawing was inferred from an analysis of pollen, phytoliths and charcoal from the Lake Caranã core and the SMD1 archaeological site associated with the three phases discussed in the main text.*
management practices aimed to increase subsistence demands associated with the Santarém culture expansion. As modern deforestation and agricultural plantations expand across the Amazon Basin, coupled with the intensification of drought severity driven by warming global temperatures, these data provide a detailed history of over four millennia of anthropogenic land use that progressively intensified, in the absence of large-scale deforestation, that has a lasting legacy on the composition of modern rainforests in the eastern Amazon. These data provide valuable new insights into the vital role that indigenous land management practices played in shaping modern ecosystems that can inform ecological benchmarks and future management efforts in the eastern Amazon.

**Methods**

**Palaeoecology methods. Regional study area.** To investigate coupled human–environment systems, we designed a multi-proxy approach integrating local (archaeological site/terrestrial palaeoecology) and regional (lake palaeoecology) spatial scales. We selected the Tapajós National Forest (FLONA), located on the eastern side of the white water Tapajós River, ~50 km south of Santarém (Pará state, Brazil), which forms part of the Creteaceous Alter do Chão Formation [1]. The climate is seasonally dry, inter-tropical humid with a distinct wet season between January and June. The mean annual rainfall ranges between 1,900 and 2,200 mm per year and the average annual temperatures are between 21 °C and 31 °C [2]. The vegetation is composed of dense terra firme hummock evergreen rainforest [3].

An understanding of the spatiotemporal nature of the pre-ADE subsistence strategies was gained by comparing radiocarbon-dated lake sediment core data from Lake Caranã (pollen, charcoal, geochemistry and magnetic susceptibility) with AMS-dated archaeobotanical soil profile data (phytoliths) from SDM1, which allowed the reconstruction of pre-Columbian land-use and subsistence strategies for the past 5,000 years. These records provided two distinct spatial scales: pollen and charcoal from the lake sediment core provided the watershed-scale (<10 m²) vegetation composition (Sugita [4]), whereas phytoliths, which are deposited in situ [5], represent local-scale (<1 m²) vegetation structure. The data were compared with the SDM1 botanical inventory data to evaluate the legacy of pre-Columbian land use on modern vegetation in the eastern Amazon.

**Palaeoecology site selection and core collection.** Lake Caranã (2° 50’ 08” S, 55° 02’ 33” W) is a flat-bottom lake located on the fluvial terrace on the eastern bank of the Rio Tapajós. Lake Caranã is located within a small closed basin and is separated from the main river channel (except during extreme flood events) by a depositional sand berm (200 m-long, ~3 m-tall) located on the northeastern edge of the lake. A 210-cm sediment core was collected using overturning drives from a Livingston drive rod piston core [6] with a modified Bolivia surface corer to collect the sediment in situ. Cores were transported back to the University of Exeter, UK, for cold storage. Lake Caranã was selected because it is located at the base of the Belterra Plateau, which is rich in archaeological sites and ADE soils, and today, receives limited sediment inputs from the Tapajós River. Thus, Lake Caranã is ideally located to reconstruct changes in human land use around the Belterra Plateau.

**Palaeoecological age-depth model.** The chronology for the Lake Caranã sediment core relies on six radiocarbon (14C) dates, 210Pb radiocarbon analysis of recent sedimentation and an age-depth model constructed in Bacon v2.2 (ref. 57) within a Geographic Information System (GIS) environment. Cores were scanned horizontally, end-to-end through the ring sensor. Magnetic susceptibility was conducted at 1-cm intervals using a Bartington ring sensor equipped with a 75-mm aperture. Magnetic susceptibility was measured to identify magnetic variation in the sediments. The magnetic susceptibility of sediments is reflective of the relative concentration of ferromagnetic (high positive magnetic susceptibility), paramagnetic (low positive magnetic susceptibility) and diamagnetic (weak negative magnetic susceptibility) minerals or materials. Typically, sediment derived from freshly eroded rock has a relatively high magnetic susceptibility, whereas sediments that are dominated by organic debris, evaporites or sediments that have undergone substantial diagenetic alteration have a low or even negative magnetic susceptibility. Sediment cores were scanned horizontally, end-to-end through the ring sensor. Magnetic susceptibility was conducted at 1-cm intervals using a Bartington ring sensor equipped with a 75-mm aperture.

