S1 File

Supporting discussions

Discussion A: Savannah classifications

African savannah ecosystems are defined by the irregular juxtaposition of woody plants on a continuous matrix of herbaceous groundcover [16]. Within the bounds of this broad definition, savannahs are subdivided on the basis of vegetation structure [35], particularly woody cover. Accordingly, we adopted the United Nations Educational, Scientific and Cultural Organization (UNESCO) classification system for savannah ecosystems (Fig B) [95]:

i. Forest: Over 70% woody cover (areal basis) with continuous crown canopy and occasional groundcover.

ii. Woodland: Between 40–70% woody cover with regular crown canopy and frequent groundcover.

iii. Wooded savanna: Between 10–40% woody cover with irregular crown canopy and common groundcover.

iv. Grassland: Less than 10% woody cover with abundant groundcover dominated by grasses.

In this study, we limit our consideration to savannah ecosystems with mean annual precipitation of less than 1200 mm [13, 96] because woody cover has a predictable response along precipitation gradients below this amount.

Discussion B: Contemporary plant leaves

We compiled both new and literature data [32, 33, 97-99] for leaf-wax n-alkanes of 139 contemporary plant leaves, which represent 82 distinctive species of (sub)tropical African savannah vegetation communities (S1 Table). Contemporary
plants were combined into dominant PFTs based on photosynthetic pathway and
growth habit [15]: C₃ woody plants (n = 42), C₃ forbs (n = 22), and C₄ grasses (n = 75). Leaf-wax n-alkane signatures within individual PFTs (Fig 2 and S1 Table) are reported for median values (\( \bar{x} \)) and the median absolute deviation (±MAD), which are robust measures of central tendency and dispersion, respectively [100].

**Discussion C: Surface sediments**

We compiled both new and literature data about C₂₇–C₃₃ (biomarker) n-alkanes in (sub)tropical Africa surface soils and terrestrial sedimentary particles (hereafter referred to as “surface sediments”) that include aeolian dust (n = 7), litter (n = 5) soil (n = 17), river and lake sediments (n = 3), and terrigenous component of ocean sediments [n = 8 (S2 Table)]. Literature data was compiled from earlier studies with \( \delta^{13}C \) values reported for odd-numbered C₂₇–C₃₃ n-alkanes derived from (sub)tropical African savannahs (c.f., Discussion A) because this range of homologues can establish uniquely parsimonious constraints on differential PFT contributions to surface sediments [32, 33] in spite of the inherent differences in wax production within and between dominant PFTs [34]. As an added benefit, this range can ameliorate most troublesome effects created by plants with variable carbon concentrating pathways (e.g., aquatic macrophytes) for surface-sediment LEWIS values as most (semi)aquatic vegetation synthesizes much lower concentrations of C₂₇–C₃₃ n-alkanes as compared to terrestrial higher vegetation [101].

**Discussion D: Satellite estimates of fractional woody cover**

We used MODerate-resolution Imaging Spectroradiometer (MODIS) Vegetation Continuous Fields (VCF) tree cover composite datasets averaged from
2000 to 2010 [66] to develop gridded 30-m resolution [102] estimates of fractional tree cover (\(\text{MODIS}_{\text{woody}}\)) for (sub)tropical African savannahs (Fig 1). Previous research has demonstrated close correspondence between satellite-estimated fractional tree cover and local higher-resolution estimates of fractional tree cover [51, 103] despite marked differences in spatial- and thematic scale.

Gridded fractional tree cover estimates were integrated from across defining geographic source-region [43, 52] to calculate more representative (“rescaled”) estimates of fractional tree cover. For soils specifically, gridded fractional tree cover estimates were averaged from across a radius of 60 m (~10,000 m\(^2\)) because soils incorporate organic matter within source-regions of 1000–10,000 m\(^2\) [104] and the estimates were validated against high-resolution images (<5 m) from GoogleEarth to constrain sub-grid variation [105]. Geomorphological features were constrained for source-regions of 10\(^3\)–10\(^4\) km\(^2\) [43, 106] to calculate rescaled fractional tree cover estimates within individual fluviolacustrine catchments [107-110].