**Loss on ignition.** Organic and carbonate sediment composition was determined by loss on ignition conducted at 4-cm intervals throughout the core. For each sample, 1 cm³ sediment was dried in an oven at 100 °C for 24 h. The samples underwent a series of 2-h burns in a muffle furnace at 550 °C and 1,000 °C to determine the relative percentage of the sample composed of organics and carbonates. The organic composition was determined by weight following standard methodologies [7].

**Lycopodium spores.** Lycopodium spores were used as a single species that is consumed by monkeys and assumed to be a proxy [8]. The spores are small, and their presence is indicative of the local monkey population. The spores were counted in a gridded petri dish at 0.25 cm² and used to determine the age of fire events [9]. The spores were counted in a petri dish at 0.25 cm² and used to determine the age of fire events [9].

**Magnetic susceptibility.** Magnetic susceptibility was determined by measuring the alpha decay of its daughter product 210Po as a proxy [10]. Sediment cores were scanned horizontally, end-to-end through the ring sensor. Magnetic susceptibility was conducted at 1-cm intervals using a Bartington ring sensor equipped with a 75-mm aperture.

**X-ray fluorescence.** X-ray fluorescence analysis was conducted using a portable X-ray fluorescence (XRF) Niton Thermo Scientific Niton ML10 GOLDD at the University of Exeter, UK, at a step size of 2,000 or 5,000 μm. A micro-X-ray beam focused through a flat capillary waveguide was used to irradiate samples to enable both X-ray radiography and X-ray fluorescence analysis. Data were acquired incrementally at 0.25-cm contiguous intervals by advancing the split core through the X-ray beam, and results were normalized using z-scores.

**Macrocharcoal.** The Lake Caranã sediment core was subsampled for macroscopic charcoal analysis at 0.5-cm intervals from 0 to 210 cm in depth. Samples were analyzed for the presence of charcoal particles greater than 125 μm using a Nikon Diaphot photomicroscope equipped with a Reticon 4900 scanning cooling method [11]. Subsampled material (1 cm³) was treated with 5% potassium hydroxide in a hot water bath for 15 min. The residue was sieved through a 125-μm sieve. Macroscopic charcoal (particles of >125 μm in minimum diameter) was counted in a grid cell petri dish at ×40 magnification on a dissecting microscope. Count samples were converted to charcoal influx (the number of charcoal particles cm⁻² yr⁻¹) and charcoal accumulation rates by dividing by the deposition time (yr cm⁻²). Charcoal influx data (particles cm⁻² yr⁻¹) were used as an indicator of fire severity (the amount of biomass consumed during a fire episode or a period of increased burning). A regime shift detection algorithm based on sequential t-tests was applied to determine the occurrence of significant shifts in charcoal influx data [12]. Shifts were detected in both the mean fluctuations and the variance of profile iterations identified radiocarbon ages Beta-469035 and Beta-469038 as potential outliers. Rather than omit these data points, they were retained and contributed to the uncertainty distribution of the model. For example, at depths of 1.00 ± 0.005 m and 1.5 ± 0.005 m, where a possible reversal occurs, the outliers allow for a greater range of age-depth iterations, which provide age estimations (3,562 ± 423 and 4,555 ± 514 cal bp, respectively) with higher uncertainties in comparison to the youngest part of the model in which the age profile distributions were narrower and showed more certainty.

**Analytic methods for standard pre-treatments and radiocarbon analysis.** Radiocarbon ages were calibrated (Supplementary Table 1) within Bacon using IntCal13 (ref. 13) and modelled using Student’s t-tests distributions with wide tails to negate the need of identifying and removing potential outliers in the age-depth model [14]. The use of Bacon and Bayesian statistics to reconstruct the accumulation history at Lake Caranã allowed us to include every radiocarbon date that was taken throughout the Lake Caranã core and develop robust estimations of age-depth uncertainty. Age-depth model mean accumulation rate priors in Bacon were calculated using the constant rate of supply model 59, which provided ten ages for the upper Lake Caranã sediment core data from Lake Caranã (pollen, charcoal, geochemistry and magnetic susceptibility) with AMS-dated archaeobotanical soil profile data (phytoliths) from SDM1, which allowed the reconstruction of pre-Columbian land-use and subsistence strategies for the past 5,000 years. These records provided two distinct spatial scales: pollen and charcoal from the lake sediment core provided the watershed-scale (<10 m²) vegetation composition (Sugita [4]), whereas phytoliths, which are deposited in situ [5], represent local-scale (<1 m²) vegetation structure. The data were compared with the SDM1 botanical inventory data to evaluate the legacy of pre-Columbian land use on modern vegetation in the eastern Amazon.

**Palaeoecological age-depth model.** The chronology for the Lake Caranã sediment core relies on six radiocarbon (14C) dates, 210Pb radiocarbon analysis of recent sedimentation and an age-depth model constructed in Bacon v2.2 (ref. 57) within R [6]. Ages for the upper Lake Caranã sediment core were modelled using 210Pb radiocarbon analyses following standard procedures [59]. Atmospheric fallout of 210Pb can be used to estimate the age of sedimentary sequences by measuring the rate of its decay across approximately six to nine half-lives, or 150–200 years. The addition of 210Pb dating was used in this study to develop a robust chronology for the most recent palaeoenvironmental changes, which also provides an important validation tool for the youngest part of the age-depth model that otherwise relies on radiocarbon analyses. Radiocarbon ages that are younger than ~250 cal bp contain large calibration uncertainties owing to a ~200-year plateau in the 14C chronology (acc.mean (accumulation mean) = 6.4). Ages for the upper Lake Caranã sediment core data from Lake Caranã (pollen, charcoal, geochemistry and magnetic susceptibility) with AMS-dated archaeobotanical soil profile data (phytoliths) from SDM1, which allowed the reconstruction of pre-Columbian land-use and subsistence strategies for the past 5,000 years. These records provided two distinct spatial scales: pollen and charcoal from the lake sediment core provided the watershed-scale (<10 m²) vegetation composition (Sugita [4]), whereas phytoliths, which are deposited in situ [5], represent local-scale (<1 m²) vegetation structure. The data were compared with the SDM1 botanical inventory data to evaluate the legacy of pre-Columbian land use on modern vegetation in the eastern Amazon.
standard terrestrial pollen counts. Large pollen grains (>53 μm) concentrated through the fine-sieving methodology were scanned for Z. mays and other crop taxa that produce large pollen, such as M. esculenta and I. batatas. The coarse fractions were counted to a standardized equivalent count of 2,000 (0.5g/cm²) grains (~3–4 slides). The pollen in the fine fractions was counted to the standard 300 terrestrial grains. Mauritia/Mauritella were counted and totaled separately owing to high concentrations. Larger non-crop pollen that was sieved into the coarse fraction (for example, Mauritia/Mauritella), was factored back into the total terrestrial pollen sums using abundance calculations from the single density distribution. This has the advantage of including the full range of probabilities associated with calibrated dates, instead of using single point estimates. SPDs were built in OxCal using the sum function and the IntCal13 calibration curve with an original data set of 85 radiocarbon dates from the Lower Tapajós. To accommodate oversampling of selected taxa that produce large pollen, such as I. batatas and C. esculenta, and to prevent the bin width from affecting the SPD shape, the SPD bin width did not affect the final shape of the SPD. This procedure is necessary because a sum of the calibrated dates assumes that observations are independent, whereas this is not the case when multiple dates were obtained for single sites or phases within an interval. Although the radiocarbon record is inherently biased by research (privileged dating of certain sites or periods) and taphonomic factors (greater preservation of charcoal towards more-recent periods), SPDs have been shown to be a reliable method to assess past population dynamics in relative terms, provided an adequate sample size and measures of chronometric hygiene, which were used here. The trends in the SPD for the Santarém region are confirmed by cultural changes that provide independent evidence of population dynamics: the initial increase after ~4,500 cal BP coincides with the appearance of ADE in the Tapajós, and the peak after ~1,250 cal BP corresponds to the development of the Santarém culture and the proliferation of ADE sites in the Belterra plateau.