Precise source-regions of dust and terrigenous components of ocean sediments are difficult to constrain [110-113], but backwind trajectories [111] and river discharge data [109, 114] were used to calculate representative rescaled fractional tree cover estimates. We calculated 1500-km backwind trajectories during seasons with prevailing offshore winds [115] with Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) models [116]. Calculations are based on Global Data Assimilation System (GDAS) archive data for an altitude of 1500 m (~850 hPa) because most dust blown offshore of (sub)tropical Africa occurs in winds between 1000–2000 m altitude [113]. Indeed, dust from major source areas in western (sub)tropical Africa is entrained by shallow (500–1500 m) trade wind layers [111, 115]. Large scale monsoon convergence of West African near-surface air masses
uplift fresh leaf lipids abraded from vicinal terrestrial (last 100–1000 km of landside wind trajectories [115]) landscapes and transport them offshore [117, 118], such that dust-delivered lipids encode integrative abundance-weighted biomarker signals of their source-region’s vegetation even if the mineral dust itself derives from a more remote area [115, 119, 120]. Aeolian organic matter also includes more minor contributions of lipids derived from soils and dessicated lake sediments, although these “old” lipids are usually quite similar as compared to aforementioned “fresh” lipids [115, 121] since average soil turnover rates in (sub)tropical Africa are short (~10–100 years) [47, 120].

Discussion E: Biomarker reconstructions of fractional woody cover

Previous studies reveal that biomarker n-alkane δ\(^{13}\)C values in surface sediments carry quantitative information about fractional tree cover when derived from savannahs [35, 122]. Because, on average, many plants in savannahs synthesize similar concentrations of n-C\(_{31}\) [34], we used the carbon isotopic composition of this n-alkane (δ\(^{13}\)C\(_{31}\)) to reconstruct fractional tree cover \(^{31}\text{f}_{\text{woody}}\) (Fig B) [35]:

\[ ^{31}\text{f}_{\text{woody}} = \sin(-1.8353 - 0.08538 \times \delta^{13}\text{C}_{31})^2 \]

These reconstructions are representative of landscape-integrated fractional tree cover across a sedimentary plant biomarker source-region, as weighted by production (i.e., concentration) and the reciprocal transport distance (i.e., autocorrelation) [123]. Although reconstructed fractional tree cover \(^{31}\text{f}_{\text{woody}}\) estimates and rescaled fractional tree cover \(^{\text{MODIS}}\text{f}_{\text{woody}}\) estimates do not have any particular intrinsic relationship, our results demonstrate a strong positive correlation [m = 0.983, b = 0.01; r = 0.962 (Fig 4B)] between reconstructed and rescaled fractional tree cover.
Biomarker \( n \)-alkanes, in particular \( n \)-C\(_{31} \), have \( \delta D \) values (e.g., \( \delta D_{31} \)) that are offset from, but highly correlated with, the \( \delta D \) value of their source-region rainfall (\( \delta D_{\text{water}} \)) [124]. Previous studies about African savannahs suggest that this offset between \( \delta D_{31} \) and \( \delta D_{\text{water}} \) values can be accounted for with \( \delta ^{13} C_{31} \) values [48]. In respective studies, \( \delta ^{13} C_{31} \) values were used to calculate offset (i.e., apparent fractionation) factors at the landscape scale \( \varepsilon_{\text{landscape}} \) (Fig E) by pairing \( 31f_{\text{woody}} \) estimates and the reciprocal fractional \( C_4 \) grass cover \( (31f_{\text{grass}} = 3.3 + 0.1 \times \delta ^{13} C_{31}) \) with literature values of the offset between \( \delta D_{31} \) values and the \( \delta D \) values of biosynthetic water sources (i.e., rainfall) for each of the corresponding plant functional types \( [C_4 \text{ grasses (−146‰); } C_3 \text{ forbs (−124‰); } C_3 \text{ woody plants (−109‰)}] \) (Fig B). Then, \( \varepsilon_{\text{landscape}} \) calculations were applied to ice-volume corrected \( \delta D_{31} \) values (Fig E) to reconstruct \( \delta D \) values of their source-region rainfall \( [31\delta D_{\text{water}} \text{ (Fig 5E)}] \).

**Discussion F:** Estimating plant species richness

Here, we adopt the contemporary plant species richness \( (S_{\text{source}}) \) estimates of African ecoregions [22] as a representative index of plant biodiversity since associated \( S_{\text{source}} \) estimates minimize sample standardizations or conjecture [21]. Indeed, the relative ease with which savannah vegetation is differentiated by PFT and taxonomy justify published \( S_{\text{source}} \) estimates as a measure of plant biodiversity\(^{119} \). In our study, plant biodiversity is defined by published \( S_{\text{source}} \) estimates in selected \( (\text{sub})\text{tropical African ecoregions (S2 Table). Under such circumstances, plant biodiversity patterns will be relative as opposed to absolute, and thus restricted to specific ecosystems. Since modern savannahs also may not be analogous to savannahs of the past, we consider down-core sediment LEWIS trends as**
representative of the relative changes as opposed to absolute differences in plant species richness.

Individual bioclimatic variables exert dynamic influences on species richness, which can differ with observation scale, across space and through time [23, 125]. However, numerous studies suggest that this scale dependence can be accounted for in savannahs with a single species-time-area relationship (STAR) [23, 126] because savannah vegetation communities are defined by plant–water–energy dynamics regardless of the scale used for observation [127]. For instance, contemporary plant species richness (PSR) estimates at the regional (100–10,000s km²), landscape (1–100s km²) and local level (0.01–1 km²) have a strong positive relationship in savannahs [73] and throughout (sub)tropical Africa [128]. More importantly, previous studies suggest that $S_{\text{source}}$ estimates, once rescaled, function as a reflection of local PSR [127, 128]. This is important for our work because it demonstrates the robustness of STAR for comparisons of PSR between disparate scales (e.g., 1 and 100 km⁻² yr⁻¹) when rescaled to a standardized time-by-area unit (e.g., 10 km⁻² yr⁻¹).

For our study, we rescaled $S_{\text{source}}$ estimates [22] using predictive (slope) regression models [125, 126, 129] constrained by published STAR terms and African ecoregion characteristics (S2 Table):

$$S_{\text{STAR}} = S_{\text{source}} \left\{ \left( \frac{A_{\text{source}}}{A_{\text{STAR}}} \right)^{z} \left( \frac{1}{T_{\text{STAR}}} \right)^{w} \left( \frac{A_{\text{source}}}{A_{\text{STAR}} T_{\text{STAR}}} \right)^{k} \right\}^{-1}$$

$S_{\text{STAR}}$ values represent the rescaled PSR (km⁻² yr⁻¹). $A_{\text{source}}$ values represent the contemporary geographic extent for a sample’s inceptive ecoregion (km²), and $A_{\text{STAR}}$ values are representative of the source-region extent [67] for sedimentary plant biomarkers [km³ (c.f., Discussion D and S2 Table)]. $T_{\text{STAR}}$ values represent the average accumulation duration in years [43, 47] for each surface-sediment type. Exponential terms represent the slope of the species–area relationship ($z$), the species–
time relationship ($w$), and time-by-area interaction ($u$). The slope of the species–area relationship ($SAR_s$) was calculated from integration area [125]:

$$z = 0.264 + 0.0396 \log(A_{STAR}) - 0.0021 \log(A_{STAR})^2$$

The alternative use of a uniform $z$ of 0.25 does not have a significant impact on regression slope and linear correlation strength when plotting partial linear residuals of $S_{STAR}$ against LEWIS values, each on arcsin($\sqrt{\text{MODIS}_{woody}}$) (Discussion I). The slope of the species–time relationship ($STR_w$) was assumed to equal 0.40 in all instances [126, 129], as was the slope ($u$) of the interaction term [126]. Moreover, since all $S_{STAR}$ values are derived from straightforward least-squares regression models, respective estimates should be semi-quantitative at least [130] and internally consistent [21] and therefore applicable to sedimentary records [131]. At length, $S_{STAR}$ values were compared to current Biogeographical Information System on African Plant Diversity (BISAP) data (0.25×0.25-degree resolution) for gross validation [132] and baseline species richness assessments of Lower Zambezi ecosystems [133].

**Discussion G: Potential drivers of high sediment LEWIS values**

We use a simplified three end-member model that has equivalent total leaf tissue $n$-alkane concentrations in each dominant PFT ($C_4$ grasses, $C_3$ woody plants and $C_3$ forbs) [15] to understand how mixing proportions can influence sediment LEWIS independent from integration scale. As modelled, $C_4$ grasses represent a stable end-member since most African species show consistent biomarker $n$-alkane $\delta^{13}C$ values of $-21.8\pm1.7\%o$ [$n = 293$ (S1 Table)] and higher relative abundances of $n$-$C_{33}$ (30±10%). Woody plants and forbs show a wider range of biomarker $n$-alkane $\delta^{13}C$ values $[-33.3\pm2.4\%o$ ($n = 142$) and $-34.4\pm2.4\%o$ ($n = 81$), respectively] because of the influences of (micro)climate conditions (e.g., water availability and light level).
on carbon isotopic fractionation among plants with C$_3$ photosynthesis [134], but both show much higher relative abundances of n-C$_{29}$ (~30±10%). Importantly, previous studies suggest that biomarker n-alkanes usually peak at n-C$_{31}$ (~40±10%) for dominant African savannah PFTs [34]. Further, our contemporary plant leaf data reveal that dominant African savannah PFTs also show near-identical LEWIS index values of ~1.9‰ (S1 Table). This consistent LEWIS index value in contemporary plant leaves creates a conceptual basis for interpretations of LEWIS in soils or sediments vis-à-vis mixing proportion models with finite unimodal distributions [5, 127].