**Modern vegetation methods.** Modern botanical survey. Three pairs of 0.25-ha plots (50 x 50 m) were sampled in ADE and non-ADE sites on the Belterra Plateau (Supplementary Fig. 6). The vegetation is classified as modern terra firme forests. All live trees, palms, and lianas with a diameter at breast height (~1.3 m above the ground) larger or equal to 10 cm were measured. Species were identified in the field, and the samples were collected and transferred to the herbaria Nova Xavantina Herbarium, Nova Xavantina and Mato Grosso for further identification. Botanical inventory data were grouped into cultivated edible plants (trees and palms) based on a revised list of domesticated plants from Clement, S. et al., 25 cultivated plants within the Americas—North, Central, and South America from Mansfield’s Encyclopedia of Agricultural and Horticultural Crops30—and other uncultivated trees (Supplementary Tables 5–7). The relative richness and the relative abundance of edible plants, palms and other tree species were calculated and presented in Fig. 2a. The relative abundance indicates the number of individuals of edible plants, edible palms or other plants divided by the total number of individuals found in the plot, and the relative richness is the number of edible plants, edible palms or other plants divided by the total number of species found in the plot. Bar charts for the frequency of edible plants, edible palms and other plants that occur in the vegetation plots are presented in Supplementary Fig. 7. Data generated or analysed during this study are included in this published article (see Supplementary Information).

**Reporting Summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** The botanical and archaeological source data used to support the findings of this study are published as Supplementary Information along with the repository pollen, charcoal and radiocarbon chemical data from the Lower Tapajós. All live trees, palms, and lianas have been made publicly available through Neotoma and the Latin American Pollen Database.

Received: 29 December 2017; Accepted: 15 June 2018; Published online: 23 July 2018

**References**

1. Clement, C. R. et al. The domestication of Amazonia before European conquest. *Proc. Biol. Sci.* **282**, 20150813 (2015).

2. Schaan, D. P. Sacred Geographies of Ancient Amazonian Historical Ecology of Social Complexity (Left Coast Press, Walnut Creek, CA, 2012).

3. Balle, W. *Cultural Forests of the Amazon*. A Historical Ecology of People and Their Landscapes (Univ. Alabama Press, Tuscaloosa, AL, 2013).

4. Roberts, P. H., Hunt, C., Arroyo-Kalin, M., Evans, D. & Boivin, N. The deep prehistory of global tropical forests and its relevance for modern conservation. *Nat. Plants* **3**, 17093 (2017).

5. de Souza, J. G. et al. Pre-Columbian earth-builders settled along the entire southern rim of the Amazon. Nat. Commun. **9**, 1125 (2018).

6. Irriarte, J. et al. Fire-free land use in pre-1492 Amazonian savannas. *Proc. Natl Acad. Sci. USA* **109**, 6473–6478 (2012).

7. Watling, J. et al. Impact of pre-Columbian ‘geoglyph’ builders on Amazonian forests. *Proc. Natl Acad. Sci. USA* **114**, 1868–1873 (2017).

8. Heckenberger, M. J. et al. Pre-Columbian urbanism, anthropogenic landscapes, and the future of the Amazon. *Science* **321**, 1214–1217 (2008).

9. Roosevelt, A. C. The Amazon and the Anthropocene: 13,000 years of human influence in a tropical rainforest. *Anthropocene* **4**, 69–87 (2013).

10. Erickson, C. L. in *The Handbook of South American Archaeology* (eds Silverman, H. & Isbell, W. H.) 157–183 (Springer, New York, NY, 2008).