Isotopic mass-balance calculations indicate n-C$_{31}$ can be used to reconstruct PFT distribution abundance (i.e., cover) [35] because of its comparable concentrations in dominant African savannah PFTs and the characteristic relationship of δ$^{13}$C$_{31}$ values against fractional tree cover (c.f., Fig 3) and C$_4$ grass cover. However, δ$^{13}$C$_{29}$ and δ$^{13}$C$_{33}$ values are skewed toward C$_3$ plants and C$_4$ grasses, respectively [122]. As a consequence, end-member abundance-weighted δ$^{13}$C$_{29}$ and δ$^{13}$C$_{33}$ values show differing quadratic (unimodal) trajectories against δ$^{13}$C$_{31}$ values and demonstrate maximum isotopic difference at δ$^{13}$C$_{31}$ values indicative of low variance (i.e., maximum evenness) in dominant PFT distribution abundances: $f_{\text{woody}}, f_{\text{grass}}$ and $f_{\text{forb}}$ estimates of 0.35, 0.40 and 0.25, respectively.

Precipitation exerts a consistently positive influence on apparent $^{13}$C fractionation in contemporary plants [134]. This influence could lead to additional LEWIS differences of ~2‰ throughout the lower Zambezi, assuming precipitation averages range about 500–1500 mm within its catchment [137]. Altitude, in contrast, has a negative influence on contemporary plant $^{13}$C fractionation [134] that translates into a maximal LEWIS difference of ~2‰ throughout the lower Zambezi (<2000
Considering precipitation correlates with altitude in southeast Africa [135, 138], their net influences on sediment LEWIS balance out. Atmospheric carbon dioxide $\delta^{13}$C values decreased $<0.5\%$ between about 25 kya and today (pre-industrial) [139], and furthermore do not develop parallel to core-sediment LEWIS. However, rising $p$CO$_2$ mirrors core-sediment LEWIS throughout this interval (Fig 5D). Although the specific mechanism(s) for such a relationship are debatable, we exclude any photosynthetic influences caused by $p$CO$_2$ because such changes would have a strictly positive influence on apparent $^{13}$C fractionation of C$_3$ plants [140] and have a negligible negative influence on coeval C$_4$ plants [70]. Consequently, $p$CO$_2$ increases would lead to attenuated, low sediment LEWIS during periods with lower atmospheric carbon dioxide concentrations (e.g., glacial–deglacial transition) [141], and therefore cannot be a cause of the dramatic declines in core-sediment LEWIS between about 19.5–11.7 kya.

**Discussion H**: Down-core data integration

We integrated the $n$-alkane datasets of Schefuß et al. [39] and Wang et al. [40] for marine cores recovered from adjacent locations off the mouth of the Zambezi River (Fig 1):

GeoB9307-3 (18°34.0’S, 37°22.9’E; 542 m water depth)

GIK16160-3 (18°14.5’S, 37°52.1’E; 1339 m water depth)

Both cores were retrieved from zones with high (de)glacial–Holocene sedimentation of Zambezi-derived terrigenous material [114, 142]. The chronologies of both cores were previously established based on accelerator mass spectrometry (AMS) $^{14}$C dates on mixed foraminifera, which were converted to calendar ages using CALIB 6.0 and
Marine09 calibration curve and a 405-yr reservoir age correction. Therefore, sample age-depth interpolations should be comparable between cores.

There are remarkably parallel trends in down-core sediment C_{27}–C_{33} n-alkane δ^{13}C and δD records (Figs D–E). However, earlier studies [40, 68] note a difference in respective molecular isotopic records amid Heinrich Stadial 1 (c.f., Fig 5), which was attributed [142, 143] to a stronger influence of hinterland tributaries at GIK16160 than at GeoB9307 (Fig F). This difference is absent in our core-sediment LEWIS record, which we consider as meaning plant biomarkers were discharged from rivers in a single ecoregion or in multiple ecoregions with similar source-vegetation communities (e.g., eastern, southern, and Zambezian miombo woodlands) [144]. This consideration is consistent with the similar species in ecoregions of the so-called Zambezian domain [144, 145], despite notable differences in estimated tree cover and leaf biomass [103, 146]. Source-vegetation communities derived from a single or similar ecoregions is also consistent with river-dominated transmission of Zambezian-derived terrigenous matter (e.g., plant biomarkers and detrital lithogenic clays [142, 147] to coastal Mozambique margin sites [39, 148] because near all the rivers that discharge into coastal Mozambique margin waters flow through Zambezian miombo woodlands [133] within their final 100s km overland flow [64] such that local, intermittent differences in sub-basin sources [148] would have a nominal influence on interpretations of down-core sediment LEWIS due to transport mixing [142] despite more uncertain effects on individual biomarker n-alkane δ^{13}C records.

Previous studies suggest that Zambezi river suspension contributed ~40-80% of the downcore clay particles at Mozambique margin sites during glacial termination, about 20–10 kya [142, 143]. Although changes in inferred Zambezi clay particle contributions show some similarities with core-sediment LEWIS trends amid this
interval (Fig F), there is an inverse relationship such that high Zambezi clay particle contributions occur when core-sediment LEWIS has much lower values. This relationship is counterintuitive, since one would theoretically presume core-sediment LEWIS featured higher values when derived from multiple clay provenances. We suggest that differences in clay provenance would have a nominal influence on core-sediment LEWIS because of the widespread Zambezi floristic affinity present throughout the southeast African region [95], which features a shared Quaternary phylogenetic and biogeographical history [59]. This suggestion is supported by parallels in core-sediment LEWIS trends and the coinciding palynological diversity given at Lake Malawi (Fig 5E), which would be unaffected by perturbations in clay provenance at GeoB9307-3 and GIK16160-3 [114, 142]. Although both records could be influenced by geographic changes in catchment basin accumulation area [114, 148], rescaling (power) relationships in sedimentary processes [43, 52] and during proxy propagation in marine and fluviolacustrine systems [149] make simultaneous changes improbable.

Explicit terrestrial (air) temperature estimates during glacial termination and the mid-Holocene are scant for southeast Africa [150, 151], but there is a strong parallel between mean annual lower Zambezi air temperature and local Indian Ocean sea-surface temperature (SST) [39, 152]. Therefore, we used Mozambique Channel SSTs reconstructed from alkenones (SST_{UK}) [94] to reconstruct Lower Zambezi temperature changes (Fig 5D). Although these reconstructed temperature changes are relative, our multivariate regression models are robust to differences in absolute value and their explanatory power will be unchanged [100].

**Discussion I**: Regression analyses
We assessed bivariate relationships between estimated fractional tree cover ($^{\text{MODIS}} f_{\text{woody}}$) and biomarker $n$-alkane signatures using ordinary (partial) least-squares regression with the statistical $R$ (http://CRAN.R-project.org) computing packages named \{lmer4\}, \{mgcv\} and \{vegan\}. Regressions use arcsin($^{\text{MODIS}} f_{\text{woody}}$) – a statistical transformation used to improve linearity in ecological datasets [100] – as the explanatory (predictor) variable. Bivariate regression models were used to assess secondary predictor influences on surface-sediment LEWIS (S3 Table) by plotting partial (linear) residuals of LEWIS from its regression on arcsin($^{\text{MODIS}} f_{\text{woody}}$) against those of $S_{\text{source}}$ and $S_{\text{STAR}}$ residuals on arcsin[$^{\text{MODIS}} f_{\text{woody}}$] (Fig C). The strong parallel, linear relationships of the corresponding partial residual plots emphasize the significance of the relationship shared between LEWIS and PSR in spite of differences in sedimentary properties and time–area integration scale. Even so, “sediment type” (S2 Table) has significant higher-order explanatory power ($R_1^2 = 0.203–0.246$) as a random variable (slopes and intercept) during predictive multilevel hierarchical linear regression analyses [53] of $S_{\text{source}}$ and $S_{\text{STAR}}$ and their residuals on arcsin($^{\text{MODIS}} f_{\text{woody}}$) that highlight the significance of deposition dynamics (e.g., proxy propagation rates and turnover [149]) independent from simple species-time-area relationships [126] and fractional tree cover. It is important to understand that this significance does not confound the interpretation of down-core records unless a record features incommensurate sediment types.