11. Barlow, J., Gardner, T. A., Lees, A. C., Parry, L. & Peres, C. A. How pristine are ‘pristine’ tropical forests? An empirical perspective on the pre-Columbian human footprint in Amazonia and implications for contemporary conservation. *Biol. Conserv.* **151**, 45–49 (2012).
75. Holst, I., Moreno, J. E. & Piperno, D. R. Identification of teosinte, maize, and *Tripsacum* in Mesoamerica by using pollen, starch grains, and phytoliths. *Proc. Natl Acad. Sci. USA* **104**, 17608–17613 (2007).
76. Burn, M. J. & Mayle, F. E. Palynological differentiation between genera of the Moraceae family and implications for Amazonian palaeoecology. *Rev. Palaeobot. Palynol.* **149**, 187–201 (2008).
77. Dickau, R. et al. Differentiation of neotropical ecosystems by modern soil phytolith assemblages and its implications for palaeoenvironmental and archaeological reconstructions. *Rev. Palaeobot. Palynol.* **193**, 15–37 (2013).
78. Watling, J. & Iriarte, J. Phytoliths from the coastal savannas of French Guiana. *Quat. Int.* **287**, 162–180 (2013).
79. Iriarte, J. et al. Late Holocene neotropical agricultural landscapes: phytolith and stable carbon isotope analysis of raised fields from French Guianan coastal savannahs. *J. Archaeol. Sci.* **37**, 2984–2994 (2010).
80. Piperno, D. R. Identifying crop plants with phytoliths (and starch grains) in Central and South America: a review and an update of the evidence. *Quat. Int.* **193**, 146–159 (2009).
81. Iriarte, J. 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. *J. Archaeol. Sci.* **30**, 1085–1094 (2003).
82. Pearsall, D. M., Chandler-Ezell, K. & Chandler-Ezell, A. Identifying maize in neotropical sediments and soils using cob phytoliths. *J. Archaeol. Sci.* **30**, 611–627 (2003).
83. Piperno, D. R., Andres, T. C. & Stothert, K. E. Phytoliths in *Cucurbita* and other neotropical Cucurbitaceae and their occurrence in early archaeological sites from the lowland American tropics. *J. Archaeol. Sci.* **27**, 193–208 (2000).
84. Travassos, D. Dark Earth Plant Management in the Lower Tapajós. PhD thesis, Univ. Exeter (2018).
85. Downey, S. S., Haas, W. R. & Shennan, S. J. European Neolithic societies showed early warning signals of population collapse. *Proc. Natl Acad. Sci. USA* **113**, 9751–9756 (2016).
86. Goldberg, A., Mychajliw, A. M. & Hadly, E. A. Post-invasion demography of prehistoric humans in South America. *Nature* **532**, 232–235 (2016).
87. Shenna, S. et al. Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nat. Commun.* **4**, 2486 (2013).
88. Timpson, A. et al. Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. *J. Archaeol. Sci.* **52**, 549–557 (2014).
89. Zahid, H. J., Robinson, E. & Kelly, R. L. Agriculture, population growth, and statistical analysis of the radiocarbon record. *Proc. Natl Acad. Sci. USA* **113**, 931–935 (2016).
90. Ramsey, C. B. & Lee, S. Recent and planned developments of the program OxCal. *Radiocarbon* **55**, 720–730 (2013).
91. Alves, D. T. Ocupação Indígena na Foz do Rio Tapajós (1610 a.C.–1020 d.C.): Estudo do sítio Porto de Santarém (Novas Edições Acadêmicas, Saarbrücken, 2014).
92. Quinn, E. R. Excavating “Tapajós” Ceramics at Santarém: their Age and Archaeological Context. PhD thesis, Univ. Illinois at Chicago (2004).
93. Roosevelt, A. C. The development of prehistoric complex societies: Amazonia, a tropical forest. *Archeol. Pap. Am. Anthropol. Assoc.* **9**, 13–33 (1999).
94. National Register of Archaeological Sites (IPHAN, 2018).
95. WinklerPrins, A. M. G. A. & Aldrich, S. P. Locating Amazonian dark earths: creating an interactive GIS of known locations. *J. Lat. Am. Geogr.* **39**, 33–50 (2010).

Acknowledgements

Funding for this research was supported by the PAST (Pre-Columbian Amazon-Scale Transformations) European Research Council Consolidator Grant to J.I. (ERC_Cog 616179). Research was conducted under permit 01506.004836/2014-69 from the Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN) and ICMBio permit 106/14-FNT. We thank all residents of the Maguarí and Jamaraquá community for their hospitality and help.

Author contributions

J.I., S.Y.M. and D.S. designed the research. S.Y.M., J.I., D.A. and M.R. carried out the palaeoecological and archaeological fieldwork. E.A.d.O. carried out the botanical inventories. S.Y.M. carried out the pollen, charcoal, geochemistry and magnetic susceptibility analyses. D.A. carried out the analysis of the archaeological data. R.L.B. built the age-model chronology. J.G.d.S. compiled and analysed the archaeological radiocarbon dates. C.L. carried out the analysis of the modern vegetation and compiled the list of edible plants. S.Y.M. and J.I. led the writing of the paper with inputs from all other authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41477-018-0205-y.
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to S.Y.M.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Experimental design

1. Sample size

Describe how sample size was determined.

Data combined from 1 lake record (60 pollen samples and 417 charcoal samples), 3 soil profiles, and six modern vegetation plots (3 pairs from ADE and non-ADE soils).

Lake samples were processed, identified, and counted at the University of Exeter Archaeobotany Laboratory following standard procedures (see Methods and SI). We used the standard pollen count size of 300 grains per sample to capture pollen diversity and characterise major vegetation communities in the Neotropics (Gosling et al., 2009). An additional sieving step was added to concentrate crop pollen. Crop pollen was identified using taxonomic key and the reference collection housed at the University of Exeter (see SI methods). Total charcoal counts from 1cm² at every 0.5cm depth were counted.