We resampled down-core records [LEWIS, $\delta^{13}C_{31}$ ($^{31}f_{\text{woody}}$), $p$CO$_2$, SST$_{UK}$, and $^{31}\delta$D$_{\text{water}}$] at binned 250-yr steps to create uniform interpolated time-series, and then compared them with Fourier cross-correlation. Considering $p$CO$_2$ had the highest linear correlation coefficient ($r = -0.940$) with LEWIS during single-factor analyses, all further multiple regression models were constrained to include $p$CO$_2$. This
common constraint befits observations of the influence $pCO_2$ has on modern savannah vegetation communities [153] and the inferred importance of $pCO_2$ on (sub)tropical African succession since at least 25 kya [68]. Since carbon dioxide concentrations can correlate with other important bioclimatic variables, we assessed the independent influence of these secondary predictors on core-sediment LEWIS by partial (multivariate) regression models, which account for covariance among predictor variables. Only partial regression models with SST$_{UK}$ and $^{31}\delta D_{water}$ as secondary predictors showed higher explanatory power as compared to only $pCO_2$ for variation in core-sediment LEWIS. The combination of $pCO_2$, SST$_{UK}$, and $^{31}\delta D_{water}$ together account for about 90% of the variance shown in core-sediment LEWIS between 25 kya and the Common Era ($r = 0.947$).

**Discussion J: Zambezi River catchment**

The Zambezi River originates in western Zambia (1450 m.a.s.l.) and flows almost 3000 km southeast before it discharges into the western Indian Ocean (Mozambique Channel). In conjunction with an extensive network of tributaries, the Zambezi River catchment drains a cumulative area of 1,570,000 km$^2$ [154], which is separated by geomorphic characteristics into three sub-catchments: the Upper, Middle, and Lower [155]. Here, our discussion will focus on the distal Lower Zambezi sub-catchment (Fig 1B) because it comprises the endmost ~500 km of the main watercourse [64], and exerts a major influence on discharge material to near-shore sediments of the coastal Mozambique Channel [142, 143].

The majority of Lower Zambezi discharge is derived from one of two main rivers – the Shire and Luangwa – and hydrodynamic models suggest there is nominal transfer of Upper and Middle sub-basin sediments into downriver reaches of the
Lower Zambezi [156, 157]. Although earlier studies suggest that Cyperaceae could have caused increased \( \delta^{13}C_{31} \) values in discharged Zambezian organic matter amid deglaciation (Fig 5C) [39], C\(_4\) sedges (e.g., papyrus) likely give a nominal influence on our interpretations of LEWIS because dissolved organic carbon derived from macrophytes is usually quite low in rivers as compared to coeval terrestrial taxa, and – based on modern observations and historical data – decreases during periods of drier conditions [158]. Thus, although there will be some uncertainties in absolute reconstructed biodiversity patterns, our approach should be internally consistent and likewise captures systematic changes in ecoregion estimates of both raw and (re)scaled taxonomic richness estimates among phytozones of (sub)tropical Africa (Fig 4) [22].
Supporting tables

S1 Table: Median and median absolute deviation (MAD) values of C_{27}–C_{33} n-alkane δ^{13}C data and LEWIS values in contemporary plant leaves [32, 33, 97-99] in three overarching plant functional types (PFTs) for savannahs [15]: C_{3} woody plants (n = 42), C_{3} forbs (n = 22), and C_{4} grasses (n = 75). Important taxonomic distinctions within each PFT are also shown. In accordance with convention, average chain length (ACL) and weighted-mean average δ^{13}C (δ^{13}C_{WMA}) were characterized:

\[
\text{ACL} = \sum \left\{ \frac{x c_x}{c_x} \right\}
\]
\[
\delta^{13}C_{WMA} = \sum \left\{ \frac{\delta^{13}c_x c_x}{c_x} \right\}
\]

Nomenclature refers to the concentration (C_x) or composition (δ^{13}C_x) of an n-alkane homologue with x carbons.

S2 Table: Surface-sediment locations alongside characteristics of their respective source (eco)regions, surface material terms, species–time–area relationship variables, measured C_{27}–C_{33} n-alkane δ^{13}C values, LEWIS values, and fractional tree cover (\textit{MODIS}_f_{\text{woody}} and \textit{31}_f_{\text{woody}}) estimates.

The composition, age and distribution of terrestrial biomarkers in soils or sediments are entwined with the source(s) and transformation of these compounds during progressive source-to-sink transmission [43, 106, 121]. This in turn is influenced by particle mobilization and transport dynamics [149]. Although sedimentary particles are often defined by physicochemical features (e.g., grain size or roundness), we use more mechanistic definitions vis-à-vis integration area and transport histories to define seven sediment types (Discussion F): dust, litter, soil, lake, river and marine sediments with abundant terrigenous material from riverine or wind-blown sources.

We use ecoregion names as defined by global WWF terminology (http://www.worldwildlife.org/science/data/terreco.cfm). Soil and terrestrial-derived sediment (“surface sediments”) types were assigned tentative surface integration areas (A_{\text{STAR}}) and timescales of formation [T_{\text{STAR}} (Discussion F)]. We rescaled S_{\text{source}} estimates using predictive models of the species-time-area relationship (Discussion F) for modern savannahs. The slope (z) of the species–area relationship (SAR_z) was calculated from integration area [125]:
The slope ($w$) of the species–time relationship (STR$_w$) was assumed to equal 0.40 in all instances [129], as was the slope ($u$) of the interaction term [126], which is essential for accurate rescaled biodiversity predictions [69].

**S3 Table**: Measured biomarker $n$-alkane $\delta^{13}$C values and LEWIS values derived from data reported by Schefuß et al. [39] and Wang et al. [40] for sediment cores recovered from off the Zambezi River mouth (Discussion H).
Supporting figures

**Fig. A**: Visual (field) counts of plant species richness ($S_{\text{field}}$) against (A) moderate-resolution imaging spectroradiometer estimates of fractional woody cover ($^{\text{MODIS}}f_{\text{woody}}$) at 30-m resolution [102] over 250 m$^2$ transect plots in the lower Zambezi sub-catchment [144], and (B) counts of the corresponding plant functional richness ($F_{\text{field}}$) [159] for identical transect plots [144]. Associated $F_{\text{field}}$ counts are derived from empirical distinctions apparent between relevant functional traits in co-occurring plants (e.g., growth form and habit) [159]. Larger circle sizes and darker shading (green) both represent increased fractional tree cover (c.f., Fig 1). Blue shaded bounds indicate empirical 90% confidence intervals as calculated from a Monte Carlo method [92]. Asymptotic significance ($p$-value) is less than 0.0001 for all the relationships shown.
Fig. B: Schematic depiction of how individual biomarker $n$-alkane $\delta^{13}$C values are related to ecosystem structure and landscape deuterium offset factors ($\varepsilon_{\text{landscape}}$) as functions of fractional tree cover (c.f., Discussion E). A, Sediment $\delta^{13}$C$_{31}$ values were used to reconstruct the cover of three overarching plant functional types (PFTs): C$_4$ grasses, C$_3$ forbs, and C$_3$ woody plants. Fractional tree cover ($^{31}f_{\text{woody}}$) was reconstructed from the equation [35]:

$$^{31}f_{\text{woody}} = \sin(-1.8353 - 0.08538 \times \delta^{13}\text{C}_{31})^2$$