Archaeological soil samples were processed for phytoliths, identified, and counted at the University of Exeter Archaeobotany Laboratory following standard procedures (Piperno, 2006). The standard count size of a minimum of 200 phytoliths were identified in each sample. In order to maximize the recovery of important phytoliths of different size classes, such as those that derive from the rinds of Cucurbita fruits (Bozarth, 1987; Piperno et al., 2000) and from leaves and cobs of maize (Iriarte, 2003; Pearsall et al., 2003; Piperno and Pearsall, 1993), soil sediments were separated by wet-sieving into silt (2-50 um) and sand (50 to 2000 um) fractions.

We set up three pairs of 0.25 ha plots (50 x 50 m) vegetation plots, three in Amazonian Dark Earths (ADE) and three in adjacent non-ADE locations. In each plot, we sampled all live trees with diameter at breast height \( \geq 10 \) cm (DBH) and at least 1.30 m tall. Botanical inventory data are available in SI. All statistics were calculated in R open source software using publicly available packages (e.g. vegan community ecology package) and detailed in the SI methods.

2. Data exclusions

Describe any data exclusions.

No data were excluded from this analysis.

Lake sediment cores were taken using a Livingston piston corer. Sediment samples showed no signs of bioturbation. This interpretation is supported by the 210Pb and 14C AMS age-depth model (see SI methods for details).

Standard methods, protocols, and taxon identification were used in pollen and phytolith analysis and detailed in the SI methods section. Three replicate soil profiles were analysed and three pairs of vegetation plots were used as replicates. All botanical inventory data are available in SI along with the updated plant list of edible plants (compiled from publicly available materials). Plants collected for identification were accessioned into the Herbarium in Manaus are publicly available. All statistics were calculated in R open source software using publicly available packages (e.g. vegan community ecology package) and detailed in the SI methods.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

Lake sediment cores were taken using a Livingston piston corer. Sediment samples showed no signs of bioturbation. This interpretation is supported by the 210Pb and 14C AMS age-depth model (see SI methods for details).

Standard methods, protocols, and taxon identification were used in pollen and phytolith analysis and detailed in the SI methods section. Three replicate soil profiles were analysed and three pairs of vegetation plots were used as replicates. All botanical inventory data are available in SI along with the updated plant list of edible plants (compiled from publicly available materials). Plants collected for identification were accessioned into the Herbarium in Manaus are publicly available. All statistics were calculated in R open source software using publicly available packages (e.g. vegan community ecology package) and detailed in the SI methods.

Archaeobotany. Samples were carefully collected by experience archaeologists using column sampling and targeted archaeological features from undisturbed archaeological contexts and soil profiles.

Palaeoecology. Four different lakes were cored in the region. Lake Carana was selected based on proximity to archaeological site and distance to river (decreased riverine transport).

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Archaeobotany. Samples were carefully collected by experience archaeologists using column sampling and targeted archaeological features from undisturbed archaeological contexts and soil profiles.

Palaeoecology. Four different lakes were cored in the region. Lake Carana was selected based on proximity to archaeological site and distance to river (decreased riverine transport).
5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Blinding was not relevant to our study.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

- The exact sample size \( n \) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- Test values indicating whether an effect is present
- Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on statistics for biologists for further resources and guidance.

Software

Describe the software used to analyze the data in this study.

- Tilia 2.0.37 was used for pollen analysis, C2 was used for phytolith analysis and R-Studio was used for statistical analysis, the Bacon R package was used in R to develop the age-depth model for lake sediment core.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). Nature Methods guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

- No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

- No antibodies were used.

10. Eukaryotic cell lines

- No eukaryotic cell lines were used.

  a. State the source of each eukaryotic cell line used.
  - No eukaryotic cells lines were used.

  b. Describe the method of cell line authentication used.
  - No eukaryotic cells lines were used.

  c. Report whether the cell lines were tested for mycoplasma contamination.
  - No eukaryotic cells lines were used.

  d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.
  - No commonly misidentified cell lines were used.
Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines.

11. Description of research animals
Provide all relevant details on animals and/or animal-derived materials used in the study.

No animals were used.

Policy information about studies involving human research participants; when reporting human research, follow the STROBE guidelines.

12. Description of human research participants
Describe the covariate-relevant population characteristics of the human research participants.

The study did not involve humans.