The fraction of C$_3$ forbs was assumed to equal the difference between $^{31}f_{\text{woody}}$ and reconstructed C$_4$ grass cover ($3.3 + 0.1 \times \delta^{13}\text{C}_{31}$ [35]). B, Fractional tree cover reconstructions were related to savannah ecosystem structure with United Nations Educational, Scientific, and Cultural (UNESCO) terminology [95]. C, At length, biomarker $n$-alkane $\delta^{13}$C values (i.e., $\delta^{13}$C$_{31}$) were used to calculate $\varepsilon_{\text{landscape}}$ values by pairing plant functional type reconstructions with literature values of the characteristic offset between $\delta$D$_{31}$ values and the $\delta$D values of biosynthetic water sources in each PFT (i.e., isotopic mass-balance) [48].
Fig. C: Partial (linear) residual plots of surface-sediment LEWIS against (A) the contemporary plant species richness estimates of African ecoregions ($S_{source}$) and (B) time–area rescaled $S_{source}$ estimates [$S_{STAR}$ (S2 Table)] as modeled by predictive species–time–area relationships (Discussion F). Residuals were calculated from respective single-factor regression models on $\arcsin(\sqrt{\text{MODIS f}_{\text{woody}}})$ to account for their common covariance with fractional tree cover. Larger circle sizes and darker shading (green) both represent increased fractional tree cover (c.f., Fig 1). Blue shaded bounds indicate empirical 90% confidence intervals as calculated from a Monte Carlo method [92]. Asymptotic significance ($p$-value) is less than 0.0001 for all the relationships shown.
Fig. D: Core-sediment δ¹³C values of biomarker n-alkanes: A, n-C₂₉ (δ¹³C₂₉); B, n-C₃₁ (δ¹³C₃₁); and C, n-C₃₃ (δ¹³C₃₃) for data compiled from Schefuß et al. [39] (filled circles) and Wang et al. [40] (open circles). Because n-C₂₇ data were sometimes unavailable, we corrected for δ¹³C₂₇ values with a 2nd degree polynomial function ($y = 0.056x^2 + 0.530x + 0.749$) based on relationships apparent for down-core sediment δ¹³C₂₉ and δ¹³C₃₁ values of in Wang et al. [40]. Core-sediment LEWIS values (i.e., C₂₇–C₃₃ n-alkanes) and the corrected C₂₉–C₃₃ n-alkane values show a very strong linear relationship ($m = 0.954; b = 0.076; r = 0.996; p$-value < 0.0001). Abbreviations are written as: Younger Dryas (YD); Heinrich Stadial 1 (H1); Bolling-Allerød (BA); Last Glacial Maximum (LGM).
Fig. E: Downcore records of hydroclimate change in southeast Africa over the last 25 kya. A, Down-core sediment δD values of \( n\)-C\(_{31} \) (\( \delta D_{31} \)) from Schefuß et al. [39] (filled circles) and Wang et al. [40] (open circles). B, Core-sediment \( \delta^{13}C_{31} \) values were used to calculate apparent landscape fractionation factors (\( \varepsilon_{\text{landscape}} \)) by paired fractional tree cover reconstructions (\( 31_{\text{woody}} \)) and the reciprocal fraction of C\(_3\) forb and C\(_4\) grass cover [35] with literature values of the offset between measured \( \delta^{13}C_{31} \) values and the δD values of biosynthetic water sources in dominant PFTs (Discussion E). Bold lines show a continuous averaged 5-pt timeseries. Abbreviations are written as: Younger Dryas (YD); Heinrich Stadial 1 (H1); Bolling-Allerød (BA); Last Glacial Maximum (LGM).
Fig. F: Downcore records of environmental change in southeast Africa over the last 25 kya.

A, Core-sediment LEWIS (Discussion H) for correlative marine cores recovered from off the Zambezi River mouth [GeoB9307 (filled circles) [39]; GIK16160 (open circles) [40]], which reflect the changing plant species richness of lower Zambezi vegetation communities [39]. Shaded bold lines show a combined 250-yr Gaussian-smoothed time-series. Shaded bounds indicate empirical 90% confidence intervals as calculated from a Monte Carlo method [92].

B, Downcore records of δ¹³C₃¹ values from at GeoB9307 (filled circles) [39] and GIK16160 (open circles) [40], which are indicative of C₃/C₄ plant functional type dominance [39]. Associated δ¹³C₃¹ values were also used to reconstruct fractional tree cover [³¹f_woody (dashed lines)] [35] through time.

C, Neodymium isotopic signatures (εNd) of the clay fraction of sediments recovered from near GIK16160-3 [142] that document changes in sediment provenance. Higher values indicate increased contribution of clays from the Zambezi River, and lower values indicate more northern river (e.g., Licungo) clay contributions.

D,
Normalized (relative) contributions of total remnant magnetization (SIRM) from end-member acquisition curves on magnetic mineral concentrations [148]. End-member 1 (EM1) derives from coastal regions of the Zambezi catchment, but EM2 derives from the hinterland.
Supporting references (N.B., both main [1-94] and supporting [95-159] references are included herein for comprehensiveness)


49. Kelly CK, Woodward FI. Ecological correlates of carbon isotope composition of leaves: a comparative analysis testing for the effects of temperature, CO2 and O2


67. Ryan CM, Williams MM, Hill TC, Grace J, Woodhouse IH. Assessing the phenology of southern tropical Africa: a comparison of hemispherical photography, scatterometry,

68. Khon VC, Wang YV, Krebs-Kanzow U, Kaplan JO, Schneider RR, Schneider B. Climate and CO2 effects on the vegetation of southern tropical Africa over the last 37,000 years. Earth Planet Sc Lett. 2014;403(0):407-17. doi: http://dx.doi.org/10.1016/j.epsl.2014.06.043.


70. Schubert BA, Jahren AH. Global increase in plant carbon isotope fractionation following the Last Glacial Maximum caused by increase in atmospheric pCO2. Geology. 2015;43(5):435-8.


142. van der Lubbe JJL, Frank M, Tjalllingi R, Schneider RR. Neodymium isotope constraints on provenance, dispersal, and climate-driven supply of Zambezi sediments along the Mozambique Margin during the past ~45,000 years. Geochem Geophys Geosyst. 2015;17:181-98.


